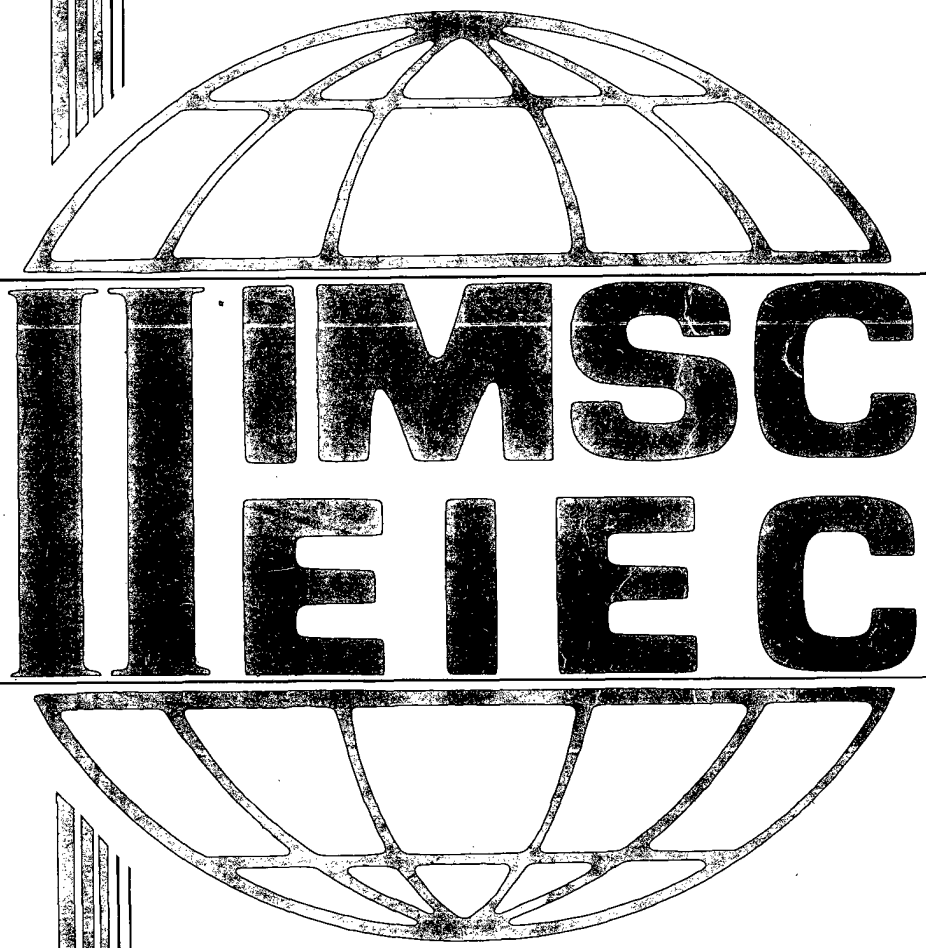
II INTERNATIONAL MEETING ON STATISTICAL CLIMATOLOGY



II Encontro Internacional de Estatística Aplicada à Climatologia

Sep. 26-30 * 1983
Lisboa * PORTUGAL

INSTITUTO NACIONAL DE METEOROLOGIA E GEOFÍSICA

II INTERNACIONAL MEETING ON STATISTICAL CLIMATOLOGY



Primary sponsors:

World Meteorological Organization
U.S. National Science Foundation
U.S. Office of Naval Research

Co-sponsors:

Academy of Sciences of Lisbon (Portugal),
American Meteorological Society,
American Statistical Association (Section on Physical and Engineering Sciences),
Bernoulli Society for Mathematical Statistics and Probability (International Statistical Institute),
Faculty of Sciences of the University of Lisbon (Department of Statistics, Operations Research and Computation),
Meteorological Society of Japan
and
National Institute of Scientific Research (Portugal).

Host Organization:

National Institute of Meteorology and Geophysics

Sep. 26-30 * 1983
Lisboa * PORTUGAL

INSTITUTO NACIONAL DE METEOROLOGIA E GEOFÍSICA
Rua C Aeroporto 1700 LISBOA — PORTUGAL

LIBRARY
ATMOSPHERIC ENVIRONMENT SERVICE
4905 DUFFERIN ST.
DOWNSVIEW, ONTARIO
CANADA M3H 5T4

SEP 3 1987

NOTE FROM THE SECRETARIAT

Some contributions to the II International Meeting of Statistical Climatology, including some lead papers, were not received in time. The Secretariat tried to include the maximum of papers for each session and this is the reason why the pages of this book could not be numbered. Instead, the papers received until 20 July 1983 are numbered in the order of presentation within a session and each page of each paper has received a number composed as follows:
Session A, Paper B, Page C is numbered A.B.C
(See the Table of contents).
At the Authors index only the numbers of session and paper are indicated.

FOREWORD

This preprint volume contains the papers to be presented at the Second International Meeting on Statistical Climatology, which will be held near Lisbon, Portugal from 26-30 September 1983 (the meeting frequently is referred to as "Lisbon-1983"). The First International Meeting on Statistical Climatology was held in Hachioji, Japan from 29 November - 1 December 1979. Selected papers presented at the meeting in Japan appeared in a book entitled Statistical Climatology published by Elsevier in 1980.

Statistical methods have played an important role in climate research, in the analysis of meteorological and climatological data, and in the application of climate information. The areas in which statistical methods currently are making significant contributions include:

(1) analysis and modeling of climatological data;
(2) spatial and temporal variability in meteorological elements on climatological time-scales;
(3) climate prediction;
(4) detection of changes in real or simulated climates; and
(5) climate impacts and applications of climate information.

The primary purpose of Lisbon-1983 is to bring together climatologists, meteorologists, statisticians, and others to describe and discuss recent work in these areas and to become more familiar with both the nature of the problems of interest and the statistical methodology available to attack them.

The scientific program for Lisbon-1983 has been arranged by an international organizing committee consisting of the following individuals: Dr. A. H. Murphy, Chairman; Dr. R. Czelnai; Dr. O. M. Essenwanger; Dr. L. A. Mendes Victor; Dr. E. L. Scott; Dr. R. Sneyers; Dr. E. Suzuki; and Dr. E. J. Wegman. The program for the meeting has been divided into seventeen sessions on fourteen different topics. One or two lead speakers have been invited to introduce each topic. The program in its current form consists of 14 lead papers, 71 contributed papers, and 30 reserve papers. The participants will represent more than thirty countries from the six major continents.

The primary sponsors of the meeting are the World Meteorological Organization (WMO), the U.S. National Science Foundation (NSF), and the U.S. Office of Naval Research (ONR). Funds provided by these organizations were used primarily to support the participation of lead speakers. Cosponsors of Lisbon-1983 include the Academy of Sciences of Lisbon (Portugal), the American Meteorological Society, the American Statistical Association (Section on Physical and Engineering Sciences), the Bernoulli Society for Mathematical Statistics and Probability (a section of the International Statistical Institute), the Faculty of Sciences of the University of Lisbon (Department of Statistics, Operations Research, and Computation), the Meteorological Society of Japan, and the National Institute of Scientific Research (Portugal).

The organizing committee would like to express its sincere appreciation to the host organization, the Instituto Nacional de Meteorologia e Geofisica (INMG) (a cosponsor of Lisbon-1983), for its kind and generous hospitality and for its many other contributions to the overall success of the meeting. Dr. Luis A. Mendes Victor is the Director-General of INMG. Special thanks are extended to the Secretariat of Lisbon-1983, under the able leadership of Mr. Anthimio Azevedo, for its efforts in connection with local arrangements and related activities. Dr. Mendes Victor and Mr. Azevedo also arranged for the compilation and publication of this preprint volume.

Finally, I would like to express the sincere wish that all participants will derive the greatest possible satisfaction and benefit from the meeting and from the most pleasant surroundings in Portugal.

June 20, 1983

Allan H. Murphy
Chairman, Organizing Committee

TABLE OF CONTENTS

		Pages
Foreword		3
Authors index		19

SESSION 1: DESIGN OF CLIMATOLOGICAL NETWORKS

Chairmen: R. Czelnai and L.A. Mendes Victor

1.1	Statistical approach to climatological network optimization - G. Der Megreditchian (Lead Paper)	*
1.2	Regionalization concept in meteorological network planning for Saudi Arabia - M.A. Ukayli, T. Husain, and H.U. Khan	**
1.3	Statistical determination of meteorology districts: a computer-based technique - B.A. Gornish	**
1.4	A quantitative basis for a solar radiation network evaluation - D.E. Petzold	1.4.1-5

SESSION 2: PROBABILISTIC AND STATISTICAL MODELS OF CLIMATOLOGICAL DATA

Chairmen: K. Cehak and E.J. Wegman

2.1	Modeling climatology of areal coverage - I.I. Gringorten (Lead Paper)	2.1.1-8
2.2	Fitting generalized lambda distribution to climatological data - A. Ozturk and A.M. Abouammoh	2.2.1-6

* - Manuscript included in back of book

** - Manuscript not available

2.3	Modeling visibility for data void regions - S.J. Bean and P.N. Somerville	2.3.1-4
2.4	Analysis of observed wind directions by statistics of directional data - A. Hiroh and M. Itoh	2.4.1-6
2.5	Application of regionalized variable analysis to climatic data - K.Y. Amegee and R.H. Cuenca	2.5.1-8
2.6	Fitting of gamma family distributions to wind speed time series: selection of the minimum number of parameters - F.R. Ferrão and L.E.V. Pessanha (Reserve)	*
2.7	Aspects of oceanic whitecap coverage dependence on wind speed: heteroscedasticity in the data, and the estimation of the Beaufort velocity - I.G. O'Muircheartaigh and E.C. Monahan (Reserve)	2.7.1-4

SESSION 3: PROBABILISTIC AND STATISTICAL MODELS OF PRECIPITATION DATA

Chairmen: Rodriguez-Iturbe and R. Sneyers

3.1	Climatic characteristics of monsoon rainfall over India - S.V. Singh (Lead Paper)	**
3.2	A general stochastic model of the precipitation process with applications to integrated precipitation, maximum daily precipitation, etc. - H. Alexandersson	3.2.1-8
3.3	Exploratory analysis of precipitation events in western Oregon - B.G. Brown, R.W. Katz, and A.H. Murphy	3.3.1-6
3.4	A statistical model for precipitation data - with application to interpolation problems - P. Allerup and H. Madsen	3.4.1-8

* - Manuscript included in back of book

** - Manuscript not available

3.5	Study of the variability in space and time of Upper Volta rainfall - T.A.B. Snijders (Reserve)	3.5.1-7
3.6	The standardized bivariate lognormal probability distribution simplifies the storm-flood relationship - L.A.V. Hiemstra (Reserve)	3.6.1-2
3.7	Five-day periods (pentads) maximum rainfall: probability table for Pelotas, RS, Brazil - J. Baptista da Silva (Reserve)	3.7.1-5
3.8	Annual rainfall variability in the northeast, Brazil - J. Ellis and A. de Seixas Valença (Reserve)	3.8.1-4

SESSION 4: MULTIVARIATE ANALYSIS OF METEOROLOGICAL AND CLIMATOLOGICAL DATA I

Chairmen: K.H. Bergman and I.T. Jolliffe

4.1	An Approach to analysis and modeling of multivariate spatial data in meteorology and climatology - K.R. Gabriel (Lead Paper)	4.1.1-8
4.2	Reconstructing surface pressure patterns using principal components regression on temperature and precipitation data - P.D. Jones, T.M.L. Wigley, and K.R. Briffa	4.2.1-8
4.3	The use of multivariate analysis to establish oceanic climatic regions - P.R. Lowe	**
4.4	A multiple climatic reduction technique applied to the extension of precipitation series - V.R. Barros and I. G. Schmidt	4.4.1-6
4.5	EOF spectral estimation in climate analysis - R.F. Cahalan	4.5.1-7
4.6	Weather map pattern classification: refinements of a correlation method - D.E. Petzold (Reserve)	4.6.1-2

* - Manuscript included in back of book

** - Manuscript not available

SESSION 5: MULTIVARIATE ANALYSIS OF METEOROLOGICAL AND CLIMATOLOGICAL DATA II

Chairmen: K.R. Gabriel and J. Rinne

5.1	Specification of complex modes of circulation with T-mode factor analysis - M.B. Richman	5.1.1-8
5.2	Statistical aspects of empirical orthogonal functions based on small sample sizes - H. von Storch	5.2.1-6
5.3	Meteorological aspects of principal component analysis in meteorological fields - W.M. Vargas and R.H. Compagnucci	5.3.1-9
5.4	A three-dimensional approach to the study of diurnal precipitation patterns in the American plains - R.C. Balling	5.4.1-5
5.5	Multivariate data analysis applied to natural rainfall regions of Portugal - F. da Costa Nicolau (Reserve)	**

SESSION 6: STOCHASTIC MODELING OF PRECIPITATION IN SPACE AND TIME(1)

Chairmen: V.K. Gupta and A.H. Murphy

6.1	Probabilistic modeling of the precipitation process - I. Rodriguez-Iturbe (Lead Paper)	6.1.1-7
6.2	Stochastic modeling of rainfall in space and time - V.K. Gupta	6.2.1-6
6.3	Trend effects on long time scales - E. Waymire	6.3.1-2
6.4	Precipitation dynamics as a stochastic process - C.A. Leary, D.R. Haragan, and H.-R. Cho	6.4.1-8

* - Manuscript included in back of book

** - Manuscript not available

6.5	Are "scattered" thundershowers randomly distributed in space? An appraisal based on correlation analysis - D. Sharon	6.5.1-4

(1) Special session organized jointly with the Committee on Precipitation of the American Geophysical Union

SESSION 7: STATISTICAL METHODS IN APPLICATIONS OF CLIMATOLOGICAL DATA

Chairmen: O.M. Essenwanger and T.D. Potter

7.1	Statistical methods for the analysis of climatic data - R.D. Stern (Lead Paper)	7.1.1-6
7.2	Useful models of temperature data - R. Coe	7.2.1-6
7.3	Joint distributions of wind speed and direction as tools for urban climatology - K. Cehak	**
7.4	Calibration in paleoclimatology when the biological response is binary - S.R. Esterby and L.D. Delorme	7.4.1-3
7.5	An analysis of the space and time variation of growing season rainfall in the central United States - P.J. Lamb and M.B. Richman	7.5.1-6
7.6	The application of climatological data to determine the wind climate of a non-mountainous region - E.L. Petersen and I. Troen	**
7.7	Sampling meteorology and statistics in an air quality model - R.A. Porter and L. Canelas	**
7.8	Seasonal variation in flooding in the UK and its relationship to soil moisture status and seasonal variation in intense rainfall - A.D. Hewson (Reserve)	**
7.9	Fuzzy multifactorial evaluation in climatological assessment - G. Chen (Reserve)	7.9.1-2

* - Manuscript included in back of book

** - Manuscript not available

SESSION 8: STATISTICAL METHODS IN CLIMATE IMPACT ASSESSMENT

Chairmen: P.J. Lamb and R.D. Stern

Chairmen: P.J. Lamb and R.D. Stern

8.1	The role of statistics in climate impact analysis - T.M.L Wigley (Lead Paper)	8.1.1-10
8.2	Assessing the impact of climate change on the development of surface water resources - V. Klemeš and J. Němec	8.2.1-8
8.3	Climatic variation in south-east England and the effect of short-term associations between weather parameters on experimentation with perennial crops - I.G. Cumming	8.3.1-8
8.4	Climate-transport relationship in Great Britain J.P. Palutikof	8.4.1-7
8.5	A regional comparison of seasonal African rainfall anomalies - T.L. Peterbaugh, R.P. Motha, N.D. Strommen, W.L. Decker, and S.K. Leduc	8.5.1-8
8.6	Cyclic components in precipitation and insolation time series in Alto-Alentejo (Portugal) - J.A.M. Corte-Real, C. Portugal da Câmara, and M.E. Costa	8.6.1-8
8.7	The role of the Bologna meteorological station in the European network during the XVIII century and the statistical analysis of its temperature and precipitation time series - S. Comani, E. Baiada, and R. Finzi (Reserve)	**
8.8	Approach to estimating climate fluctuation in relation to agriculture - F.N. Kogan (Reserve)	8.8.1-5
8.9	Aridity fluctuations over central south Indian region - A.A.L.N. Sarma and M. Ravindranath (Reserve)	8.9.1-5

* - Manuscript included in back of book

** - Manuscript not available

SESSION 9: USE AND VALUE OF CLIMATE INFORMATION

Chairmen: E.E. Zhukovsky and E.S. Epstein

9.1	Optimal use and economic value of prior data and categorical and probabilistic climate forecasts - R. Krzysztofowicz (Lead Paper)	9.1.1-7
9.2	Statistical models of economically optimal application of climatological data - E.E. Zhukovsky	9.2.1-8
9.3	Use of climatological information in river basin planning under risk - L. Duckstein, D.R. Davis, and E.J. Plate	9.3.1-10
9.4	Use of decision analysis to assess the actual and potential value of short-range climate forecasts - A.H. Murphy, R.L. Winkler, and R.W. Katz	9.4.1-6
9.5	Credibility of climate estimates - A.R. Boehm	**
9.6	On the present and potential use of climate information by the U.S. private agricultural sector - P.J. Lamb, S.T. Sonka, and S.A. Changnon (Reserve)	9.1.1-6

SESSION 10: EXTREME VALUE STATISTICS IN METEOROLOGY AND CLIMATOLOGY

Chairmen: I.I. Gringorten and E. Suzuki

10.1	Extreme values and meteorology - J. Tiago de Oliveira (Lead Paper)	10.1.1-8
10.2	Stochastic process analysis of rainfall totals and extremes - K.J.A. Revfeim	10.2.1-5
10.3	On the estimation in the case of an exponential distribution with two parameters with an application to the problem of large values (exceedances) - R. Sneyers and M. Vandiepenbeeck	10.3.1-2

* - Manuscript included in back of book

** - Manuscript not available

10.4	The effect of seasonal variation and serial correlation on the extreme value distribution of rainfall data - T.A. Buishand	10.4.1-6
10.5	Comparison of Gumbel's and Jenkinson's models for estimation of probable maximum daily rainfall intensities under different return periods - A. Krishnan	**
10.6	The distribution of design storms from exceedance series - W. Zucchini and P. Adamson	10.6.1-4
10.7	An approach to the determination of extreme snow loads - N. Djordjevic and N. Aleksic (Reserve)	**
10.8	The choice of extremal models for meteorological data - K.F. Turkman (Reserve)	**

SESSION 11: TIME SERIES ANALYSIS

Chairmen: R.A. Madden and J. Tiago de Oliveira

11.1	Time series analysis in climatic system models - M.B. Okamoto (Lead Paper)	11.1.1-6
11.2	High-resolution frequency analysis and applications to the prediction of almost-periodic functions - G.W. Brier, M.M. Siddiqui, M.W. Haurwitz, and R. Biondini	11.2.1-8
11.3	Spectral auto variance, cross variance and coherence analysis of long air temperature series - C.-D. Schönwiese	11.3.1-6
11.4	Dynamic MESA-spectra of climatological time series - H.P. Junk	11.4.1-6
11.5	Red and white noise and turbulence of low-level winds for time intervals of seconds - O.M. Essenwanger	11.5.1-8
11.6	Noninteger spectral analysis of climatological data - G. Padmanabhan and A.R. Rao (Reserve)	11.6.1-8

* - Manuscript included in back of book

** - Manuscript not available

11.7	Spectral analysis of summer precipitation series - R.H. Compagnucci and W.M. Vargas (reserve)	11.7.1-7
11.8	The 16-day peak in the second EOF of the 500 mb geopotential height J. Rinne and A. Sarkanen (Reserve)	11.8.1-8
11.9	Search for volcanic and CO_2 signals in surface air temperature data - A. Robock (Reserve)	**
11.10	Observational evidence of the relations between large-scale orographic forcing, blocking and stratospheric winter circulation - G. Hartjenstein (Reserve)	**

SESSION 12: STATISTICAL CLIMATE PREDICTION I

Chairmen: R. Livezey and N. Nicholls

12.1	On the prospects for climate prediction - R.A. Madden (Lead Paper)	12.1.1-6
12.2	500 mb geopotential variability in a 10-year climate simulation experiment - M. Déqué	12.2.1-4
12.3	Objective classification of synoptic pressure patterns over southeast Australia - J.D. Jasper and H. Stern	12.3.1-6
12.4	The synoptic climatology of winter surface temperature in the United States in relation to the monthly mean 700 mb circulation - W.H. Klein and J.M. Kline	12.4.1-8
12.5	Specification of monthly precipitation in the western United States from monthly mean circulation - B.C. Weare and M.A. Hoeschele	12.5.1-6
12.6	The stochastic-dynamic theory of air-sea interaction applied to the quasibiennial and semi-annual oscillation of sea surface temperature - M.-T. Li and J.-Y. Huang (Reserve)	12.6.1-5

* - Manuscript included in back of book

** - Manuscript not available

SESSION 13: STATISTICAL CLIMATE PREDICTION II

Chairmen: G. Der Megreditchian and W.H. Klein

13.1	Statistical climate prediction: the state of the art - N. Nicholls (Lead Paper)	13.1.1-7
13.2	Bayesian prediction applied to simple linear regression: monthly mean North American snow cover - E.S. Epstein	13.2.1-5
13.3	Prediction of seasonal temperatures based on persistencies in cyclone frequency EOFs - B.P. Hayden	13.3.1-5
13.4	Some results of applications of statistical methods to the climate changes and short-term climate prediction in China - M.-T. Li and D.-R. Yao	13.4.1-8
13.5	An application of the threshold autoregression procedure to climate analysis and forecasting - J. Shi, Q. Zhou, and J. Xiang	13.5.1-4
13.6	A new forecasting method of time series - J. Zhou	13.6.1-6
13.7	Further development of seasonal climate forecasts for the territory of Hungary - J. Bartholy and M. Kaba (Reserve)	13.7.1-5
13.8	An adaptative method of periodic analysis and its application to climate prediction - H. Cao and Z. Lu (Reserve)	13.8.1-4

SESSION 14: STATISTICAL METHODS IN DETECTING CLIMATE CHANGE

Chairmen: G.W. Brier and G.C. Tiao

14.1	Statistical analysis of general circulation model climate simulation, sensitivity, and prediction experiments - R. Livezey (Lead Paper)	**

* - Manuscript included in back of book

** - Manuscript not available

14.2	Techniques for making inferences about climatic variability - R.W. Katz	**
14.3	Evaluation of climate change experiments - F.W. Zwiers	14.3.1-8
14.4	A study of the statistical significance of the response of a spectral general circulation model to a modification of solar radiation - J.-M. Veysseire	14.4.1-6
14.5	On the interpretation and estimation of the equivalent sample size - H.J. Thiebaux and F.W. Zwiers	14.5.1-10
14.6	Evaluation of the statistical significance of climatic changes - C. Goossens and A. Berger	14.6.1-8
14.7	On the verification of January GCM simulations - H. von Storch and E. Roeckner (reserve)	14.7.1-8

SESSION 15: TIME SERIES MODELS

Chairmen: R.W. Katz and K.J.A. Revfeim

15.1	Recent developments in time series methods with application to meteorological data - G.C. Tiao (Lead Paper)	**
15.2	Bayesian model selection - A.R. Rao and R.A. Cantiller	15.2.1-8
15.3	A Markov model for variations in cloud cover - H. Madsen, H. Spliid, and P. Thyregod	15.3.1-7
15.4	Statistical approach to soil moisture content based on a Markov chain model for precipitation - T. Faragó	15.4.1-6
15.5	Application of a regular Markov chain to computations of runs of historical flood and drought years - C.S. Yao	15.5.1-6

* - Manuscript included in back of book

** - Manuscript not available

15.6	A Box-Jenkins analysis of series of winter-stength in the Danube region - K. Cehak (Reserve)	15.6.1-4

SESSION 16: STATISTICAL ANALYSIS AND MODELING OF METEOROLOGICAL AND CLIMATOLOGICAL DATA I

Chiarmen: M.B. Okamoto and E.L. Scott

16.1	Evidence for 40,000 year scaling regime in climatological temperatures - S. Lovejoy and D. Schertzer	16.1.1-5
16.2	Level and interannual variability of seasonal mean temperatures in Europe - C.J.E. Schuurmans and A.J. Coops	16.2.1-5
16.3	The 500 mb geopotential height as a first-order Markov process - J. Rinne and A. Sarkanen	16.3.1-6
16.4	Why do we get spurious correlations in climatology: some explanations - I.T. Jolliffe	16.4.1-6
16.5	Progressive analysis of Kendall's coefficient for depicting a trend or a jump in long time series - J.C. Mendes and F.R. Ferrão (Reserve)	**
16.6	Trend correction in climatological time series obtained by hard limiting procedure - M.B. Okamoto and K. Iwase (Reserve)	16.6.1-2

SESSION 17: STATISTICAL ANALYSIS AND MODELING OF METEOROLOGICAL AND CLIMATOLOGICAL DATA II

Chairmen: S.V. Singh and P.N. Somerville

17.1	Stochastic modeling of rainfall temporal patterns - V.T.V. Nguyen	17.1.1-4
17.2	A method of analysing implied periods od climatological data using step-wise regression - F.Y. Wei, Z. Zhao, and X.G. Zhang	17.2.1-4

* - Manuscript included in back of book

** - Manuscript not available

17.3	Examples of use of statistical techniques for various climatic data - C.-F. Hsu, S.A. Changnon, and K.R. Gabriel	**
17.4	Some applications of fuzzy sets to analysis of climate and climate simulation - H. Cao	17.4.1-4
17.5	The distributions of dry, wet, warm and cold spells in monthly climatic series in Argentina. Feasibleness of a Markov's model application - W.M. Vargas (Reserve)	**
17.6	On the regularization of the pluviometric series - G. Flocchini and C. Bagnasco (Reserve)	17.6.1-8

* - Manuscript included in back of book

** - Manuscript not available

AUTHORS INDEX

	Paper
Abouammoh, Abdulrahman M.	2.2
Adamson, Peter	10.6
Aleksic, Nenad	10.7
Alexandersson, Hans	3.2
Allerup, Peter	3.4
Amegee, Kodjo Y.	2.5
Bagnasco, Cristina	17.6
Baiada, E.	8.7
Balling, Robert C., Jr.	5.4
Baptista da Silva, João	3.7
Barros, Vicente R.	4.4
Bartholy, Judit	13.7
Bean, S.J.	2.3
Berger, A.	14.6
Biondini, R.	11.2
Boehm, Albert R.	9.5
Brier, Glenn W.	11.2
Briffa, K.R.	4.2
Brown, Barbara G.	3.3
Buishand, T. Adri	10.4
Cahalan, Robert F.	4.5
Canelas, L.	7.7
Cantiller, Rosalinda	15.2
Cao, Hongxing	13.8; 17.4
Cehak, Konrad	7.3; 15,6
Changnon, Stanley A., Jr.	9.6; 17.3
Cho, Han-Ru	6.4
Coe, Richard	7.2
Comani, Silvia	8.7
Compagnucci, Rosa H.	5.3; 11.7
Coops, A.J.	16.2
Cornish, Brain A.	1.3
Corte-Real, João A.M.	8.6
Costa, Maria Emília S.	8.6
Costa Nicolau, Fernando	5.5
Cuenca, Richard H.	2.5
Cumming, Ian G.	8.3
Davis, Donald R.	9.3
Decker, W.L.	8.5
Delorme, L.D.	7.4
Déqué, Michel.	12.2
Der Megreditchian, Guy	1.1

	Paper
Di-Rong, Yao	13.4
Djordyevic, N.	10.7
Duckstein, Lucien	9.3
Ellis, Jacques	3.8
Epstein, Edward S.	13.2
Essenwanger, Oskar M.	11.5
Esterby, Sylvia R.	7.4
Faragó, Tibor	15.4
Ferrão, Fernando R.	2.6; 16.6
Finzi, R.	8.7
Flocchini, Giuseppe	17.6
Gabriel, K. Ruben	4.1; 17.3
Goossens, Christine	14.6
Gringorten, Irving I.	2.1
Guofan, Chen	7.9
Gupta, Vijay K.	6.2
Haragan, Donald R.	6.4
Martjenstein, Gisela	11.10
Haurwitz, M.W.	11.2
Hayden, Bruce P.	13.3
Hewson, A.D.	7.8
Hiemstra, Lourens A.V.	3.6
Hiroh, Atsushi	2.4
Hoeschele, Marc A.	12.5
Hsu, C.F.	17.3
Huang, J.-Y.	12.6
Husain, Tahir	1.2
Itoh, Masashi	2.4
Iwase, Kōsei	16.6
Jasper, J.D.	12.3
Jolliffe, Ian T.	16.4
Jones, P.D.	4.2
Junk, Hans-Peter	11.4
Kaba, Magdolna	13.7
Katz, Richard W.	3.3; 9.4; 14.2
Khan, H.U.	1.2
Klein, William H.	12.4
Klemes, Vit	8.2
Kline, Joann M.	12.4
Kogan, Felix N.	8.8
Krishnan, A.	10.5
Krzysztofowicz, Roman	9.1

	Paper
Lamb, Peter J.	7.5; 9.6
Leary, Colleen A.	6.4
Leduc, S.K.	8.5
Li, Mai-Tsun	12.6; 13.4
Livezey, Robert E.	14.1
Lovejoy, Shavn	16.1
Lowe, Paul R.	4.3
Lu, Zhijun	13.8
Madden, Roland A.	12.1
Madsen, Henning	3.4
Madsen, Henrik	15.3
Mendes, José C.	16.6
Monahan, Edward C.	2.7
Motha, Raymond P.	8.5
Murphy, Allan H.	3.3; 9.4
Nemec, Jaromir	8.2
Nguyen, Van-Thanh-Van	17.1
Nicholls, Neville	13.1
Okamoto, Masanori B.	11.1; 16.6
O'Muircheartaigh, Iognaid G.	2.7
Ozturk, Aydin	2.2
Padmanabhan, G.	11.6
Palutikof, Jean P.	8.4
Pessanha, Luís E. Vasconcelos	2.6
Pettersen, Erik L.	7.6
Petzold, Donald E.	1.4; 4.6
Plate, Erich J.	9.3
Porter, Richard A.	7.7
Portugal da Câmara, Carlos	8.6
Puterbaugh, Thomas L.	8.5
Padmanabhan, G.	11.6
Rao, A. Ramachandra	11.6; 15.2
Ravindranath, M.	8.9
Revfeim, Kristian J.	10.2
Richman, Michael B.	5.1; 7.5
Rinne, Juhani	11.8; 16.3
Robock, Alan	11.9
Rodriguez-Iturbe, Ignacio J.	6.1
Roeckner, Erich	14.7

	Paper
Sarkanen, Annakaisa	11.8; 16.3
Sarma, A.A.L.N.	8.9
Schertzer, Daniel	16.1
Schmidt, Irma G.	4.4
Schönwiese, Christian-Dietrich	11.3
Schuurmans, C.J.E.	16.2
Sharon, David	6.5
Shi, Jiuen	13.5
Siddiqui, M.M.	11.2
Singh, Shyamvir V.	3.1
Sneyers, Raymond	10.3
Snijders, Tom A.B.	3.5
Sommerville, Paul N.	2.3
Sonka, Steven T.	9.6
Spliid, Henrik	15.3
Stern, H.	12.3
Stern, Roger D.	7.1
Strommen, Norton D.	8.5
Thiebaux, H.J.	14.5
Thyregod, Poul	15.3
Tiago de Oliveira, José	10.1
Tiao, George C.	15.1
Troen, I.	7.6
Turkman, Kamil F.	10.8
Ukaily, M.A.	1.2
Valença, Ademilde de S.	3.8
Vandiepenbeeck, M.	10.3
Vargas, Walter M.	5.3; 11.7; 17.5
Veisseire, Jean-Michel	14.4
von Storch, Hans	5.2; 14.7
Xiang, Jingtiang	13.5
Yao, C.S.	15.5
Yao, D.R.	13,4
Zhang, X.G.	17.2
Zhao, Zhen	17.2
Zhou, Jiabin	13.6
Zhou, Qingang	13.5
Zhukovsky, E.E.	9.2
Zucchini, Walter	10.6
Zwiers, Francis W.	14.3; 14.5

	Paper
Waymire, Edward	6.3
Weare, Bryan C.	12.5
Wei, F.Y.	17.2
Wigley, T.M.L.	4.2; 8.1
Winkler, Robert L.	3.3; 9.4

A QUANTITATIVE BASIS FOR A SOLAR RADIATION NETWORK EVALUATION

Donald E. Petzold

Department of Geography
University of Maryland
College Park, Maryland 20742, U.S.A.

1. INTRODUCTION

Solar radiation measurements are being used extensively for investigations of hydrometeorological phenomena in small drainage basins. However, there have been relatively few cases in which such data have been considered for practical forecasting purposes over large basins, since the number of measurement sites needed to obtain an adequate sample was not known.

The design of any solar radiation data network must consider certain characteristics of solar radiation: (1) Its intensity at the surface of the earth may vary considerably over a short distance due to weather and cloud conditions. (2) Latitude determines the solar intensity at the outer edge of the atmosphere, thus it may affect solar receipts at the surface in basins of great north-south extent. (3) Surfaces at high elevations generally receive more intense solar radiation, due to a thinner overlying atmosphere. Hence, large basins of great topographic relief may exhibit considerable variations in solar receipts. (4) Certain sections of large basins may be dominated by persistent storm tracks or air masses which would cause distinct differences in seasonal solar radiation receipts across the basin. If there are no systematic differences in solar radiation (caused by (3), (4) or (5), above), then radiation differences found over short distances may be thought of as random occurrences and the radiation data network may be evaluated statistically.

The analysis of this type of spatial variation is not new to geographical climatology, and several approaches are possible. First, a complete surface of radiation values may be mapped, as is done routinely with pressure, temperature and precipitation. Unfortunately, such a dense network of solar radiation measuring stations is not feasible economically and logistically. Second, a network of moderate density may be analyzed by some form of triangulation, common in precipitation analysis. As an example, Erickson and McCorquodale (1966) suggested the use of the Thiesson polygon method to estimate radiation available to melt snow in the Manicouagan River basin. A major disadvantage of this methods is that it requires several stations along the perimeter of the basin which is often inaccessible and removed from the focus of study. Finally, a network may be defined by determining areas of the basin which are represented by single radiation measurements. This is clearly the most economical and suitable for most purposes of analysis.

The evaluation of a representative network must include the consideration of three factors: error, distance and time. Error represents the difference between the radiation flux at the representative site and that at a remote location. If systematic differences in regional solar radiation can be ruled out, then radiation differences found over short distances may be thought of as random occurrences, and the error will increase with distance and decrease as the time interval increases.

These thoughts prompted a study of radiation characteristics of the Churchill River basin of 89,000 km^2, located in the eastern Canadian subarctic, approximately 1850 km NNE of Washington, D.C. The basin is roughly triangular in shape, with the major east-west axis being approximately 450 km and the major north-south approximately 410 km. The average elevation of the basin is approximately 550 m above sea level with mountainous areas in the western and southern parts of the basin averaging between 770 m and 950 m.

2. SYSTEMATIC RADIATION DIFFERENCES

The size of the Churchill basin necessitates the examination of each possible cause of systematic differences in solar radiation receipt. The most obvious possibility arises from the effect of latitude, since the basin spans more than 3 degrees of latitude. This

influence may be examined by calculating solar radiation receipts at the top of the atmosphere for the 15th of each month at latitudes 52.0° N and 55.0° N, assuming a solar constant of 1396 W m^{-2}. This range covers the north-south extent of the basin. The results, presented in Table 1, show variations of 0 to 30% across the basin. However, variations during the summer months are very small; less than 4% between April and August. Since this summer period includes the active snowmelt and evaporation seasons, it appears that latitudinal effects can be neglected for most hydrometeorological forecasting purposes in the basin.

Table 1. Solar radiation ($MJm^{-2}day^{-1}$) at the top of the atmosphere on the 15th of each month.

Month	Lat. (°N): 52.0	55.0	Variation Across Basin (%)
Jan	7.3	5.6	26.4
Feb	13.9	11.9	14.8
Mar	21.9	20.2	8.0
Apr	31.5	30.4	3.6
May	39.2	38.6	0.6
Jun	42.6	42.4	0.0
Jul	41.1	40.7	0.8
Aug	35.3	34.5	2.4
Sep	26.7	25.1	5.8
Oct	17.3	15.5	10.8
Nov	9.8	8.0	20.0
Dec	6.6	4.9	29.8

Areas of the basin at high elevations receive more solar radiation due to the shallower atmosphere above. To assess elevational differences, data collected during the snowmelt season from two adjacent stations on the periphery of the basin were analyzed on an hourly basis (Petzold, 1974). The relationship indicated a systematic difference in the order of 11% for every 200m in elevation, based on measurements taken at 517 m. However, the mountainous areas of the basin are only 220 to 400m higher than the mean elevation of the basin and the areas of mountainous terrain are quite small compared to the total area of the basin. Thus, the actual differences involved are probably small enough to be ignored.

The question of systematic solar radiation variation caused by frontal or air mass persistence can be answered only from data collected over relatively long periods of time. A comparison of monthly averages of solar radiation over a 10 year period at stations at opposite ends of the basin (Schefferville, Quebec and Goose, Labrador) exhibits no consistent differences (see Fig. 1). Instead, equal amounts of scatter are found around the 1:1 line, despite the large elevational differences between the two stations (Schefferville, 517 m; Goose, 50 m). In addition, there is a negligible difference in average daily values during this period of comparison (Schefferville 10.93 $MJm^{-2}day^{-1}$; Goose 11.01 $MJm^{-2}day^{-1}$).

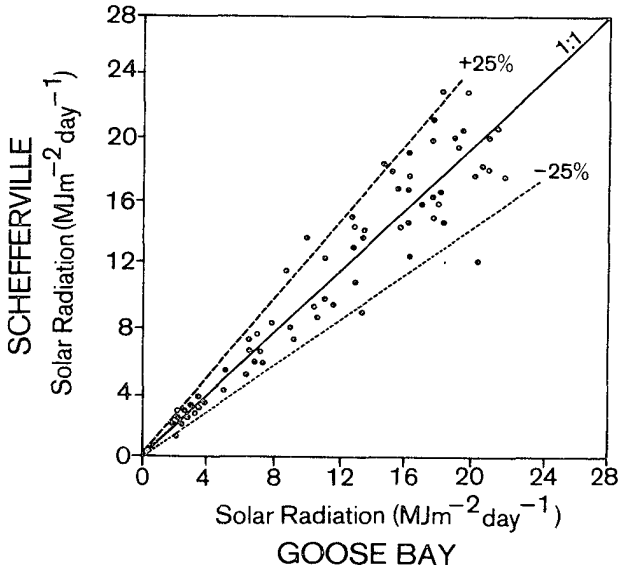


Fig. 1 Comparison of mean monthly solar radiation receipts at Schefferville, Quebec and Goose, Labrador.

Although conditions exist which could produce systematic differences in solar radiation across the basin, it appears that none is evident on an annual basis. However, it is not possible to state conclusively that there are no systematic differences until more data are available for analysis.

3. DISTANCE ANALYSIS

Assuming the random nature of solar radiation differences, it is now possible to assess the variations of solar radiation across the basin. Only the summer months (April through August) will be analyzed due to their hydrometeorological importance.

The scatter of data exhibited in Fig. 1 indicates that reliable short-term solar radiation data for this basin cannot be provided by measurements taken only at Schefferville and Goose. To answer questions concerning the necessary number of and location of stations, pyranometers were installed at a number of settlements across the basin such that eight different combinations of station pairs, separated by distances between 39 and 237 km, could be used to evaluate the effect of distance on a representative solar radiation network (see Fig. 2 for station locations). A summary of the distance analysis appears in Table 2 which includes the mean daily differences ($\bar{d}$) in recorded radiation totals for selected pairs of stations, and standard deviations (S) of the differences for two summer seasons.

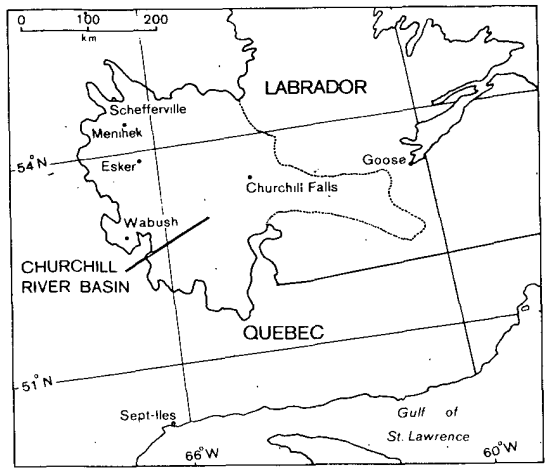


Fig. 2 Stations used in the distance analysis.

Table 2. Distance analysis of radiation differences between stations. (Radiation values in $MJm^{-2}day^{-1}$)

Stations**	Distance (km)	Days	$\bar{K}\downarrow$	Differences $\bar{d}$	S
1973					
Esk-Men	68	29	18.1	0.582	2.663
Wab-Esk	109	119	16.1	1.239#	4.396
ChF-Men	163	118	17.3	0.067	4.400
ChF-Esk	200	37	18.1	0.862	4.995
ChF-Wab	203	137	16.7	1.357#	5.309
ChF-Sch	230	42	15.1	1.022	7.545
ChF-Goo	237	98	18.3	0.653	6.016
1974					
Wab-Esk	109	92	18.1	1.243	4.300
ChF-Esk	200	110	18.8	0.423	4.204
ChF-Wab	203	94	18.0	0.825	4.513
ChF-Sch	230	89	18.0	0.842	5.460
ChF-Goo	237	70	19.9	0.419	5.192

d greater than ± 7.1% of $\bar{K}\downarrow$.

** Station abbreviations: Schefferville (Sch), Menihek (Men), Wabush (Wab), Churchill Falls (ChF), Goose (Goo), Esker (Esk).

If it is assumed that the measurement at any one station may be in error by ± 5%, then any difference less than ± 7.1% (= $(5^2 + 5^2)^{1/2}$) can be attributed to measurement error. Only two data sets, as noted in Table 2, have $\bar{d}$ values that exceed this limit. Both are associated with the Wabush station which produced generally poor quality data during 1973. The mean differences are small enough to conclude that all stations have similar radiation climates. Therefore, it is possible to assume that $\bar{d}$ approaches zero, and the standard deviation of the differences (S) may be used to indicate the probable error involved in extending a point measurement of solar radiation into the surrounding region.

The variation of S with distance is shown in Fig. 3. These data points generally agree with those derived by Wilson and Petzold (1972) for stations in Southern Canada, with the difference between the two data sets being small, less than 0.84 $MJm^{-2}day^{-1}$. The calculated best-fit regression equation relating S to distance (D) is

$$S = 2.05 + 1.80 \times 10^{-2}D - 1.675 \times 10^{-5}D^2 \quad MJm^{-2}day^{-1} \qquad (1)$$

which has a correlation coefficient of 0.84. For both the Churchill Basin and southern Canada data combined, the calculated relation is

$$S = 1.00 + 2.74 \times 10^{-2}D - 4.815 \times 10^{-5}D^2 \quad MJm^{-2}day^{-1} \qquad (2)$$

which has a correlation coefficient of 0.9[illegible]. Since Eqs. (1) and (2) produce similar results as shown in Fig. 3, Eq. (2) has been used in subsequent analyses since it was derived from considerably more data (34 points as compared to 12).

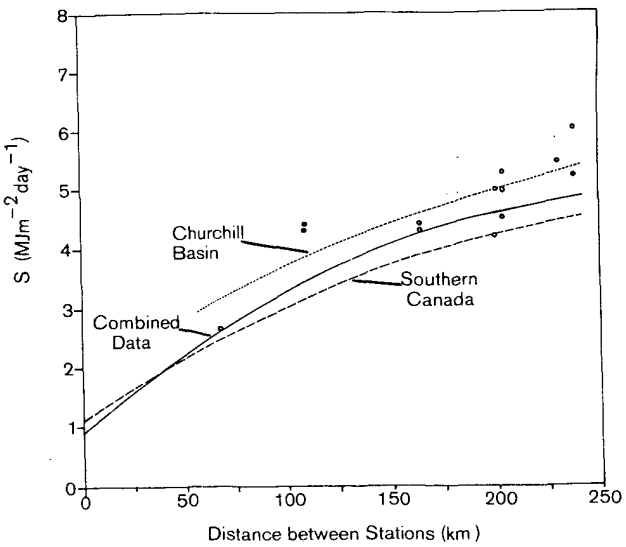


Fig. 3 Variations of the standard deviation with distance

Frequency distributions of each data set were subjected to a 5% chi-square test to determine if the daily differences between stations were distributed normally. As shown in Table 3, only three cases produced negative results.

Table 3. Calculated chi-square values and decisions of tests for normality (Critical value for 5% chi-square: 12.592).

Stations	Calculated Chi-square (χ^2)	Decision (5% level)
Esk-Men	9.708	Yes
Wab-Esk	18.988	No
ChF-Men	9.892	Yes
ChF-Esk	7.478	Yes
ChF-Wab	9.153	Yes
ChF-Sch	12.014	Yes
ChF-Goo	3.874	Yes
Wab-Esk	12.468	Yes
ChF-Esk	20.161	No
ChF-Wab	6.646	Yes
ChF-Sch	14.555	No
ChF-Goo	7.243	Yes

An interesting feature of the frequency distributions, particularly those which tested negatively, is that the observed peak is higher than would be expected to occur in a normal distribution. This indicates that there is a slightly greater chance of finding a small difference between stations than would be predicted on the basis a normal distribution.

The last step in the present analysis, that of determining the effect of using averaged radiation totals for periods of more than one day, may be approached by using the Central Limit Theorem. Although the mean differences between stations will remain the same in any data set whether 1-day, 2-day or n-day values are used, the standard deviation of the differences will decrease as the averaging time increases. The Central Limit Theorem applies to this case as follows: if the daily differences occur randomly from a normal distribution, the standard deviation for n-day averages, S_n, will be given by:

$$S_n = Sn^{-\frac{1}{2}} \qquad (3)$$

Therefore, the expected standard deviation for differences between stations using two-day averages would be $S_2 = S/1.41$, and so on.

The expected summertime variations of S_n with distance for radiation averages of 1 to 4 days are shown in Fig. 4. Since these values represent a variation of ±1.0 standard deviation about the mean and the differences are distributed normally, the curves shown in Fig. 4 would apply to approximately 67% of individual values. Calculation of 90% probability limits results in the distance variations shown in Fig. 5. These last two diagrams show the integrated effects of error, distance, and time which may now be incorporated into the design of an effective representative network.

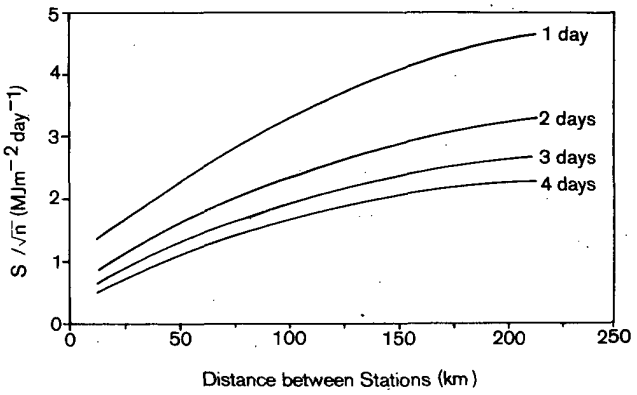


Fig. 4 Expected variations of S_n at the 67% probability level.

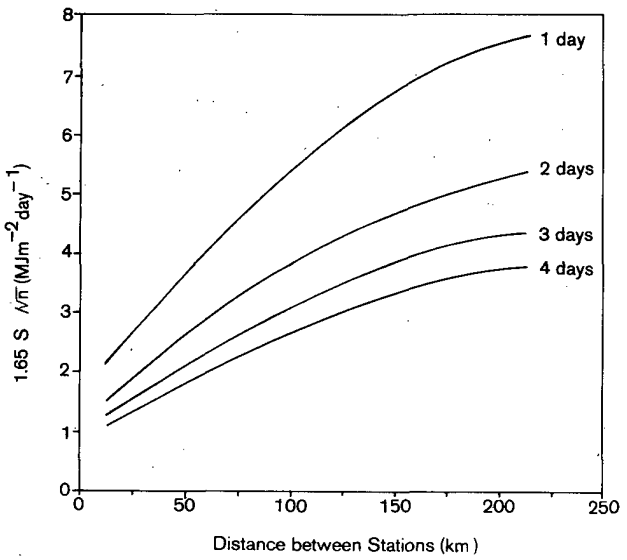


Fig. 5 Expected variations of S_n at the 90% probability level.

4. EVALUATION OF EXISTING NETWORK

Snowmelt runoff forecasts may be one of the primary uses for the basin solar radiation data, so this evaluation concerns the main snowmelt period in May. Table 4 presents various alternatives of averaging time, probability levels, and acceptable error, and shows corresponding distances to which a point measurement may be extended, rounded to the nearest 5 km as determined from Figs. 4 and 5. Error levels of ±10%, ±15%, and ±20% amount to 1.88, 2.85, and 3.81 $MJm^{-2}day^{-1}$, or heat sufficient to produce 6, 9, and 11 mm of meltwater in a snowpack, respectively. It should be noted that these values pertain to a mean solar radiation value for May of 18.93 $MJm^{-2}day^{-1}$, so that the actual percentage errors will be greater when radiation totals are smaller (and therefore less important), but the errors will be smaller when the effect of radiation is most important.

Table 4. Recommended distances (km) for extending solar radiation data based on mean May conditions.

No. of Days	67% Probability Error: ±10%	67% Probability Error: ±15%	90% Probability Error: ±15%	90% Probability Error: ±20%
1	35	75	30	55
2	70	150	60	95
3	100	--	90	145
4	130	--	110	200

The first decision to be made concerns the acceptable error limit. Measurement error itself can be assumed to have an upper limit of ±5%. A limit of ±10% might be acceptable for small basins, but is clearly out of the question for a basin of this size. At the same time an error of ±20% generally would be considered too large to be of any practical use. Therefore the following analysis is based on a designed error of ±15%.

The second decision concerns the acceptable period for which the solar radiation data may be averaged. Runoff predictions would be updated each day although there would be a lag time of several days before the meltwater from remote areas reached one of the reservoirs. This point is significant since it means that the time element can be used to decrease uncertainties involved in the predictions for distant locations. The most effective use of this element will depend on the actual lag times involves in various parts of the basin, but for the present purpose it is assumed that an optimum period for radiation data does not exceed 3 or 4 days. Allowing 1 day for data compilation and snowmelt estimation, this would provide 2 or 3 days notice in the case of a location having a one week lag.

Having made these basic decisions, Table 4 now can be used to evaluate the radiation network. On the first day of prediction there would be a "probable" estimate (67% probability level) for a maximum distance of 75 km from each station. As shown in Fig. 6, large areas would be left uncovered. The majority of the basin would recive a "probable" estimate by the second day since coverage from each station would extent to 150 km.

The response for "definite" estimates (90% probability level) for periods of 2, 3 and 4 days is shown in Fig. 7. The western section of the basin is provided with "definite" estimates after only two days. Three regions have poor coverage and would not be classified as "definite" until at least the fifth day: (1) an area of approximately 900 km^2 in the north; (2) an area of about 5800 km^2 in the south; and (3) an area of about 1500 km^2 in the lower basin between Churchill Falls and Goose Bay. The latter two could be most critical for runoff prediction since there is a very large uncovered area in the south, and lag times of water flow in the lower basin will be considerably shorter than in most parts of the upper basin due to the shorter distances involved.

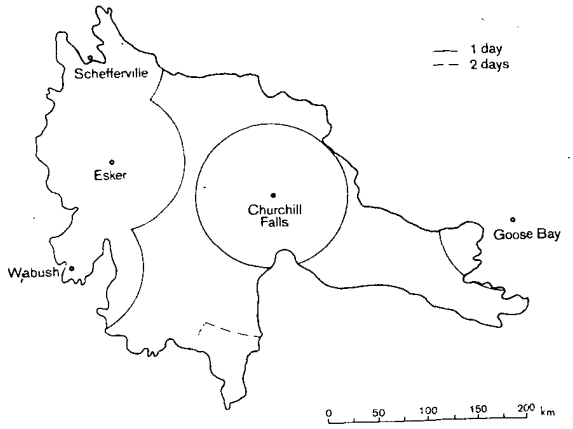


Fig. 6 Basin coverage based on the 67% probability level.

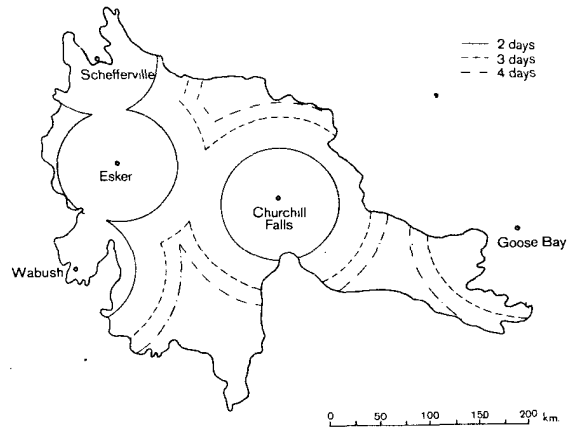


Fig. 7 Basin coverage based on the 90% probability level.

5. CONCLUSIONS

On the basis of this analysis, it appears that: (1) two new stations (in the northern and southern regions) be added in the central basin, and (2) one or more stations be considered for the lower basin if runoff forecasts are required there.

Although this evaluation was performed for the snowmelt period, the recommendations are equally applicable for other purposes during the summer season defined earlier. The solar radiation data for example, could be used to estimate evaporation losses. An evaluation for the period from June to September shows that the average radiation income decreases during the summer months, so that the areas with good coverage are smaller than those for the snowmelt period, and the need for additional stations is accentuated.

The technique used here provides a quantitative basis for a solar radiation network evaluation and it can also serve in a more limited capacity to indicate the reliability of data extrapolated from a single station to another site nearby. Its use should be re-

stricted however to regions having similar radiation climates.

6. REFERENCES

Erickson, O.M. and J.A. McCorquodale, 1966: Applications of computers to the determination of snowmelt runoff. Stat. Methods in Hydrol., Proc., Hydrol. Sympos., 5, 361-376.

Petzold, D.E., 1974: Solar and Net Radiation Over Snow. Clim. Research Ser. No. 9, McGill Univ., Montreal, 77 pp.

Wilson, R.G. and D.E. Petzold, 1972: Daily solar radiation differences between stations in southern Canada. Clim. Bull., 11, 15-22.

MODELING CLIMATOLOGY OF AREAL COVERAGE

Irving I. Gringorten

Air Force Geophysics Laboratory (LYT)
Hanscom AFB, MA 01731

1. INTRODUCTION

Throughout a given area, at any instant in time, the weather, such as rainfall depicted on a radar screen, or clouds in a satellite picture, produces a pattern that is unique to that instant in time, not likely to be duplicated at another instant. The change, from one instant to another, is likewise unique. Both the pattern and its change, however, are governed by basic restraints that characterize the climate of the area. The relevant climatology must consist of the statistics of the events or the probabilities of events in the area.

If the probability is known, say, of 24-hour rainfall exceeding 10 mm at a single observational point, we must ask: what is the corresponding probability that the same quantity will be exceeded everywhere in a surrounding area, or some fraction of a surrounding area? It is this type of climatology that is the subject of this paper.

Satellite pictures, recording as they do the areal coverage of cloudiness, eventually should help to provide the kind of cloud climatology that we seek. Radar has been a valuable asset in recent years because of the panoramic view that is provided of precipitation within a circular area surrounding the radar station. Satellite and radar, however, are recent tools. Generally, the existing climatic records contain single-instant observations of weather at single-point weather stations. Recorded rainfall amount is that at a rain gauge. Temperature is that in the thermometer shelter, and so on. The climatology of a city usually is that at an observational point in the city.

With respect to time, most observations might be considered as instantaneous, observed at regular intervals. But there are also time-integrated observations, like 24-hour totals of rainfall. There are observations of the average temperature, and maximum and minimum temperatures in a 24-hour period. All such observations contribute to the climatological character of a location as it relates to time or the duration of events. Is there a similar climatic record of spatial coverage of climate?

In most previous studies, correlation coefficents are the principal object of investigation, with notable exceptions. Schreiner and Riedel (1978) have collected rainfall data integrated over areas ranging in size from 26 km^2 to 2600 km^2, and have found frequencies of extremes of rainfall as a function of size of the ground covered. Other studies, by Court (1961), Briggs (1972) and Roberts (1971) have linked the rain in a circular area to the central single-station amounts by idealized models. In general, however, efforts at areal coverage have been few and models of areal coverage have been rare. The following examples illustrate the potential of such studies.

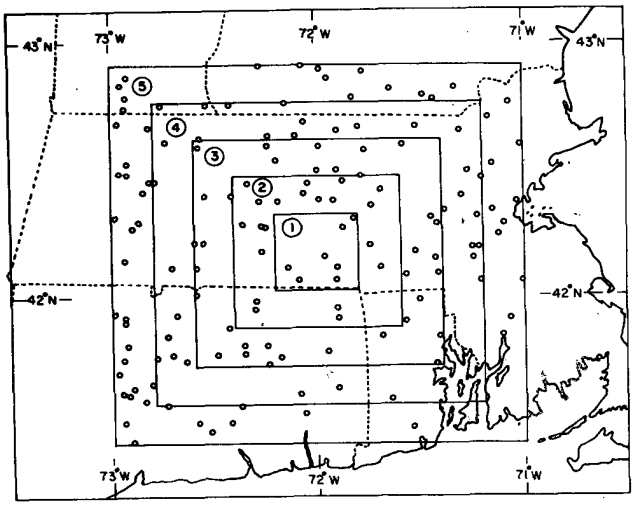


Fig. 1. A map showing the location of 145 stations equipped with rain gauges in New England.

Fig. 1 shows southern New England, centered in Massachusetts. There are 145 little circles showing locations of stations that report 24-hour rainfall. There are 5 nesting squares of areas ranging from 1000 km^2 to 28,000 km^2. Some 20 years of record have provided the estimates (Table 1) of the maximum 24-hour rainfall amounts as a function of areal size. The frequency of no rain at a single station is 62%, but within a sizeable area this frequency must drop steadily with increasing size of the area. The single-point probability of 25 mm of rain is 2.8%, or the probability that it will not be exceeded is 97.2%. As we look at areas of increasing size, this cumulative probability must decrease. Or conversely the prob-

Table 1. New England July maximum 24-hour rainfall, in square areas ranging in size from 1,000 km^2 to 28,000 km^2.

Rain Amount (mm)	Single-Station Frequ.	Frequency of Areal Maximum (Area in units of 10^3 km^2)		
		1	10	28
none	.62	.44	.17	.07
≤ 25	.97	.94	.76	.63
≤ 75	.999	.997	.984	.965

ability that rainfall will exceed 25 mm, somewhere in an area, must increase with increasing size of the area.

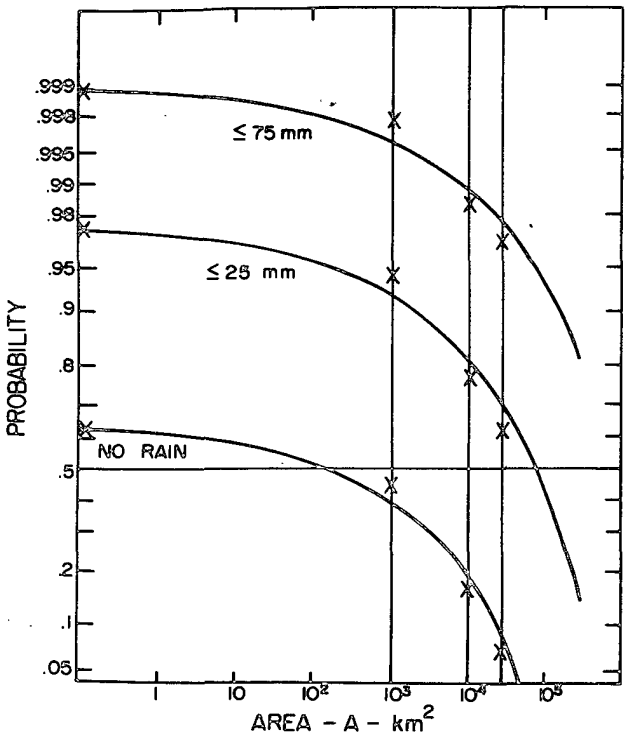


Fig. 2. The cumulative frequency of the maximum 24-hour rainfall in the areas of 10^3 km^2, 10^4 km^2 and 2.8×10^4 km^2, July in New England. The curves were fitted by a model of areal coverage.

The numbers that appear in Table 1 also appear in Fig. 2 at the X's. The solid curves have been drawn by means of a model of areal cover, which I shall try to describe shortly. In order to fit the curves to the observational data, a value had to be found for the model's parameter, called scale distance (2.63 km in the example). To understand this parameter, for the present, we describe it as the distance over which correlation coefficient is 0.99.

Another question is: What is the probability of rain over a fraction of the whole area? Fig. 3 is a sketch of one picture of a radar PPI-scope, showing scattered areas of stormy weather. Fig. 4 is a graphical presentation of frequencies of

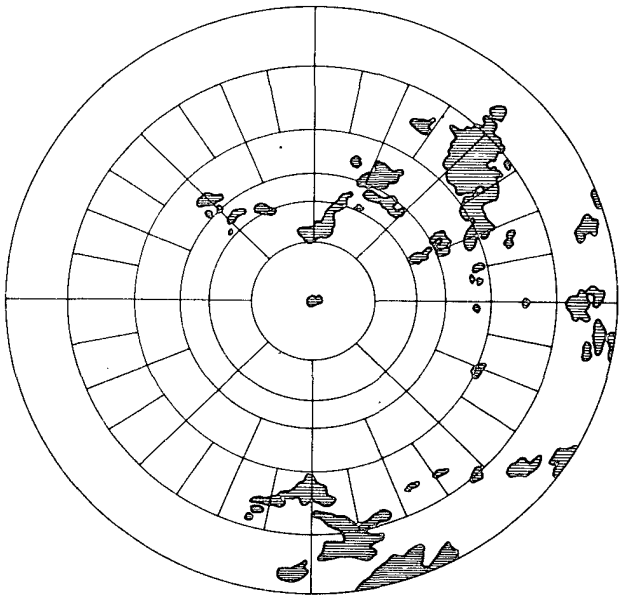


Fig. 3. Illustration of a PPI-radar picture, in which there were scattered storms producing radar echoes.

rainfall echoes. For example, at Key West, Florida, in summer, the single-point probability of a radar echo is 2%. But the radar sees rain over an area. We had divided the radar picture into cells, each cell of size 1580 km^2. Two cells have twice that area, all 64 cells have an area

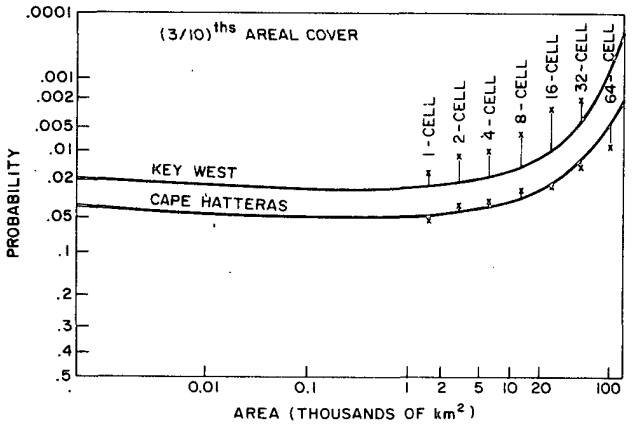


Fig. 4. The frequencies of radar echoes covering, at least, (3/10)ths of the areas ranging from 1.6×10^3 km^2 (1 cell) to 10^5 km^2 (64 cells). The plots are for Key West, Florida and Cape Hatteras, N.C. in summer (all hours included).

of approximately 100,000 km^2. The frequency of echoes, in at least 3/10ths of small areas, is nearly the same 2%. If the area becomes large, then a 2% singlepoint probability will not yield 30% cover very often. The solid curves of this diagram, again, are provided by a model of areal coverage.

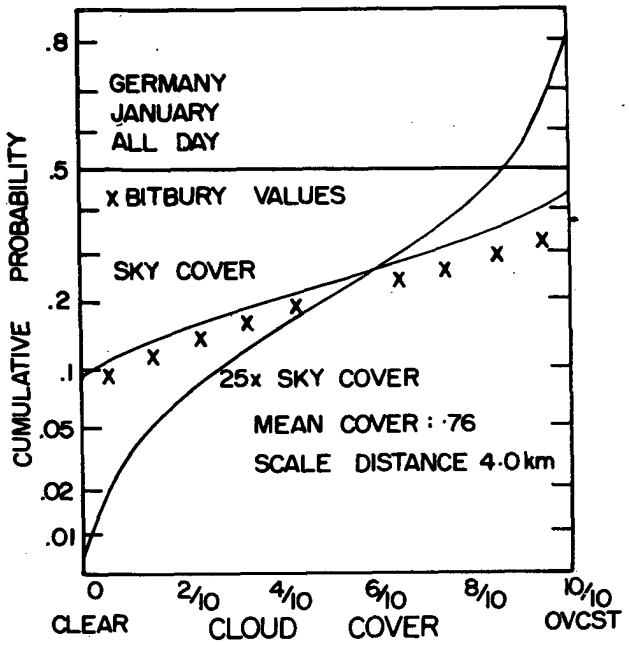


Fig. 5. The cumulative probability of cloud cover, from clear to overcast, for the area of the sky dome that an observer sees, and for 25 times that area. The example is for German stations with mean cloud cover 0.76 in January.

For partial cover let us look at Fig. 5, for an example of cloudiness. In this diagram we have plotted probabilities of cloud cover from clear to overcast. Germany has a satisfying high frequency of clouds in winter, averaging 76%. The summaries of ground observations at Bitburg are shown by the X's. The model provided a reasonable fit to this data, but the model does more. Consider a larger area that a satellite might see. Obviously the probability of all-clear must be less, likewise of complete overcast. The larger area must experience more cases of partial cover. The model provides such estimates that are otherwise not available.

One characteristic that we must simulate with our model is the change of correlation coefficient with the distance between stations. Fig. 6 is an example taken from a 1962 report (Bertoni and Lund). Correlation begins high for short distances, but drops to zero and becomes negative at substantial distances between stations. The example shows cc of surface temperatures plotted against distances between stations from a few km to several thousand km apart.

Theoretically an abundance of data, at an abundance of stations in a small area, with a rapid succession of observations, could provide the kind of areal climatology that we seek. Practically such abundance would confront us with the enormous task of mapping and recording events as a history of the weather, prior to its summarization

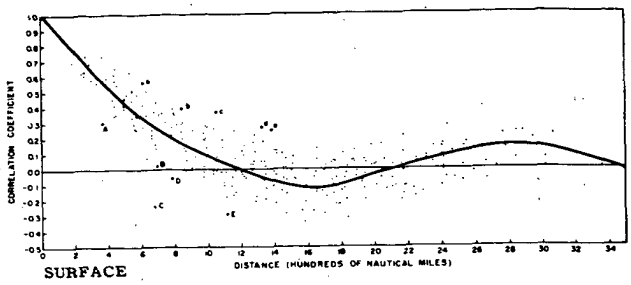


Fig. 6. Correlation coefficient versus distance between stations, of wintertime surface temperatures (Ref: Bertoni and Lund, 1964).

in statistical format. This is the difficulty that we wish to overcome by modeling. We want to be able to simulate a sequence of changes in the weather covering an area, whenever we choose, and we want, ultimately, to provide usable statistics of areal coverage, including cc's, and probabilities of fractional cover by weather conditions such as cloud cover.

By and large we must depend on statistical models which I must yet describe. But there is a concept that needs to be examined first, concerning the equivalent normal deviate.

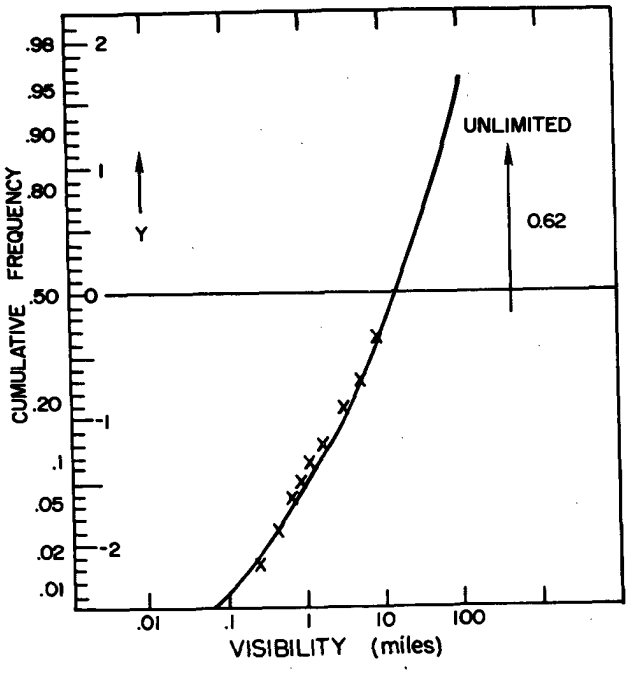


Fig. 7. Illustration of transformation of a weather element into its equivalent normal deviate (or vice versa). Example is for January noontime visibility at Bedford, MA.

2. EQUIVALENT NORMAL DEVIATE (END)

Fig. 7 is a sample plot of the cumulative frequency of a weather element, in this case of visibility at Bedford, MA, in January, at noon-time. The X's mark the frequencies that have been summarized from 20 years of data. They show, for example, a frequency of 2% for visibility less than 1/4 mile, 7% for visibility less than 1 mile, also 62% frequency of visibility greater than 10 miles. Alongside the scale of cumulative probability we can plot another scale of the equivalent normal deviate, (y). The correspondence of y to F(y) can be found in nearly every textbook on statistics. By means of a diagram such as Fig. 7, or alternatively by equations, we obtain a transformation of the variable (V in this case) into its END (y), or vice versa.

Our models are derived using the END's. The results are then transformed to the appropriate weather variables, when so required.

3. APPROACH TO AREAL CLIMATOLOGY

An effective model of areal cover will produce a simulated horizontal field, stochastically, without recourse to physical laws or dynamics. Two early models (Gringorten, 1979) were devised in the manner illustrated in Fig. 8. Imagine that our horizontal space is filled with a fine-mesh grid of points. Assign, initially, random numbers to all points. Around point (i,j) draw a circle and take a weighted average of all the random numbers within the circle, then replace the random number at (i,j) with the weighted average, y(i,j). If we do this for all points within a large square, we shall obtain a pattern wherein the new values, at all points, will be correlated with each other. Models that have been composed in this way differ from each other in the weighting function that is selected for each model.

Once a model is selected, the adopted procedure can be repeated many times to produce many simulated synoptic situations, thus comprising a large sample that can be surveyed for spatial correlations or associations.

This early modeling procedure has been set aside in favor of a new model that requires the generation of only a few dozen random numbers per map, instead of many thousands.

4. BOEHM'S SAWTOOTH WAVE MODEL (BSW)

The presently preferred model was devised by Major Albert R. Boehm, USAF. He has given it the descriptive name Sawtooth Wave. We shall refer to it as the BSW-model.

We begin with a horizontal space (Fig. 9) with axes (U,V) and point of origin (0,0). We introduce a stationary wave formation with waves that are straight and uniform in wavelength. The wavelength (Λ) becomes the parameter of the model, which, in our climatic fields, measures several hundred to several thousand km. In the following derivation, however, the wavelength is the unit of distance.

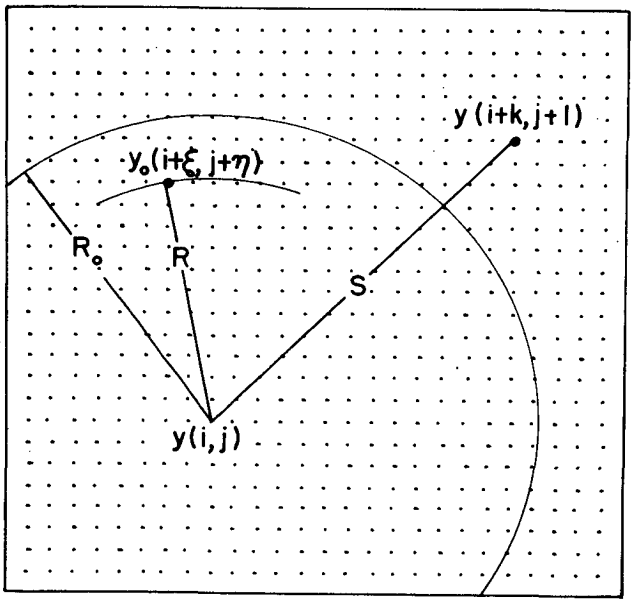


Fig. 8. To illustrate the former development of stochastic modeling.

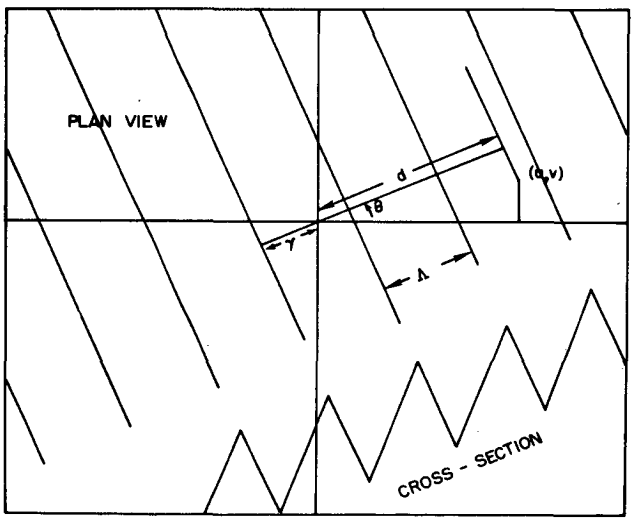


Fig. 9. To illustrate the stochastic procedure in the BSW-model.

In cross-section the waves are saw-tooth in form. The left-hand edge of each wave has zero height, the right-hand edge has a height of unity. At any point (u,v) in between the wave phase, x(u,v), is uniformly distributed between zero and one.

The wave formation has two degrees of freedom, and thus allows for selection of two random numbers: one for the direction (θ) of the wave formation and one for the phase at the point of origin, x(0,0), which must be equal to the fractional distance γ shown in the diagram. At every point in the U-V field there is a wave phase, x(u,v).

Now, consider many (N) wave formations, each with a random direction (Θ_i) and each with a random phase (γ_i) at the origin. Then, at each point (u,v) there will be N values of

$$x_i(u,v) \text{ for } 1 \leq i \leq N$$
$$0 \leq \Theta_i \leq 2\pi$$
$$0 \leq \gamma_i \leq 1$$

The average, y(u,v), of all N values of $x_i(u,v)$, as N becomes large, will have an asymptotically Gaussian distribution. When, for convenience, N is made equal to 12, the equation becomes

$$y(u,v) = \sum_{i=1}^{12} x_i(u,v) - 6$$

There will be a value of y(u,v) at each point (u,v) in the horizontal space. A sample of such generated fields is shown in Fig. 10, with distinct low and high cells and saddle points.

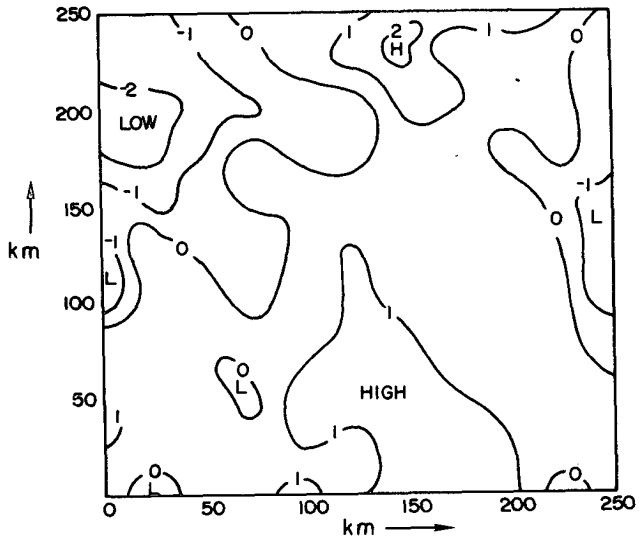


Fig. 10. Example of horizontal field generated stochastically by the BSW-model.

4a. Correlation

An analytical solution has been found for the correlation coefficient ρ(s) between points separated by distance (s), in units of the wavelength:

$$\rho(s) = 3s^2 - (12/\pi)s + 1 \qquad \text{for } 0 \leq s \leq 1$$
$$= 3s^2 - (12/\pi)s + 1 + (24/\pi)[\cos^{-1}(1/s) - \sqrt{s^2-1}] \qquad \text{for } 1 < s \leq 2$$

The graph of ρ(s) versus s is shown in Fig. 11. For our purposes we shall rarely need to consider distances of more than one wavelength. The distance (s) will usually be a number between zero and one, given by

$$s = s'/\Lambda$$

where s' is the distance (km) and Λ is the parameter wavelength (km).

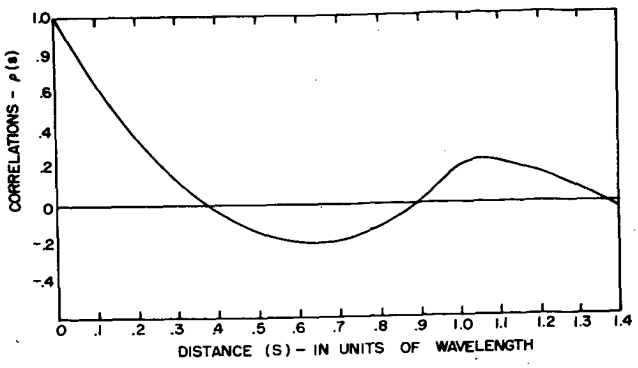


Fig. 11. Graphical presentation of the change of cc with distance in the BSW model.

4b. Areal coverage

While we have an analytical solution for the correlation decrease with distance we lack such good fortune, at this time, for the primary goal of this paper: to find the probability of the maximum condition in an area or a fraction of the area. We have obtained approximate, although good, solutions by Monte Carlo simulation. The following is a brief description of our procedures.

In a field generated by the model, consider a square area of side s (in units of the wavelength). At some point in the area (s^2) there will be a maximum. There will be, also, a tenth largest, a two-tenths largest, --- and finally the minimum in the area. In terms of a fraction (f) of the area the maximized minimum in such fraction, $y(f,s^2)$, becomes the primary subject. We want to find the cumulative probability distribution of the threshold value: $P(y:f,s^2)$.

A large number (M) of stochastically generated maps will provide a frequency distribution of $y(f,s^2)$ that will approach the true probability distribution asymptotically as M is increased.

This kind of Monte Carlo simulation was conducted with M = 25,000. On each map 12 square areas were surveyed for the largest value of y, the (1/10)th largest, (2/10)th largest, and so on, down to the smallest value. The linear size of each square, in units of wavelength, was

$$s = 2^z/340 \qquad \text{where } z = -1(1)10$$

From all 25,000 maps cumulative frequencies were found for each

$$y(f,s^2) = -3.5(0.5)3.5$$

$$\text{for } f = 0(0.1)1.0$$

$$\text{and } z = -1(1)10.$$

Using the frequencies thus obtained the charts of Fig. 12 were drawn to show the probability estimates of y-maxima for f = 0.0, 0.1, 0.3, 0.5 successively, and for z from -1 to 8. Smoothing was necessary, especially at the extremes of the distributions.

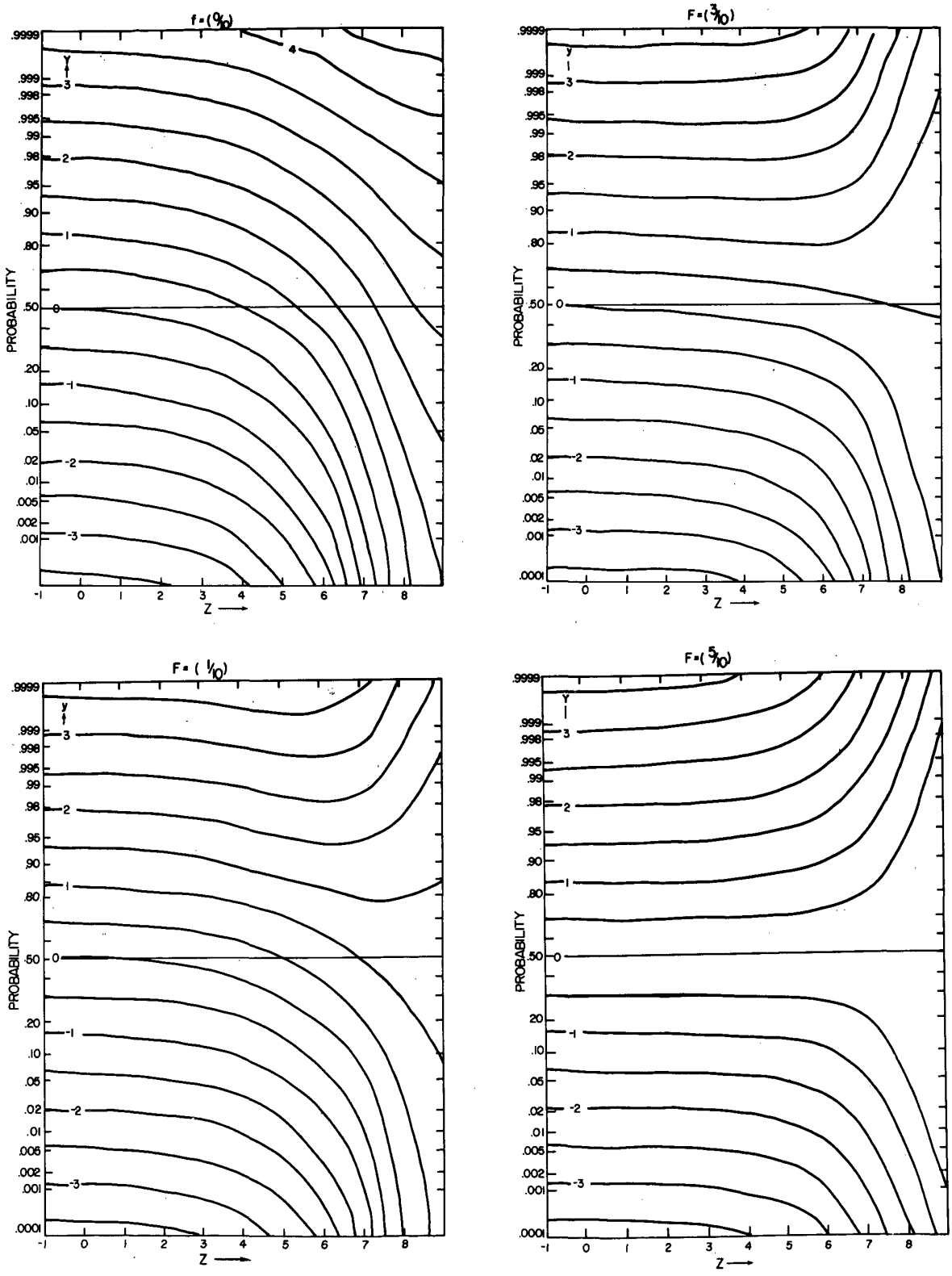


Fig. 12. BSW-model charts of the probabilities, $P(y;f,s^2)$, of fractional cover (f) of an area (s^2) of END-values not exceeding y.

4c. Applications

For a square area, A km^2, with linear dimension s' km we set

$$s = s'/\Lambda$$

for which

$$z = \ell n(340s)/\ell n\ 2$$

When z is entered in a chart of Fig. 12, the vertical line upward from the abscissa provides estimates of the probabilities of various thresholds of y corresponding to the area (A).

Instead of the wavelength (Λ) it is convenient to substitute the parameter (r), called the scale distance. In the BSW-model it is a distance over which the cc is 0.9888 (close to 0.99). By and large, it is a more meaningful parameter in application, given by

$$r = s'/340s \text{ km} = \Lambda/340 \text{ km} = \sqrt{A}/2^z \text{ km}.$$

Or, for z:

$$z = \ell n(\sqrt{A}/r)/\ell n\ 2$$

For the New England 24-hour winter precipitation, we found scale distances of 9 to 10 km. For radar echoes they ranged from 1 to 2 km in summer, 2 1/2 to 4 1/2 km in winter. Cloud cover scale distances have ranged from 1/2 km to 10 km depending on season, time of day as well as geographic location.

5. TIME CHANGES

The above-described model enables us to generate "snapshot" pictures stochastically of a horizontal field. So, far, the END-value y(u,v) at a point (u,v) on one map is independent of that generated on any other map.

Let us suppose that one map has been generated at time t with values of $y_t(u,v)$ at each point (u,v) of the space. Let another map be generated, for the time (t+δt), in terms of an END (η) to give values of $\eta_{(t+\delta t)}(u,v)$ at each point (u,v). Make the modification to $y_{(t+\delta t)}(u,v)$ by means of the Markov model:

$$y_{(t+\delta t)}(u,v)=\rho_t \cdot y_t(u,v)+\sqrt{1-\rho_t^2}\cdot \eta_{(t+\delta t)}(u,v)$$

where the cc (ρ_t) for time lag (δt) is given by

$$\rho_t = \exp(-\ t/T\ )$$

where T (in hours) is a parameter called relaxation time.

Between the y-values of the later map, the cc can be proved to be still $\rho(s)$, the same as it was on the map at time t. That is, the cc-field on the new map, δt hours later, is preserved. This is a necessary condition to assure the continuity of the model's characteristics. An example of hourly changes is shown in Figs. 13 for T = 20 hours.

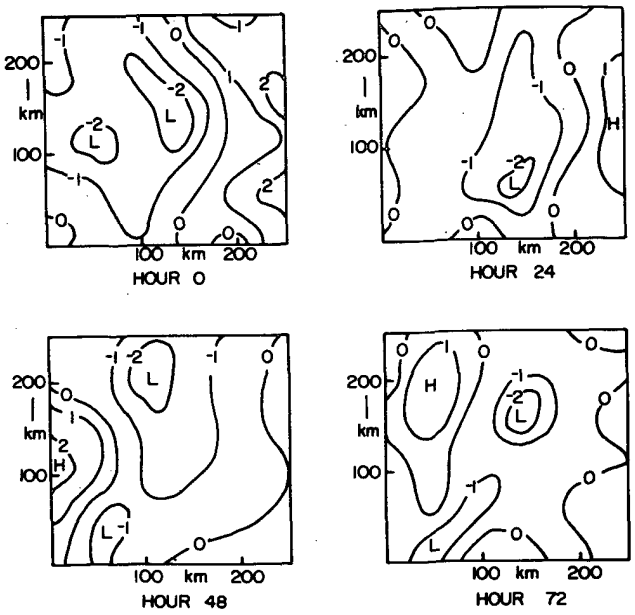


Fig. 13. Example of stochastic horizontal fields following one another at 24-hour intervals in a Markov process.

6. REMAINING PROBLEMS

These efforts have raised more problems than they have solved. We have devised a procedure to generate synoptic pictures of the weather, stochastically, without regard for the underlying physical processes. We have made no effort to use the models of Numerical Weather Prediction (NWP) to produce successive pictures of the meteorological events. At this time, however, we do not know if the hourly sequences of NWP will, in fact, yield a climatology that approximates that of the real climate with respect to such elements as cloud cover, especially mesoscale features. At this point in time we are justified in adopting a procedure that will produce maps that, collectively, will yield a climatology consisting of a good approximation to the horizontal field of correlation and satisfactory probability estimates of the maximum and minimum of the weather conditions in fraction (f) of an area (A). There is nothing in the model, yet, that will cause a synoptic situation to be advected as well as modified. An ultimate model should be 4-dimensional, to yield probabilities of conditions in 3-dimensional space and to include events in time. Moreover, the probability of events or conditions should be estimated analytically instead of by Monte Carlo simulation. We have made a beginning, but much remains to be done.

7. ACKNOWLEDGEMENTS

The author's associates, Donald D. Grantham and Charles Burger, have contributed generously to the substance of this paper and have provided critical insight into the problems. The Sawtooth Wave Model is the invention of Major Albert R. Boehm, USAF, Environmental Technical Applications Center. Much of the background effort was performed by the computer division, AFGL, and by students of Regis College, Weston, MA.

8. REFERENCES

Bertoni, E. A., and Lund, E. A., 1964: Winter space correlations of pressure, temperature and density to 16 km. Environ. Res. Pap., No. 75, AFCRL-64-1020.

Briggs, J., 1972: Probability of aircraft encounters with heavy rain. Metorol. Mag., 101, pp 8-13.

Court, A., 1961: Area-depth rainfall formulas. J. Geophys. Res., 66, pp 1823-1831.

Gringorten, I. I., 1979: Probability models of weather conditions occupying a line or an area. J. of Appl. Meteorol., 18, pp 957-977.

National Bureau of Standards, 1961: Handbook of Mathematical Functions, Applied Mathematics Series, 55, Government Printing Office, Washington, D.C., 1045 pp.

Roberts, C. F., 1971: A note of the derivation of a scale measure for precipitation events. Mon. Wea. Rev., 99, pp 873-876.

Schreiner, L. C. and Riedel, J. F., 1978: Probable maximum precipitation estimates United States east of 105th meridian, Hydrometeo. Rep. No. 51, NOAA-NWS-HR-51.

FITTING GENREALIZED LAMBDA DISTRIBUTION TO CLIMATOLOGICAL DATA

Aydın Öztürk and Abdulrahman M. Abouammoh

Department of Statistics
King Saud University
Riyadh, Saudi Arabia

1. INTRODUCTION

One of the main goals of fitting a probability distribution to a set of data of a climatic event is to describe the general statistical properties of the underlying event. So far many probability models have been proposed to describe climatological data. Suzuki (1980) has reviewed most of the work that has been done to find some reasonable probability distribution models for various climatic events. It has been pointed out that some of the climatic events and the corresponding proposed distributions are:

Temperature: Normal, Pearson Type I.

Precipitation: Gamma, Log-normal, Kappa type.

Relative Humidity: Beta type.

Wind speed: Gamma, Weibull, Log-normal type.

The nature of large variability of the random sample usually requires the use of different distributions. The discrepancies between the probability models which are fitted to data of the same climatic event may be explained as the different characteristic caused by different factors such as locations and periods. If a certain probability distribution can be specified for a given data set, then the problem of fitting a probability model and interpreting the results would simplified considerably. However in many practical situations one often faces the difficulty of determining in advance the type of the model which gives the best fit to the data. One common approach to overcome the above difficulty is to derive the probability model which reflects the real mechanism of the underlying event. This kind of model should include some components to explain the variation due to location, periods ... etc.

Another approach could be used is to find a general probability distribution which covers almost all of the possible distributions of the underlying event. Tere are few systems of distributions which cover most of the well known distributions. Among these distributions the generalized lambda distribution (GLD) which has a unique functional form and covers a wide range of distributions.

The GLD has been used for various purposes. Kottas and Lau (1979)suggested it for modelling lead time demands. Silver (1977) used GLD to estimate the safety factor in an inventory control model. Öztürk and Dale (1982a) used GLD as a probability model for the weekly solar radiation data at West Lafayette Indiana U.S.A. and obtained satisfactory results. The GLD can also be used for Monte Carlo studies (O'Mera (1976) and Schmeiser (1977)).

The main objective of this study is to intrroduce GLD as a general probability model for some climatic factors. In section 2 a breif definition of the GLD distribution and estimation procedures are explained. In section 3, some results

based on maximum, minimum and average temperature and relative humidity data are discussed. The last section is devoted to some remarks on using GLD. In the appendix two of the estimation methods are briefly explained.

2. THE GENERALIZED LAMBDA DISTRIBUTION

The GLD was developed by Ramberg and Schmeiser (1972, 1974) as a generalized form of Tukey's three-parameter lambda distribution, Tukey (1960). The quantile function of GLD is given by

$$X = Q(p) = \lambda_1 \{ p^{\lambda_3} - (1 - p)^{\lambda_4} \}/\lambda_2; \quad 0 \leq p \leq 1$$

where λ_1 and λ_2 are location and scale parameters while λ_3 and λ_4 are the shape parameters and p is the value of a uniformly distributed random variable P in the interval 0 to 1 . The probability density function of x is

$$f(x) = f\{Q(p)\} = 1/Q'(p)$$

$$= \lambda_2 \{ \lambda_3 p^{\lambda_3 - 1} + \lambda_4 (1-p)^{\lambda_4 - 1} \}^{-1}$$

$$0 \leq p \leq 1$$

where $Q'(p)$ is the derivative of Q(p).

The plot of the density curve may be obtained by plotting f(x) on the y-axis versus X=Q(p) on the x-axis. Similarly plot of the distribution function may be obtained by plotting p on the y-axis versus X=Q(p) on the x-axes.

The GLD includes a wide range of curve shapes , for example an approximation can be made for standard normal distribution by selecting the parameter values as $\lambda_1 = 0, \lambda_2 = 0.1975$ and $\lambda_3 = \lambda_4 = 0.1349$. Similary a good approximation to the exponential distribution is obtained for $\lambda_1 = \lambda_2 = \lambda_4 = 0.0004$ and $\lambda_3 = 0$, (see Ramberg et al. (1979) and Ramberg and Schmeiser (1974) for more detailed discussion on the properties of the distribution).

Estimation procedure of the parameters of GLD is somehow difficult. Ramberg et al. proposed the method of moments as a procedure to estimate the parameters. Considering that the sample moments (especially the third and fourth moments) are sensitive to extreme observations and are subject to large sampling variability. Mykytka and Ramberg (1979) generalized the method of moments to allow the use of surrogate measures of location, scale, symmetry and tail-weight. Öztürk and Dale (1982b) developed a procedure to obtain least squares estimates of the parameters. Defining the guantity R_i as

$$R_i = \left(\frac{i}{n+1}\right)^{\lambda_3} - \left(\frac{n-i+1}{n+1}\right)^{\lambda_4}$$

the least squares estiments of λ_3 and λ_4 are obtained by minimizing .

$$\min_{\lambda_3 \lambda_4} \quad (1 - r_{XR}^2)$$

where r_{XR} is the correlation coefficient between the sample order statistic X_i and corresponding R_i values. Least squares estimates of λ_1 and λ_2 are given by

$$\hat{\lambda}_1 = \bar{x} - b_{XR} \cdot \bar{R}$$

$$\hat{\lambda}_2 = 1/b_{XR}$$

where x is the sample mean and $\bar{R}$ is the mean of R_i for a given set of values of λ_3 and λ_4 b_{XR} stands for the regression coefficient of X on R.

Making some Monte Carlo experiments Öztürk and Dale(1982b) have also indicated that the least squares procedure is less sensitive to sampling variations and give slightly better fit to the true underlying distribution than the other methods. Moreover, variances and covariances of the estimates of the parameters may be computed by using the least squares techniques.

3. APPLICATION OF THE DISTRIBUTION

In this section we intend to demonstrate how well the GLD fits the different sets of data based on different periods of the year.

For this purpose, a meteorological station namely Dahran (in Saudi Arabia) was selected. The data based on 27 years records which was available only for minimum temperature, maximum temperature, average temperature and relative humidity. It was not possible to get the complete records of some other factors such as precipitation, wind speed and solar radiation. The GLD could of course be fitted for these data sets if they were available.

For each factor, the mean values were computed using 27 obsevations for each month. Then, minimum, maximum, mean and standard deviation are obtained from the monthly means of the corresponding factors. Results are given in Table 1. Statistics for each month. which were used to compute the values in the table are based on different sample sizes because of the missing observations.

Table 1. Some basic statistics for each factor.

Factor	Min.	Max.	Mean	St.Dev.
Minimum Temp.	5.96	27.48	16.75	8.02
Maximum Temp.	25.62	46.00	37.83	7.52
Average Temp.	15.74	35.77	26.53	7.60
Average Humidity	38.67	66.41	55.33	8.58

Parameter estimates were computed using three different methods, namely the method of moment, the Mykytka-Ramberg (M-R) method and the least squares method. Table 2 gives the means of the estimates for the 12 months of the the year. It was noticed that although the estimates of the parameters obtained by each method appear to differ considerably, the differences between the overall fittings were much less than one might expect.

In order to have some idea about the goodness of fit of GLD a statistic which is similar to Kolmogorov-smirnov statistic is defined as

$$D^* = \max \left| Q^{-1}(x_i, \hat{\underline{\lambda}}) - \frac{i}{n+1} \right|$$

Table 2. Mean values of the parameter estimates

Factor	Method	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$	$\hat{\lambda}_4$
Minimum Temp	Moment	16.885	0.095	0.210	0.182
	M-R	16.525	-0.003	0.113	0.181
	Least squares	16.456	-0.170	0.009	0.084
Average Temp.	Moment	26.974	0.129	0.247	0.085
	M-R	26.960	0.117	0.202	0.256
	Least squares	27.091	0.023	0.107	0.031
Maximum Temp.	Moment	37.712	0.121	0.197	0.235
	M-R	37.990	0.083	0.171	0.111
	Least squares	37.990	0.047	0.111	0.048
Average Humidity	Moment	53.774	0.036	0.292	0.244
	M-R	55.455	0.018	0.256	0.226
	Least squares	53.307	-0.010	0.052	0.198

where $\hat{\underline{\lambda}} = (\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4)$. This statistic was computed for each month and each factor using three different methods of estimation. Results in Tables 3, 4 and 5 give an overall comparison between the three methods, also means and standard deviations of D* were computed and put at the end of each table. Since the distribution of D* is not known a goodness of fit test based on D* can not be made. However one may empirically conclude that the fittings are generally satisfactory.

4. CONCLUDING REMARKS

The GLD is proposed as a general probability distribution for the climatological factors. Because of its flexibility it can fit most of the data sets easily. Therefore, if one faces the difficulty of finding

Table 3. Values of D* based on the method of moments

Month	Minimum Temp.	Average Temp.	Maximum Temp.	Average Humidity
Jan.	0.177	0.137	0.110	0.311
Feb.	0.127	0.152	0.127	0.089
March	0.151	0.130	0.100	0.087
April	0.178	0.116	0.089	0.056
May	0.126	0.170	0.160	0.083
June	0.100	0.167	0.191	0.061
July	0.239	0.148	0.148	0.088
Aug.	0.151	0.178	0.199	0.077
Sept.	0.157	0.110	0.117	0.107
Oct.	0.114	0.115	0.120	0.103
Nov.	0.105	0.180	0.135	0.098
Dec.	0.098	0.139	0.141	0.071
Mean	0.144	0.174	0.136	0.102
S.D.	0.041	0.084	0.035	0.067

Table 4. Values of D* based on the method of Mykytka-Ramberg .

Month	Minimum Temp.	Average Temp.	Maximum Temp.	Average Humidity
Jan.	0.155	0.130	0.113	0.311
Feb.	0.117	0.157	0.122	0.089
March	0.154	0.141	0.100	0.087
April	0.190	0.199	0.088	0.056
May	0.129	0.187	0.162	0.083
June	0.101	0.165	0.204	0.061
July	0.303	0.293	0.149	0.087
Aug.	0.176	0.152	0.201	0.077
Sept.	0.164	0.173	0.109	0.107
Oct.	0.115	0.112	0.128	0.103
Nov.	0.107	0.157	0.156	0.089
Dec.	0.099	0.138	0.147	0.071
Mean	0.151	0.162	0.140	0.092
S.D.	0.057	0.047	0.037	0.035

Table 5. Values of D* based on the least squares

Month	Minimum Temp.	Average Temp.	Maximum Temp.	Average Humidity
Jan	0.162	0.140	0.114	0.156
Feb.	0.121	0.165	0.120	0.091
March	0.149	0.139	0.090	0.077
April	0.192	0.114	0.154	0.058
May	0.135	0.195	0.193	0.083
June	0.097	0.189	0.149	0.071
July	0.338	0.246	0.053	0.085
Aug.	0.157	0.156	0.203	0.074
Sept.	0.183	0.189	0.121	0.103
Oct.	0.112	0.114	0.124	0.092
Nov.	0.103	0.171	0.156	0.090
Dec.	0.092	0.122	0.136	0.079
Mean	0.153	0.163	0.137	0.083
S.D.	o.067	0.039	0.043	0.030

a suitable probability distribution to fit a given set of data, then the GLD could be used. However the following limitations should be taken into account;

(i) The GLD has 4 parameters and hence reliable results could be obtained for large samples only.

(ii) The GLD is defined in a certain region by its parameters. Samples falling outside this region may cause a difficulty in computing the parameter estimates. Although the proposed method overcomes this difficulty to some extent one should be cureful for these cases.

APPENDIX: ESTIMATION METHODS OF THE PARAMETERS

(i) Method of moments:

For the GLD, the mean, variance, third and fourth central moments are derived by Ramberg and Schmeiser (1974) as

$$\mu = \lambda_1 + A/\lambda_2$$

$$\sigma^2 = (B-A^2)/\lambda^2$$

$$\mu_3 = (C - 3AB + 2A^3)/\lambda_2^3$$

$$\mu_4 = (D - 4AC + 6A^2B - 3A^4)/\lambda_2^4$$

where

$$A = 1/(1 + \lambda_3) - 1/(1 + \lambda_4)$$

$$B = 1/(1 + 2\lambda_3) + 1/(1 + 2\lambda_4) - 2\beta(1 + \lambda_3, 1 + \lambda_4)$$

$$C = 1/(1 + 3\lambda_3) - 3\beta(1 + 2\lambda_3, 1 + \lambda_4) + 3\beta(1 + \lambda_3, 1 + 2\lambda_4) - 1/(1 + 3\lambda_4)$$

$$D = 1/(1 + 4\lambda_3) - 4\beta(1 + 3\lambda_3, 1 + \lambda_4) + 6\beta(1 + 2\lambda_3, 1 + 2\lambda_4) - 4\beta(1 + \lambda_3, 1 + 3\lambda_4) + (1 + 4\lambda_4)$$

In the equations above β stands for the usual beta function. The coefficients of skewness and kurtosis are

$$\alpha_3 = \mu_3 / \sigma^3 \qquad (1)$$

$$\alpha_4 = \mu_4 / \sigma^4 \qquad (2)$$

Sample estimates of α_3 and α_4 are obtained as

$$\alpha_3 = m_3/(m_2)^{3/2},\ \alpha_4 = m_4/m_2^2 \qquad (3)$$

where

$$m_r = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^r,\quad r = 1, 2, \ldots \qquad (4)$$

moment estimates of the parameters are obtained from the two simultaneous equations:

$$\hat{\alpha}_3 = \alpha_3(\lambda_3, \lambda_4) \qquad (5)$$

$$\hat{\alpha}_4 = \alpha_4(\lambda_3, \lambda_4) \qquad (6)$$

The solution of the above equations may be obtained by a function minimazation procedure by defining the objective function as

$$g(\lambda_3, \lambda_4) = [\hat{\alpha}_3 - \alpha_3(\lambda_3, \lambda_4)]^2 + [\hat{\alpha}_4 - \alpha_4(\lambda_3, \lambda_4)]^2 \qquad (7)$$

with the constraint $\lambda_3 \lambda_4 > 0$.

Once the values of $\hat{\lambda}_3$ and $\hat{\lambda}_4$ are computed then $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are computed as

$$\hat{\lambda}_1 = \bar{x} - A/\lambda_2 \qquad (8)$$

$$\hat{\lambda}_2 = \pm [(B - A^2)/m_2]^{1/2} \qquad (9)$$

where the sign of λ_2 is the same as the sign of λ_3. Ramberg et al. (1979) prepared tables to obtain the parameter estimates for a given setof values of α_3 and α_4.

(ii) Mykytka-Ramberg method;

Mykytka and Ramberg (1979) generalized the method of moments to allow the use of surrogate measures of location, scale, symmetry and tailweight. They defined the measures of symmetry and tailweight as

$$Q_3 = \frac{\bar{U}(0.05) - \bar{M}(0.50)}{\bar{M}(0.50) - \bar{L}(0.05)}$$

$$Q_4 = \frac{\bar{U}(0.05) - \bar{L}(0.05)}{\bar{U}(0.05) - \bar{L}(0.05)}$$

where $\bar{U}(0.05)$, $\bar{M}(0.50)$ and $\bar{L}(0.05)$ are respectiely the averages of the largest five percent of the x_i's, the middle fifty percent of the x_i's and the smallest five percent of the x_i's ...etc.

The asymptotic expectations are given by

$$E(Q_3) = [\mu_{\bar{U}(0.05)} - \mu_{\bar{M}(0.50)}] / [\mu_{\bar{M}(0.50)} - \mu_{\bar{L}(0.05)}]$$

$$E(Q_4)=[\mu_{\bar{U}}(0.05) - \mu_{\bar{L}}(0.05)]/[\mu_{\bar{U}}(0.50) - \mu_{\bar{L}}(0.50)]$$

where

$$\mu_{\bar{U}(\alpha)} = (\alpha\lambda_2)^{-1}\left[\frac{1-(1-\alpha)^{\lambda_3+1}}{\lambda_3+1} - \frac{\alpha^{\lambda_4+1}}{\lambda_4+1}\right] + \lambda_1$$

$$\mu_{\bar{L}(\alpha)} = (\alpha\lambda_2)^{-1}\frac{\alpha^{\lambda_3+1}}{\lambda_3+1} + \left[\frac{(1-\alpha)^{\lambda_4+1}-1}{\lambda_4+1}\right] + \lambda_1$$

$$\mu_{\bar{M}(\alpha)} = (\alpha\lambda_2)^{-1}\left[\frac{(1+\alpha)^{\lambda_3+1} - (1-\alpha)^{\lambda_3+1}}{\lambda_3+1}(2)^{\lambda_3+1}\right]$$

$$+\left[(2)^{\lambda_4+1}\frac{(1-\alpha)^{\lambda_4+1} - (1+\alpha)^{\lambda_4+1}}{\lambda_4+1}\right] + \lambda_1$$

Hence the parameters λ_3 and λ_4 may be eatimated by minimizing the objective function

$$f(\lambda_3,\lambda_4)=[\,Q_3-Q_3(\lambda_3,\lambda_4)]^2 + [Q_4-Q_4(\lambda_3,\lambda_4)\,]^2$$

with the constraint $\lambda_3\ \lambda_4>0$. The parameters λ_1 and λ_2 can then be estimated from (8) and (9).

5. REFERENCES

Kottas, J. F. and Lau, J. 1979: A realistic approach for modelling stochastic lead time distribution. AIIE Transaction, 11, 54-60

Mykytka, E. F. and Ramberg, J. S. 1979: Fitting a distribution to data using an alternative to moments. Proceeding of the 1979 winter simulation conference, 369-374

O'Mera, P. D. 1976: An investigation of linear rank statistics for multiple linear regression model. Unpublished Ph.D. thesis University of Iowa.

Öztürk, A. and Dale, R. F. 1982a: A study of fitting the generalized lambda distribution to solar radiation data. J. Appl. Meteor., 12, 995-1004.

Öztürk, A. and Dale, R. F. 1982b: The generelaized lambda distribution; A procedure for no-linear leaśt squares esimation of the parameters. (Submitted to Technometerics).

Ramberg, J. S., Dudewicz, E. J., Tadikamalla, P. R. and Mytytka, E. F. 1979: A probability distribution and its uses in fitting data. Technometerics, 21, 201-214.

Ramberg, J. S. and Schmeiser, B. W. 1972: An approximate method for generating symmetric random variables. Comm. ACM, 15, 987-990.

Ramberg, J. S. and Schmeiser, B. W. 1974: An approximate method for generating symmetric random variables. Comm.ACM, 17, 78-82.

Schmeiser, B. W 1977: Method for modelling and generating probabilistic components in digital computer simulation when the standard distribution s are not adequate: "A survey". Proceeding of the winter simulation conference, 51-57.

Silver, E. A. 1977: A safety factor approximation based upon Tukey's lambda distribution. Operational Research Quarterly, 28, 743-746.

Suzuki, E. 1980: A summerized review of theoretical distributions fitted to climatic factors and Markov chain models of weather sequences, with some examples. Statistical Climatology, Development in Atmospheric Science, Vol. 13, Elsevier.

Tukey, J. W. 1960: The practical relationship between the common transformation of percentages of counts of amounts. Tech. Rep. No. 36, Statistical Technichal Rep. Group, Princeton University.

MODELING VISIBILITY FOR DATA VOID REGIONS

S. J. Bean and P. N. Somerville

University of Central Florida
Orlando, FL 32816

1. INTRODUCTION

Records on weather have been kept for many locations for many years. These records are available for many purposes. One important use is climatological prediction for weather elements. If one wished to predict, for example, the probability of cloud cover greater than 0.9 at 1000 hours on February 4, 1984 at Mildenhall, England, one could examine the original records or summaries of those records to get an empirical estimate of that probability. A more efficient method would be to model the data using statistical distribution functions. Somerville and Bean (1979) and others have developed models for such weather elements as visibility, ceiling, sky cover, precipitation, windspeed, etc. A particular statistical distribution function is ordinarily used for a weather element, with a specific set of parameters being used for a given month, time of day and location. For example, the Weibull distribution is frequently used for visibility, and for Mildenhall, England at 1000 hours (local standard time) in February the values $\alpha = .1322$ and $\beta = 1.1251$ may be used in the cumulative distribution function $F(x) = 1-\exp(\alpha x^{\beta})$ to give good estimates of the probability of visibility less than x miles.

The problem is more difficult for stations where no records exist. In this study we develop Weibull models for visibility for Germany in which the parameters α and β are functions of geographic and climatic predictor variables.

2. MODELING APPROACH

There are many variables which may influence visibility. Variables which were investigated in this study include the following--elevation, relative elevation, east-west elevation differences, north-south elevation differences, proximity to a major body of water, population density, relative humidity, mean windspeed, mean precipitation, latitude, longitude, and functions of, and interactions between the above variables.

In the first step, for each month and hour period (the day was divided into 8 three-hour periods), the Weibull model was used to obtain values of α and β for 60 selected German locations. The values for each of α and β were then modeled as functions of the above variables. A stepwise regression program was used to ascertain which of the variables could be used as predictors for a specified month and hour period.

For each month, and each of the eight three-hour periods, expressions for α and β were obtained which were functions of the selected set of the above variables. The predictive equations were then refined by a nonlinear regression program in which the sums of squares of the residuals minimized were not the residual differences between the values for α and β and their predicted values, but were the differences between the empirical cumulative distribution function and the model cumulative distribution function. This latter regression produced a set of equations for predicting α and β from geographic and climatic predictor variables. One set of equations was produced for each month and time of day. These models were designated collectively as the "Variables Model."

The two variables which were most useful for predictive purposes (and were available at stations for which records were not available) were elevation, and relative elevation and henceforth when we use the term "Variables Model", we chall mean the model in which only the variables elevation and relative elevation were used. (Actually, our model uses the cubes of each of elevation and relative elevation.)

A second set of models was produced by again using the above-mentioned nonlinear regression program, this time using none of the predictor variables. The result was a pair of constants α,β for each month and time of day, which gave the best fit for all stations. These models were designed collectively as the "Constants Model."

3. METHODOLGY

Let $E_j(x_i)$ be the empirical probability (step function) that the visibility is less than x_i miles for the j^{th} station, using data from the RUSSWOs. The values of x_i are the following distances in miles: 1/4, 5/16, 1/2, 5/8, 3/4, 1, 5/4, 3/2, 2, 5/2, 3, 4, 5 and 6. Let

Month	Hour Period 06-08	09-11	12-14	15-17	18-20	21-23
Jan	.186	.212	.143	.128	.131	.109
	.707e-02	.598e-02	.655e-02	.687e-02	.713e-02	.704e-02
	-.154e-02	-.272e-02	-.296e-02	-.278e-02	-.188e-02	-.509e-03
Feb	.179	.176	.961e-01	.780e-01	.831e-01	.732e-01
	.832e-02	.717e-02	.659e-02	.671e-02	.697e-02	.775e-02
	-.300e-02	-.328e-02	-.285e-02	-.274e-02	-.226e-02	-.175e-02
Mar	.121	.848e-01	.384e-01	.285e-01	.307e-01	.307e-01
	.672e-02	.580e-02	.479e-02	.461e-02	.510e-02	.554e-02
	-.164e-02	-.178e-02	-.132e-02	-.119e-02	-.973e-03	-.139e-02
Apr	.734e-01	.365e-01	.204e-01	.153e-01	.149e-01	.144e-01
	.630e-02	.486e-02	.409e-02	.400e-02	.435e-02	.433e-02
	-.133e-02	-.862e-03	-.107e-02	-.965e-03	-.770e-03	-.654e-03
May	.562e-01	.235e-01	.135e-01	.977e-02	.107e-01	.131e-01
	.463e-02	.399e-02	.333e-02	.275e-02	.322e-02	.375e-02
	-.114e-02	-.163e-02	-.129e-02	-.896e-03	-.840e-03	-.836e-03
Jun	.527e-01	.215e-01	.119e-01	.851e-02	.898e-02	.110e-01
	.452e-02	.370e-02	.283e-02	.218e-02	.275e-02	.299e-02
	-.720e-03	-.158e-02	-.107e-02	-.621e-03	-.568e-03	-.600e-03
Jul	.445e-01	.186e-01	.112e-01	.855e-02	.843e-02	.951e-02
	.365e-02	.310e-02	.245e-02	.189e-02	.205e-02	.267e-02
	.197e-03	-.107e-02	-.934e-03	-.577e-03	-.476e-03	-.785e-03
Aug	.979e-01	.278e-01	.138e-01	.103e-01	.101e-01	.121e-01
	.327e-02	.282e-02	.218e-02	.207e-02	.239e-02	.271e-02
	-.118e-02	-.580e-03	-.641e-03	-.492e-03	-.414e-03	-.355e-03
Sep	.221	.706e-01	.209e-01	.146e-01	.174e-01	.239e-01
	.336e-02	.352e-02	.238e-02	.233e-02	.305e-02	.389e-02
	-.230e-02	-.184e-02	-.981e-03	-.738e-03	-.689e-03	-.101e-02
Oct	.340	.194	.592e-01	.407e-01	.628e-01	.792e-01
	.290e-02	.388e-02	.376e-02	.395e-02	.428e-02	.507e-02
	-.411e-02	-.443e-02	-.209e-02	-.165e-02	-.198e-02	-.224e-02
Nov	.187	.170	.928e-01	.862e-01	.942e-01	.925e-01
	.850e-02	.752e-02	.724e-02	.752e-02	.794e-02	.801e-02
	-.311e-02	-.335e-02	-.278e-02	-.272e-02	-.229e-02	-.163e-02
Dec	.163	.184	.131	.126	.126	.115
	.701e-02	.589e-02	.625e-02	.678e-02	.689e-02	.742e-02
	-.146e-02	-.177e-02	-.265e-02	-.228e-02	-.123e-02	-.291e-03

Exhibit 2.1

Model Coefficients For α

Month	Hour Period 06-08	09-11	12-14	15-17	18-20	21-23
Jan	1.05	1.01	1.13	1.15	1.16	1.12
	-.920e-02	-.872e-02	-.923e-02	-.952e-02	-.105e-01	-.969e-02
	-.635e-02	-.600e-02	-.864e-02	-.803e-02	-.522e-02	-.698e-02
Feb	1.02	1.05	1.21	1.21	1.22	1.16
	-.944e-02	-.892e-02	-.971e-02	-.103e-01	-.108e-01	-.104e-01
	-.330e-02	-.564e-02	-.926e-02	-.613e-02	-.443e-02	-.419e-02
Mar	1.14	1.26	1.36	1.33	1.37	1.34
	-.960e-02	-.983e-02	-.102e-01	-.103e-01	-.111e-01	-.105e-01
	-.629e-02	-.105e-01	-.110e-01	-.924e-02	-.708e-02	-.954e-02
Apr	1.21	1.32	1.16	1.08	1.19	1.22
	-.102e-01	-.923e-02	-.937e-02	-.956e-02	-.107e-01	-.113e-01
	-.649e-02	-.159e-01	-.399e-02	.483e-03	-.560e-03	.143e-02
May	1.27	1.33	1.09	1.05	1.10	1.18
	-.876e-02	-.839e-02	-.479e-02	-.566e-02	-.762e-02	-.900e-02
	-.193e-01	-.352e-01	-.298e-01	-.218e-01	-.122e-01	-.977e-02
Jun	1.33	1.41	1.23	1.21	1.25	1.36
	-.812e-02	-.595e-02	-.538e-02	-.803e-02	-.913e-02	-.109e-01
	-.265e-01	-.418e-01	-.325e-01	-.145e-01	-.119e-01	-.661e-02
Jul	1.43	1.47	1.14	1.04	1.19	1.43
	-.849e-02	-.589e-02	-.474e-02	-.766e-02	-.968e-02	-.109e-01
	-.294e-01	-.457e-01	-.304e-01	-.588e-02	-.246e-02	-.111e-01
Aug	1.13	1.50	1.28	1.14	1.30	1.46
	-.100e-01	-.833e-02	-.872e-02	-.865e-02	-.101e-01	-.120e-01
	-.245e-02	-.291e-01	-.110e-01	-.316e-02	-.516e-02	-.506e-02
Sep	.829	1.26	1.45	1.31	1.40	1.37
	-.722e-02	-.858e-02	-.898e-02	-.954e-02	-.111e-01	-.126e-01
	-.232e-02	-.160e-01	-.201e-01	-.986e-02	-.670e-02	.400e-02
Oct	.679	.900	1.28	1.34	1.26	1.05
	-.562e-02	-.743e-02	-.933e-02	-.107e-01	-.114e-01	-.107e-01
	-.353e-02	-.568e-02	-.161e-01	-.108e-01	-.243e-02	.468e-02
Nov	.887	.938	1.11	1.10	1.11	.978
	-.823e-02	-.822e-02	-.891e-02	-.924e-02	-.985e-02	-.852e-02
	-.311e-02	-.455e-02	-.878e-02	-.664e-02	-.437e-02	-.525e-02
Dec	1.04	1.02	1.12	1.12	1.13	1.06
	-.916e-02	-.887e-02	-.968e-02	-.963e-02	-.102e-01	-.903e-02
	-.637e-02	-.616e-02	-.593e-02	-.690e-02	-.622e-02	-.876e-02

Exhibit 2.2

Model Coefficients For β

$$F_j(x;\ \alpha_j,\ \beta_j) = 1 - e^{-\alpha_j x^{\beta_j}}$$

where $\alpha_j = \gamma_0 + \gamma_1 * EL_j + \gamma_2 * AE_j$

$\beta_j = \delta_0 + \delta_1 * EL_j + \delta_2 * AE_j$.

EL_j and AE_j are transformed and scaled values for the elevation at station j, and the average elevation at 20 equispaced locations on a circle of radius 20 kilometers and centered at the station. EL is the cube of the elevation in feet, divided by 10^9. AE is the cube of the average elevation in feet, divided by 10^9. The constants γ_0, γ_1 γ_2, δ_0, δ_1, δ_2 are determined by minimizing the espression

$$\sum_j \sum_i [E_j(x_i) - F_j(x_i;\ \alpha_j,\ \beta_j)]^2$$

That is, the constants are those which minimize the sum of squares of the differences between the empirical and model probabilities over all stations and all distance for which data are available. The constants γ_0, γ_1 and γ_2 are given in Exhibit 2.1, "Model Coefficients for Alpha," and the constants δ_0, δ_1, δ_2 are given in Exhibit 2.2, "Model Coefficients for Beta."

Month	Hour Period 06-08	09-11	12-14	15-17	18-20	21-23	ALL
Jan	.118	.111	.100	.099	.115	.150	.117
Feb	.102	.089	.081	.081	.097	.132	.099
Mar	.092	.086	.079	.074	.084	.098	.086
Apr	.089	.078	.066	.064	.072	.080	.075
May	.085	.060	.040	.038	.052	.069	.060
Jun	.085	.058	.041	.046	.057	.074	.062
Jul	.087	.055	.036	.041	.056	.074	.061
Aug	.099	.069	.056	.050	.062	.088	.073
Sep	.105	.079	.067	.060	.079	.097	.083
Oct	.096	.085	.078	.078	.102	.126	.096
Nov	.094	.090	.074	.083	.102	.125	.096
Dec	.104	.102	.096	.097	.107	.151	.111
ALL	.097	.082	.071	.070	.085	.109	.087

Exhibit 2.3

RMS By Month and Hour (Sample Reuse)

The models were evaluated by sample reuse. The method of sample reuse may be briefly described as follows. If there are n stations, n separate models are developed. For each model, one station is used as the evaluation set with the remaining n-1 used as the calibration set. For each station, the root mean square (RMS) error of the modeling procedure is obtained. The RMS values for the individual stations are then pooled to obtain an overall RMS error for the procedure. The RMS values by month and hour are given in Exhibit 2.3.

Exhibit 2.4 gives the RMS for each station, averaged over all months and hour periods. The first column of RMS values gives the RMS values from sample reuse, that is, when all of the other stations were used in the model building process, and the specified station was used as the "evaluation set." The RMS values indicate how well that visibility would have been predicted if the station had been a station "for which no records exist," using a model based on all the other stations. For 54 out of the 60 stations, prediction would have been good. The stations which would not have been well predicted are Baumholder, Siegenberg, Kahler Asten, Plezen, Feldberg and Grosser Falk. The latter six stations have previously presented modeling problems. The second

WMO	Station	LAT	LONG	Sample Re-use RMS	Full Fit RMS
10616	HAHN AB	49.95	-7.27	.062	.048
10610	BITBURG AB	49.95	-6.57	.044	.042
10614	RAMSTEIN AB	49.43	-7.58	.051	.052
10607	SPANGDAHLEM AB	49.97	-6.70	.057	.055
10384	TEMPELHOF APRT	52.47	-13.40	.058	.058
10755	ANSBACH AAF	49.32	-10.63	.068	.064
10544	FULDA AAF	50.53	-9.63	.045	.046
10869	ERDING AS	48.32	-11.93	.064	.062
10765	FEUCHT AAF	49.38	-11.18	.046	.045
10618	BAUMHOLDER AAF	49.65	-7.30	.153	.154
10626	BAD KREUZNACH AAF	49.87	-7.88	.071	.071
10971	BAD TOLZ AAF	47.77	-11.60	.121	.102
10714	ZWEIBRUCKEN AB	49.22	-7.40	.041	.040
10633	WIESBADEN AB	50.05	-8.33	.036	.036
10633	FINTHEN AAF	49.97	-8.15	.048	.047
10763	FURTH AAF	49.50	-10.95	.085	.085
10642	HANAU AAF	50.17	-8.95	.037	.036
10852	GABLINGEN AAF	48.45	-10.87	.084	.082
10653	GIEBELSTADT AUX AF	49.67	-9.88	.060	.060
10687	GRAFENWOHR AAF	49.70	-11.95	.063	.066
10734	HEIDELBERG AAF	49.40	-8.65	.034	.034
10752	ILLESHEIM AAF	49.47	-10.38	.084	.084
10659	KITZINGEN AAF	49.75	-10.20	.051	.052
10763	NURNBERG	49.50	-11.08	.062	.063
10729	COLEMAN AAF	49.57	-8.47	.078	.077
10657	WERTHEIM AAF	49.77	-9.48	.095	.096
10745	SCHWAEBISCH HALL AA	49.17	-9.78	.046	.046
10712	SEMBACH AB	49.52	-7.87	.037	.037
10862	SIEGENBERG GUNNERY	48.75	-11.80	.140	.140
10738	ECHTERDINGEN ARPT	48.68	-9.22	.056	.054
10501	AACHEN, DL	50.78	-6.12	.081	.080
10224	BREMEN, GER	53.05	-8.80	.043	.043
10305	LINGEN, GER	52.52	-7.33	.040	.040
10427	KAHLER ASTEN, GER	51.18	-8.50	.204	.188
11157	AIGEN ENNSTAL, OS	47.53	-14.15	.056	.053
11448	PLEZEN/DOBRA, CZ	49.67	-13.30	.130	.131
10900	BREMGARTEN, GER	47.90	-7.63	.044	.041
10948	OBERSTDORF, GER	47.40	-10.30	.097	.080
10929	KONSTANZ, GER	47.68	-9.20	.037	.036
11120	INNSBRUCK, OS	47.27	-11.37	.069	.074
11150	SALZBURG, OS	47.80	-13.02	.048	.047
10893	PASSAU, GER	48.58	-13.50	.056	.055
10908	FELDBERG, GER	47.87	-8.02	.280	.069
10921	NEUHAUSEN, DL	47.98	-8.92	.072	.063
10510	NURBURG, GER	50.33	-6.97	.085	.077
10515	KOBLENZ, DL	50.35	-7.60	.066	.068
10532	GIESSEN, GER	50.57	-8.72	.059	.060
10542	MERSFELD, DL	50.87	-9.72	.059	.060
10658	KISSINLEN, DL	50.20	-10.10	.039	.040
10671	COBURG, GER	50.27	-10.97	.038	.039
10685	HOF, GER	50.32	-11.90	.048	.057
10704	BERUS, GER	49.27	-6.70	.056	.054
10727	KARLSRUHE, GER	49.02	-8.40	.041	.041
10742	OHRINGEN, GER	49.20	-9.53	.049	.049
10788	STAUBING, GER	48.82	-12.60	.084	.085
10791	GROSSER FALK, GER	49.08	-13.30	.259	.064
10805	LAHR, GER	48.37	-7.85	.074	.073
10837	LAUDHEIM, GER	48.22	-9.93	.053	.056
10953	KAUFBEUREN, DL	47.87	-10.63	.069	.062
10875	MUHLDORF, GER	48.25	-12.55	.053	.053

Exhibit 2.4

RMS By Station

column of RMS values are the RMS values from a full regression for the model using all 60 stations. Two things are remarkable. First, Feldberg and Grosser Falk show up as remarkably influential in the regression. This suggests that for both stations there is a prediction variable that, if used in the model, would have greatly improved prediction for these stations. Second, for the other 58 stations the RMS values are little changed. This suggests (at least to the authors) that the use of the variables of elevation and relative elevation results in a quite good model for predicting visibility probabilities for specific months and hour periods.

Exhibit 2.5 gives the RMS values that resulted when all 60 stations were used to generate the model coefficients given in Exhibits 2.1 and 2.2. These are the RMS values over all stations and are internal estimates. Exhibit 2.3, and not Exhibit 2.5, should be used as a measure of the ability to predict visibility for locations in Germany where no visibility records are available. One reason for retaining Exhibit 2.5 in this report is to indicate that RMS values obtained from the regressions producing the model are overestimates of the model accuracies and instead an independent estimate should be used. It is worth repeating that sample reuse does provide independent estimates of the model accuracies! This is accomplished despite the fact that we do not have two sets of data--one for the determination of the model accuracies and one for the evaluation of the model.

	Hour Period						
Month	06-08	09-11	12-14	15-17	18-20	21-23	ALL
Jan	.106	.098	.086	.083	.100	.140	.104
Feb	.088	.074	.061	.060	.079	.120	.083
Mar	.075	.066	.056	.049	.060	.085	.066
Apr	.069	.058	.041	.035	.042	.053	.051
May	.068	.048	.031	.026	.030	.048	.044
Jun	.070	.048	.031	.026	.032	.047	.045
Jul	.071	.046	.029	.023	.030	.047	.044
Aug	.082	.053	.036	.028	.035	.063	.053
Sep	.095	.063	.050	.036	.054	.073	.065
Oct	.087	.071	.057	.052	.081	.111	.079
Nov	.082	.076	.064	.065	.087	.115	.083
Dec	.089	.087	.079	.080	.090	.142	.097
ALL	.083	.068	.055	.051	.065	.094	.071

Exhibit 2.5
RMS By Month and Hour (Overall Fit)

4. USE OF THE MODEL

Problem 1: Find α and β for Koblenz for March 0900-1100 hours and use it to find probability of visibility less than 1.5, 2.0 and 6.0 miles. The elevation of Koblenz is 318 feet and the average elevation of the surrounding area is 900 feet.

Solution: From Exhibits 2.1 and 2.2 we have

$$\alpha = .0848 + ..00580\ (318)^3/10^9 - .00178\ (900)^3 10^9 = .0837 \text{ and}$$

$$\beta = 1.26 - .00983\ (318)^3/10^9 - .0105\ (900)^3/10^9 = 1.252$$

Estimated probability of visibility less than or equal to x miles is

$$1 - e^{-.0837x^{1.252}}$$

The probabilities of visibilities less than or equal to distiances of 1.5, 2.0 and 6.0 miles are thus estimated as .130, .181 and .546. These compare to the RUSSWO values of .90, .156 and .525.

Problem 2: Find and for Berus for September 1800-200 hours and use it to find probabities of visibility less than 1.5, 2.0 and 6.0 miles. The elevation of Berus is 1204 feet and the average elevation of the surrounding area is 1050 feet.

Solution: From Exhibits 2.1 and 2.2 we have

$$\alpha = .0174 + .00305 \times (1204)^3/10^9 - .000689\ (1050)^3/10^9 = .0219 \text{ and}$$

$$\beta = 1.40 - .0111 \times (1204)^3/10^9 - .00670\ (1050)^3/10^9 = 1.373.$$

Estimated probability of visibility less than or equal to x miles is

$$1 - e^{-.0219x^{1.373}}$$

The probabilities of visibilities less than or equal to distances of 1.5, 2.0 and 6.0 miles are thus estimated as .037, .055 and .226. These compare to the RUSSWO values of .019, .027 and .140.

5. SUMMARY

Given only month, time of day, and the elevations for an arbitrary location in Germany, models are developed for estimating probabilities of visibilities less than a specified distance. The two-parameter Weibull model is used. For a given month and hour period, the parameters are functions of elevation and average elevation of the surrounding region. The models are developed using a somewhat sophisticated nonlinear, least squares regression program. Evaluation is accomplished using "sample-reuse," a relatively new technique in which a valid independent evaluation of the model is made using the same observations used to develop the model.

6. ACKNOWLEDGEMENT

This research was supported by Contract Nos. F19628-80-C-0004 and F19628-83-K0001 with the Cambridge Research Laboratories of the U.S. Air Force.

7. REFERENCE

Somerville, Paul N. and Steven J. Bean, "Probability Modeling of Weather Elements," Sixth Conference on Probability and Statistics in Atmospheric Sciences, Banff, Alberta, Canada, October 9-12, 1979, pp. 173-175.

ANALYSIS OF OBSERVED WIND DIRECTIONS BY STATISTICS OF DIRECTIONAL DATA

Atsushi Hiroh* and Masashi Itoh**

*Kanto Gakuin University
4834 Mutsuura, Kanazawa-ku
Yokohama, JAPAN 236
**Tokyo Metropolitan Research Institute for
Environmental Protection
2-7-1 Yuuraku-cho, Chiyoda-ku
Tokyo, JAPAN 100

1. INTRODUCTION

Among studies related to wind profile, many of them treat of wind speed, and only few studies treat of wind direction. The reason is considered to be the data description of wind directions. Wind directions are usually measured in degrees, and this causes the difficulties of the analyzation. For example, let the observed wind directions in a sample of size 2 be 5° and 355°, the arithmetic mean will give the value of 180°, which is the direction opposite to the real mean direction. This shows that the mean direction of the angular data depends heavily on the choice of the zero directions. If we take the zero direction as y-axis in place of the x-axis, the observed values reduce to 275° and 265° respectively. Then, the mean direction will be 270°, and it presents the real mean direction. Thus for analyzing the angular data we need another methodology, by which the mean direction is obtained independently of the choice of the zero direction.

The data of wind directions are often analyzed by the wind rose. However, wind rose is given only at each monitoring point and at each period, and it is difficult to compare statistically, with each other.

For the analyzation of the data of wind directions, statistics of directional data is applicable. It can treat of the angular data, and can give the method to calculate the mean directions, circular variance and higher order moment. By this methodology we can analyze statistically the data on the circle in the same manner as the data on the line. The details of statistics of directional data are discussed by Mardia (1972).

The statistical inference and test are also possible by statistics of directional data, assuming the distribution on the circle. Test makes it possible to compare the characterstics between two given wind roses, and we can use these characteristics in place of wind rose as the values which represent the local wind.

For the distribution on the circle, the von Mises distribution and the wrapped normal distribution are commonly used (Collett, 1981; Gumbel, 1954; Baba, 1981). The von Mises distribution plays an important role on the circle which is similar to that of normal distribution on the line.

In this paper, the statistical methodology of directional data and its applications to the wind directions, which are observed near Tokyo area, and utility of the methodology are described.

2. STATISTICAL METHODOLOGY OF DIRECTIONAL DATA

2.1 The Mean Direction and the Circular Variance

Statistics of directional data is discussed by Mardia. For the analysis of wind directions, we show the methodology to estimate the mean direction and the circular variance cited from Mardia.

Let P_i be the point on the circumference of the unit circle corresponding to the angle θ_i $(i=1,\ldots,n)$. Then the mean direction $\bar{\theta}$ is defined to be the direction of the resultant of the unit vectors $\overrightarrow{OP}_1,\ldots,\overrightarrow{OP}_n$, and the mean direction is obtained by

$$\bar{\theta} = \arctan(C/S) \tag{1}$$

where

$$\bar{C} = \frac{1}{n}\Sigma\cos\theta_i \ , \ \bar{S} = \frac{1}{n}\Sigma\sin\theta_i \tag{2}$$

Unlike the mean on the line, the mean direction on the circumference of the unit circle is invariant under a change of the zero direction.

The length of the resultant is $R=n\bar{R}$, where

$$\bar{R} = (\bar{C}^2 + \bar{S}^2)^{1/2} \tag{3}$$

The circular variance S is defined as

$$S = 1 - \frac{1}{n}\Sigma\cos(\theta_i - \bar{\theta}) \tag{4}$$

This is the value of dispersion about $\bar{\theta}$. By the eq.(2), we have

$$S = 1 - \bar{R} \tag{5}$$

and circular variance S is also invariant under a change of the zero direction. Using $R=n\bar{R}$, the circular variance is described as

$$S = 1 - R/n \tag{6}$$

So that the value of the circular variance is

$$0 \leq S \leq 1 \tag{7}$$

If the observed directions are tightly clustered about the mean direction, the circular variance S will be nearly zero while the mean resultant length $\bar{R}$ will be 1. On the other hand, if the directions are widely dispersed, S will be nearly 1, while R will be small.

Let us consider the characteristic function

$$\phi_p = E(\exp(ip\theta)) = \int_0^{2\pi} \exp(i\theta)dF(\theta) \quad (8)$$

(p=0,±1,±2,...), where $F(\theta)$ is the circular distribution function of θ.
Let us write

$$\phi_p = \alpha_p + i\beta_p = \rho_p \exp(iw_p) \quad (9)$$

where

$$\alpha_p = E(\cos p\theta), \quad \beta_p = E(\sin p\theta) \quad (10)$$

and for p=1,

$$\alpha_1 = \alpha, \ \beta_1 = \beta, \ \rho_1 = \rho, \ w_1 = \mu$$

Then the pth central trigonometric moment which can be defined as

$$\mu_p = E(\exp(ip(\theta-\mu))) = \phi_p \exp(-ip\mu)$$
$$= \bar{\alpha}_p + i\bar{\beta}_p \quad (11)$$

We have the relations

$$\bar{\alpha}_p = \rho_p \cos(w_p - p\mu), \ \beta_p = \rho_p \sin(w_p - p\mu) \quad (12)$$

In particular case of p=1, we have

$$\bar{\alpha}_1 = \rho, \quad \bar{\beta}_1 = 0$$

The mean direction of θ and the resultant length ρ are obtained for p=1 in eq.(9) as

$$\phi_1 = \rho \exp(i\mu) \quad (13)$$

And the circular variance of θ is

$$V = 1 - \rho = 1 - E(\cos(\theta-\mu)) \quad (14)$$

The circular variance has the similar property to that of σ^2 on the line. That is V=0 if and only if the distribution is concentrated at one single point $\theta=\mu$. And V=1 shows that the distribution has maximum variation and there is no concentration around any particular direction.

2.2 The von Mises Distribution

In statistical inference on the circle, the von Mises distribution plays an important role, whose importance on the circle is somewhat similar to that of the normal distribution on the line and it sometimes is called as "circular normal distribution". The wrapped normal distribution is also important in the same sense. These distributions have periodicity which is necessary for the distribution defined on the circle. That is, for the probabilistic density function $f(\theta)$, the following periodicity is required.

$$f(\theta+2\pi) = f(\theta), \ \int_0^{2\pi} f(\theta)d\theta = 1 \quad (15)$$

In this paper, we use the von Mises distribution as the circular distribution, because the distribution of wind directions is considered to be distributed in von Mises distribution and its treatment is easier than that of the wrapped normal distribution.

A circular random variable θ is said to have a von Mises distribution $M(\mu,\kappa)$ if its p.d.f. is given by

$$f(\theta;\mu,\kappa) = \frac{1}{2\pi I_0(\kappa)} \exp(\kappa \cos(\theta-\mu))$$
$$0 < \theta \leq 2\pi, \quad \kappa > 0, \ 0 \leq \mu < 2\pi \quad (16)$$

where $I_0(\kappa)$ is the modified Bessel function of the first kind and order zero and expressed as

$$I_0(\kappa) = \Sigma \frac{1}{(r!^2)} \left(\frac{1}{2}\kappa\right)^{2r} \quad (17)$$

The von Mises distribution has two parameters, the mean direction θ and the concentration parameter κ.

The distribution is unimodal and is symmetric about the mean direction. The mode is at $\theta=\mu$ and antimode is at $\theta=\mu+\pi$. The ratio of the density at the mode to the density at the antimode is given by $\exp(2\kappa)$ so that the larger the value of κ, the greater is the clustering around the mode. That is, for large value of κ, the circular variance will be small and vice versa. For large values of κ, the points reduce to $\pm\kappa^{-1/2}$ which are comparable to the inflection points $\pm\sigma$ for normal distribution with mean zero.

Let us assume that the distribution on the circle is distributed in the von Mises distribution, then the characteristic function is obtained as

$$\phi_p = \alpha_p = I_p(\kappa)/I_0(\kappa) \quad (18)$$

where $I_p(\kappa)$ is the modified Bessel function of the first kind and of the pth order.
In particular, we have for $\mu \neq 0$,

$$\alpha = A(\kappa)\cos\mu, \quad \beta = A(\kappa)\sin\mu, \ \rho = A(\kappa),$$
$$V = 1 - A(\kappa) \quad (19)$$

where

$$A(\kappa) = I_1(\kappa)/I_0(\kappa) \quad \left(= I_0'(\kappa)/I_0(\kappa)\right)$$

Thus the mean direction, the resultant length and the concentration parameter is estimated as

$$\hat{\mu} = \bar{\theta}, \ \hat{\rho} = \bar{R}, \ \hat{\kappa} = A^{-1}(\bar{R}) \quad (20)$$

2.3 Test for Samples from von Mises Populations

Assuming that the wind directions are distributed in von Mises distribution $M(\mu,\kappa)$, the test procedure is as follows.

Let $\theta_{i1}, \theta_{i2}, \ldots, \theta_{in_i}$ (i=1,2) be independent random samples of size n_1, n_2 from $M(\mu_i,\kappa_i)$ for i=1,2 respectively. Let us consider to test

$$H_0 : \mu_1 = \mu_2 = \mu$$

against

$$H_1 : \mu_1 \neq \mu_2$$

where μ and $\kappa(\kappa_1=\kappa_2)$ are unknown.
Under the hypothesis H_0, p.d.f. of u can be written as

$$R/(\pi h_n(R)(u^2-R^2)^{1/2}) \int_{-R}^{R} h_{n_1}\left(\frac{1}{2}(u+v)\right) h_{n_2}\left(\frac{1}{2}(u-v)\right)$$
$$/(R^2-v^2)^{1/2}dv \qquad (R < u < n) \quad (21)$$

where $u=R_1+R_2$, $v=R_1-R_2$ and

$$h_n(R) = R\int_0^\infty uJ_0(Ru)J_0^n(u)du$$

$J_0(x)$ is the standard Bessel function of zero order.

We have test-statistics by approximation of eq.(21). For large κ, we have approximation from Watson and Williams'

$$\chi_1^2 = 2\kappa(R_1+R_2-R) \qquad (22)$$

For $\kappa > 2$,

$$F_{1,n-2}=(1+\frac{1}{8\kappa})((n-2)(R_1+R_2-R)/(n-R_1-R_2)) \qquad (23)$$

χ_1^2 is chi-square distribution with degree of freedom 1, and $F_{1,n-2}$ is F distribution with degree of freedom 1 and n-2.

In the case of testing κ, the following procedure can be used.

$$H_0 : \kappa_1 = \kappa_2 = \kappa$$

against

$$H_1 : \kappa_1 \neq \kappa_2$$

where μ_1, μ_2 and κ are unknown.
Under the hypothesis H_0, approximation of the distribution of R is obtained. Therefore we can test H_0 by follwing test-statistics.
For $\bar{R} < 0.45$,

$$Z=(2/\sqrt{3})(g_1(2\bar{R}_1)-g_1(2\bar{R}_2))/((n_1-4)^{-1}+(n_2-4)^{-1})^{1/2} \qquad (24)$$

where Z is distributed as N(0,1) and

$$g_1(\bar{R}_i) = \sin^{-1}(0.61237\bar{R}_i)$$

For $0.45 \leqq \bar{R} \leqq 0.7$,

$$Z = (g_2(\bar{R}_1) - g_2(\bar{R}_2))/(0.89325((n_1-3)^{-1}+(n_2-3)^{-1})^{1/2}) \qquad (25)$$

where Z is distributed as N(0,1) and

$$g_2(\bar{R}_i)=\sinh^{-1}((\bar{R}_i-1.0894)/0.25789)$$

For $\bar{R} > 0.7$,

$$F_{n_1-1,n_2-1}=((n_1-R_1)(n_2-1))/((n_2-R_2)(n_1-1)) \qquad (26)$$

F is distributed as F distribution with degree of freedom n_1-1 and n_2-1.

3. DATA DESCRIPTION

Data used in this study are of two kind. The first kind of data are wind directions which are observed at ground level in 1980. Hourly averaged wind directions are observed at 35 monitoring stations in Tokyo. The data are used for calculation of the mean wind directions and the circular variances in each month.

The second kind of data used are observed as a special program to investigate the influences of meteorological factors on air pollutions in south Kanto area (Meteorological Agency of Japan, 1977). The following meteorological factors are measured: temperature profile, wind profile, trajectory, structure of see breeze, etc. Among these factors, we use the data of upper wind observations which are observed in four seasons.

In each season the observation period is one week, and wind directions are observed hourly in degrees by pilot balloon. The observation period in each season is as follows.

(1) 11/25/74 - 12/01/74 (Fall)
(2) 08/01/75 - 08/07/75 (Summer)
(3) 12/10/75 - 12/16/75 (Winter)
(4) 03/03/76 - 03/09/76 (Spring)

The number of data by pilot balloon observation are 18 in fall, 25 in summer, 21 in winter and 26 in spring. Pilot balloon is observed at the height of every 100m from 50m to 1950m. For the application of statistics of directional data, we use the pilot balloon observation data to calculate the mean wind directions, the circular variances and to test at each height.

4. APPLICATIONS

In previous section we describe the statistical methodology of directional data, and in this section it is applied to the wind direction data. First data are observed at ground level in Tokyo in 1980. The mean directions and the circular variances are calculated in each month and the results at 4 monitoring stations, i.e. Joto, Minato, Harumi and Tochomae, are shown in Fig.1.

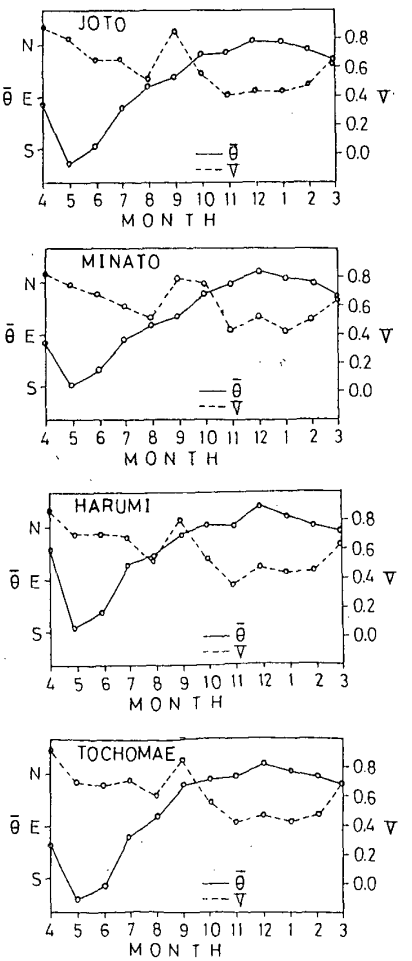


Fig. 1. The mean direction and the circular variance in each month for the data observed at ground level in Tokyo (1980)

The monthly variation of the mean wind directions at 4 monitoring stations show close resemblances. The mean direction in May is near south and it changes slowly from east to north along month. In 1980 it was cool in summer, so that the mean directions in summer are almost between east and north.

The circular variances shows the same pattern at the data of each monitoring station. In April its value is about 1, which means the variance is very large and the wind directions change so often. And then it slowly reduces except in September. During winter its value reduces about 0.4 and it increases to about 0.7 in March. That is, the circular variance is hight in spring and fall and low in winter. The data at the other monitoring stations also vary similar to those at 4 stations in Fig.1.

It is also possible to apply statistical test to these data at ground level, but in many cases the distributions of observed wind directions are not unimodal. For the multi-modal distribution the mixture is required and test-statistics become more complicated, so that in this paper we do not treat of test for the data observed at ground level.

Next data are the upper wind observations in south Kanto area in four seasons. Wind roses of summer data at each height at Machida are shown in Fig. 1. Intuitively from Fig. 1 we can see that the mean direction changes from south to west and the circular variance slowly increases as height increases.

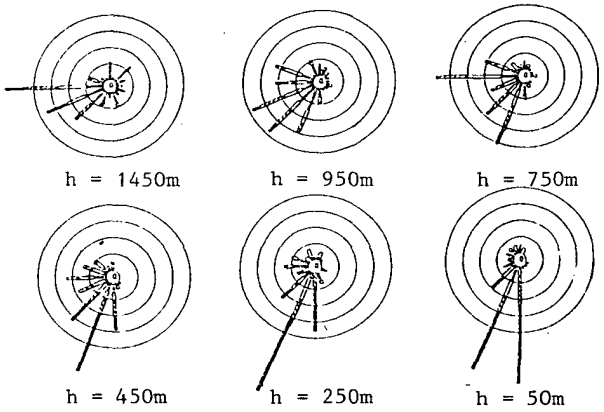


Fig. 2. Wind rose at each height at Machida (08/01/75-08/07/75)

The mean direction and the circular variance at each height are shown in Fig. 3. In summer the mean direction changes slowly from south to west and the circular variance increases from about 0.1 to 0.4 as height increases. These are the same characteristics obtained from wind rose at each height. It shows that we can treat of the mean direction and the circular variance as the representative values in place of wind rose which describe the characteristics of wind direction data. And these values make it possible to treat of wind direction data quantitatively.

In spring the mean direction rotates from north to west as height increases, i.e. wind direction has clockwise rotation. In other places the results are the same. The circular variance has maximum value at height about 200m and then

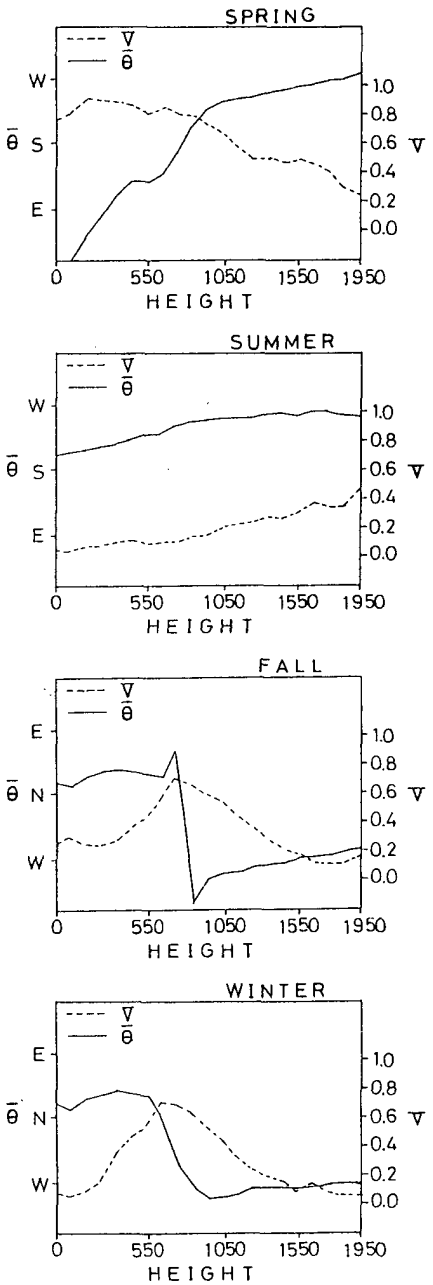


Fig. 3. The mean direction and the circular variance at each height for each season at Machida

it slowly decreases as height increases.

In summer the variation of the mean direction and the circular variance is small. The circular variances show almost the same values and the mean direction changes from south to west very slowly. It indicates that the wind is stable in summer.

In fall the mean direction shows the rapid change at height about 800m. It turns from northeast to northwest and the circular variance takes maximum value at the same height. This is becuase the boundary layer is considered to be disconnected at this height and there the atmosphere is exceedingly unstable.

In winter the circular variance shows the

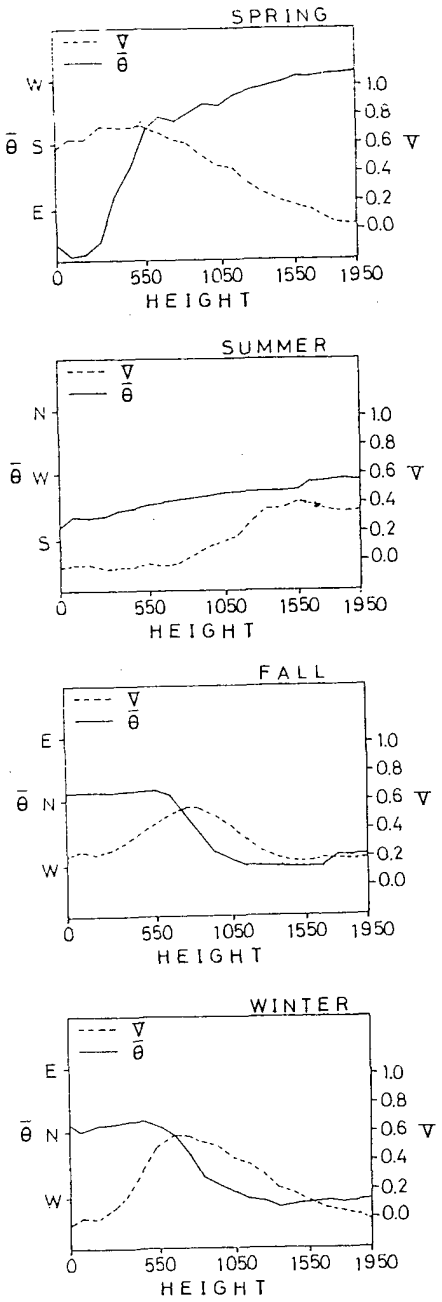


Fig. 4. The mean direction and the circular variance at each height for each season at Ootemachi

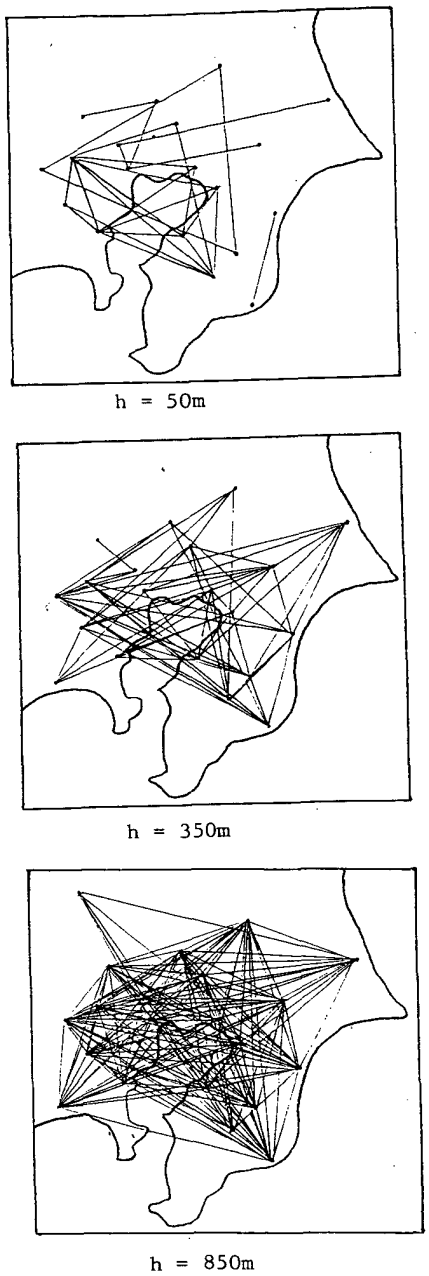


Fig. 5. Results of test where point is monitoring point and straight line shows that the mean direction and the concentration parameter are not significantly different between two points at the end of each line (12/10/75 - 12/16/75)

same pattern as that in fall, but the variatiion of the mean direction is small. This is because the date of observation in fall and winter is very close.

We also show the mean direction and the circular variance at Ootemachi in Fig. 4. In Fig. 4 the variations are almost the same as in Fig. 3. In other places the figures of the variations of the mean direction and the circular variance have resemblances to these figures.

Assuming that the distribution of wind directions are distributed in von Mises distribution $M(\mu,\kappa)$, we can apply test for the concentration parameter κ and the mean wind direction $\bar{\theta}$ between two arbitary monitoring points or heights. Test is performed for the data of upper wind observation by test procedure which is described in previous section. Wind roses of upper observation data in Fig. 2 show that the distribution of the wind directions is almost unimodal, and it shows that we do not need to consider the mixture when testing.

Results of test are shown in Fig. 5, where

Table 1. Number of cases in which the mean direction and the circular variance are not significantly different

Season		Spring		Summer		Fall		Winter	
No	Height	N	%	N	%	N	%	N	%
1	SURFACE	141	43	38	13	34	22	35	17
2	50m	152	47	28	9	63	41	27	13
3	150m	163	50	27	9	77	50	40	19
4	250m	152	47	25	8	78	51	45	21
5	350m	181	56	26	9	87	57	69	33
6	450m	192	59	29	10	87	57	97	46
7	550m	191	59	33	11	86	56	136	65
8	650m	198	61	32	11	97	63	170	81
9	750m	215	66	33	11	106	69	182	87
10	850m	222	68	28	9	98	64	187	89
11	950m	245	75	37	12	93	61	182	87
12	1050m	241	74	33	11	88	58	177	84
13	1150m	215	66	41	14	93	61	170	81
14	1250m	180	55	48	16	105	69	136	65
15	1350m	193	59	52	17	112	73	152	72
16	1450m	197	61	53	18	89	58	144	69
17	1550m	203	62	59	20	68	44	125	60
18	1650m	216	66	76	25	64	42	136	65
19	1750m	184	57	91	30	62	41	90	43
20	1850m	155	48	95	32	67	44	84	40
21	1950m	171	53	96	32	70	46	77	37

* Significant level 10%

the map covers south Kanto area and the point on the map is the monitoring point. The straight line between two points indicates that the mean wind direction and the concentration parameter are at the same time not significantly different at each point at the end of the line. In Fig. 5 results at three different height in winter are shown. The density of the line expresses the number of cases which are not significantly different. In Fig. 5 there is a tendency that the higher the height increases, the higher the density becomes. That is, in higher layer the wind is more stable. It is considered that in higher layer the influences of geostrophic wind becomes greater than in lower layer, and that the figure shows this phenomenon.

The ratio of number of significantly indifferent cases is shown in Table 1. In many cases the ratio has maximum percentage at the height about 1000m. This is related to the circular variance.

In spring the ratio is almost more than 50% and the difference of ratio is small. In summer the ratio is small and from Fig. 3,4 the circular variance is also small compared with that in other seasons. In fall the ratio is similar to that in spring. In winter the variation of the circular variance and the ratio is alike.

When the circular variance is large, the ratio is also large in lower layer and vice versa. In higher layer the circular variance is small except in summer and the ratio is not so small. This indicates that in lower layer the wind directions are mostly different at each places and that in higher layer the wind directions are not so different influenced by geostrophic wind, whether the circular variance is large or small.

5. CONSLUSIONS

In the precedent studies on wind direction many of them relate wind rose and wind directions are analyzed by wind rose. However, we can not treat of wind directions quantatively by wind roses.

In this paper we described the statistical methodology of directional data and its application to the analyzation of wind directions. And we also showed that by this methodology the treatment of wind direction data becomes much easier and that statistics by this methodology express the characteristics which wind direction data contain. That is, the mean wind direction and the circular variance can take place of wind rose. The comparison of characteristics between two arbitary places and heights can be performed by test of the mean direction and the concentration parameter.

However, our future work still remains in the treatment of test of multi-modal case, calculation of correlation coefficients, higher order of trigonometric moments and in the analyzation of the relation between obtained statistics and structure of boundary layer.

REFERENCES

Baba, Y., 1981: Statistics of angular data (in Japanese), Proc. Inst. Statis. Math., 28, 4-54.

Collett, D. and Lewis, T., 1981: Discriminating between the von Mises and wrapped normal distribution, Austral. J. Statist., 23, 73-79.

Gumbel, E. J., 1954: Applications of the circular normal distribution, J. Amer. Statist. Ass., 49, 267-297.

Mardia, K. V., 1972: Statistics of Directional Data, Academic Press.

Meteorological Agency of Japan, 1977: Report on Atmospheric Environment in South Kanto Area.

APPLICATION OF REGIONALIZED VARIABLE ANALYSIS TO CLIMATIC DATA

Kodjo Y. Amegee and Richard H. Cuenca

Department of Agricultural Engineering
Oregon State University
Corvallis, Oregon 97331, U.S.A.

1. INTRODUCTION

Regionalized variable analysis is applied when there is a quantifiable structure in the variation of a random variable. Intuitively, samples of a natural variable taken in close proximity have a higher probability of being similar in magnitude than samples taken further apart. If there is a spatial structure for the variable in relation to the coordinate system, i.e. a relationship between the magnitude of variation and distance from the sampling point, it can be described by means of a variogram.

A problem commonly encountered in analysis of hydrologic variables is interpolation of measured data at points where no measurements exist. Distribution of data values to a grid network is an example of this problem. Typical interpolating methods do not maintain the structure of variation, assuming such a structure can be quantified. A method termed kriging allows for weighting of the coefficients used in interpolation such that the observed structure of variation is maintained.

Statement of the Problem

Studies of the hydrologic balance for water resource system design, irrigation system design, and irrigation scheduling require correct determination of evapotranspiration. Evapotranspiration is a combination of water evaporated from the soil surface and water withdrawn by the roots, transported through the plant, and transpired from the leaf surface. Evapotranspiration is a function of various parameters including atmospheric conditions measured by meterological parameters, plant variety, stage of growth, and soil moisture content. To standardize the analysis, evapotranspiration is often calculated for a reference crop which is of uniform height, completely covering the ground, free of weeds, diseases and pests, and maintained at soil moisture conditions which do not limit growth. The most common reference crops are grass or alfalfa (Medicago sativa, L.). Under reference crop conditions, evapotranspiration becomes a meterological or climatic variable which varies in response to atmospheric conditions.

In applications such as water resources planning it is beneficial to be able to describe reference evapotranspiration quantitatively over large geographic or governmental areas. An example of such a presentation is indicated in Figure 1 in which isoquants of mean reference evapotranspiration (for grass) for the month of July are indicated for central California. The distances covered in Figure 1 are approximately 580 km (360 mi) in the North-South direction and 370 km (230 mi) in the East-West direction.

The most convenient means of displaying the distribution of evapotranspiration over an area such as that indicated in Figure 1 is by using computerized contour plotting routines. Such plotting programs normally require that data be available at uniformly spaced grid corners for efficient plotting of contours of equal evapotranspiration. Since meteorological stations are not generally established based on a grid system, it is necessary to interpolate between measurement sites to points of the grid system which have no data. Grid corners which are located at points where no meteorological data exist will be termed unoccupied grid corners.

The variation of evapotranspiration over a region may be considered as having a spatial structure. Measurements of evapotranspiration made in close proximity will generally be more similar quantitatively than measurements made at greater distances. In interpolating for evapotranspiration amounts at unoccupied grid corners prior to computerized contour plotting, it would be beneficial to maintain this spatial structure. A method of interpolation is therefore sought in which the spatial structure of the evapotranspiration variable determined from the measured data will be maintained for the unoccupied grid corners.

Theory of Regionalized Variables

Over the past 30 years the theory of regionalized variables has been developed mainly in reaction to stimulus of the mining industry for accurate estimation and an indication of the reliability of estimates of ore reserves. This work was originally based on practical methods used in South African coal mines by D. G. Krige (1951). A theoretical background for regionalized variable analysis has been developed by Matheron (1963) and others at the Centre de Morphologie Mathématique in Fontainebleau, France. The theory and applied methods have come to be known collectively as geostatistics. The practical estimation techniques that have been developed are termed the method of kriging in honor of the pioneer work

done by D. G. Krige (Henley, 1981).

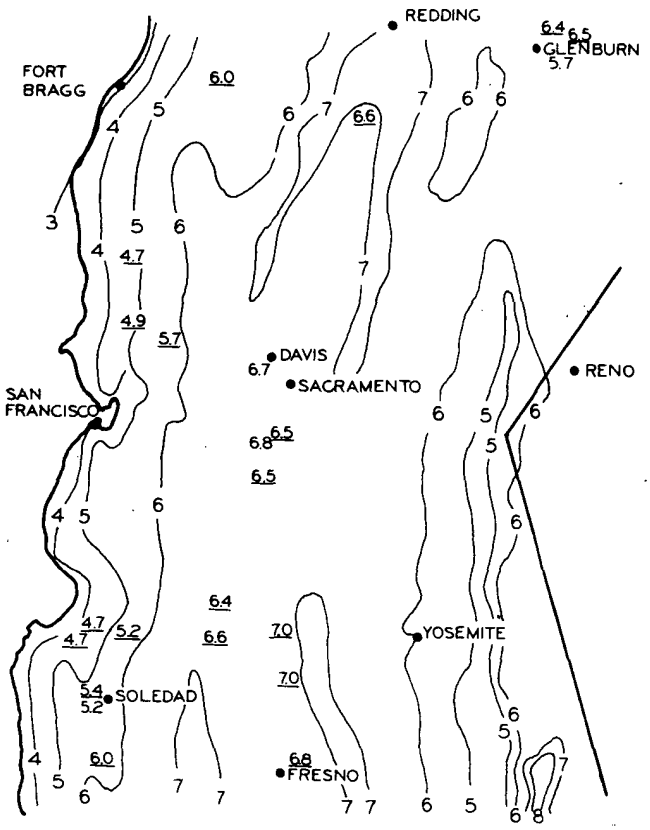


Figure 1. Monthly mean reference evapotranspiration in mm/day for grass for central California in July based on lysimetric (plain numerals) and pan evaporation (underscored numerals) data. (Taken from Cuenca et al, 1981.)

The purpose of geostatistics is to obtain a ıantitative description of the form of a surface volume sampled at discrete points. The sence of the theory of regionalized variables is use information about the auto-correlation of ta using spatial relationships which are not techniques applied in classical or Fisherian statistics (Henley, 1981). Certain methods in geostatistics may appear familiar if reference is made to time-series analysis with substitution of the space domain for the time domain.

As in the majority of physical sciences, there is an error associated with recorded climatological data over space because of the complexity of the phenomenon and observational error. An optimum discriptor of the observations will fulfill two partly contradictory objectives. It will fit exactly all recorded data points whether or not the observations include significant error since the amount of error is unknown. In giving this exact fit, it will provide an efficient and economical interpolation between the data points. At the same time, it will provide a stable description by minimizing the disturbing effect of extreme values (Henley, 1981).

Regionalized variable theory is the statistics of a particular type of variable - the regionalized variable which differs from an ordinary scaler random variable in that it has a defined spatial location. Consider two observations Z_1 and Z_2 recorded at two locations x_1 and x_2, where $x_2 = x_1 + h$ and h is the distance between measurements, termed the lag. If the true surface of measured values has the same mean value everywhere, i.e. over any subdomain, the average value of $(Z_1 - Z_2)$ would be zero. Stating this concept in terms of expected values,

$$E\,[Z(x)] = m \tag{1}$$

$$E\,[Z(x) - Z(x+h)] = 0 \tag{2}$$

Equations (1) and (2) define stationarity of the first order; the random variable Z(x) has the same mean value m regardless of location.

The covariance of Z_1 and Z_2 may be defined as the expected value of the product of deviations of the two observations from the mean,

$$E\,[\{Z(x) - m\}\,\{Z(x+h) = m\}] = C(h) \tag{3}$$

where C(h) is the covariance for two observations separated by distance h. If for every h this spatial covariance is independent of the locations of Z_1 and Z_2, the data exhibit second-order stationarity. If second-order stationarity holds, the covariance as h tends to zero approaches the variance of the random variable Z,

$$C(0) = E[\{Z(x) - m\}\,\{Z(x) - m\}] = \mathrm{var}(Z) \tag{4}$$

Full second-order stationarity may be difficult to justify in practical applications. In addition, the mean value m is not always known (and may not be constant), so that the covariance and variance may not be computed directly. An alternative statistic, not based on the assumption of second-order stationarity, may be defined which does not require the mean value. Although the expected value of Z(x) - Z(x+h) may be zero, the expected value of the square of this difference is not necessarily zero. A consistant set of assumptions may be made which are weaker than second-order stationarity and termed the intrinsic hypothesis by Matheron (1963),

$$E\,[Z(x) - Z(x+h)] = 0 \tag{2}$$

$$E[\{Z(x) - Z(x+h)\}^2] = \mathrm{var}[Z(x) - Z(x+h)] = 2\gamma(h) \tag{5}$$

where $\gamma(h)$ is a statistic known as the semivariance. The value of γ is dependent on h, the lag between observations. At h = 0, γ is necessarily zero, while at increasing h, γ will tend to increase. $\gamma(h)$ takes the following computational form,

$$\gamma(h) = \frac{1}{2\,N(k)} \sum_{i=1}^{N(k)} (Z_i - Z_{i+k})^2 \tag{6}$$

where N(k) is the number of data pairs at lag k, separated by distance $h = k\,\Delta\,x$.

If second-order stationarity also holds, there is a constant and finite mean m and the covariance C(h) and variance C(0) both exist. The semivariance, variance and covariance are related by the following,

$$\gamma(h) = C(0) - C(h) \quad (7)$$

[A certain amount of confusion has arisen regarding terminology in geostatistics. In this paper, use will be made of a combination of the terminology indicated by Hajrasuliha et al (1980) and Henley (1981). This will be done mainly to avoid the awkward use of the suffix -ogram, which has the connotation of pictorial representation in English, to represent a data point. The value of single variable computed using equation (6) will be termed the semivariance. The set of all possible solutions to equation (6) will be termed the semivariogram function. The pictorial representation of the semivariogram function (see Figure 2) will be termed the variogram.]

A representation of the relationship between the semivariance, variance, and covariance is indicated in Figure 2. In this figure, there is some structural relationship in space of the variation of the random variable given by the shape of the semivariogram function up to a limit termed the range. At the range, the semivariogram function approaches the variance of the random variable within some error criteria ε. Beyong the range, there is no spatial dependence of the random variable, i.e. the covariance is equal to zero, and application of traditional statistical techniques is recommended.

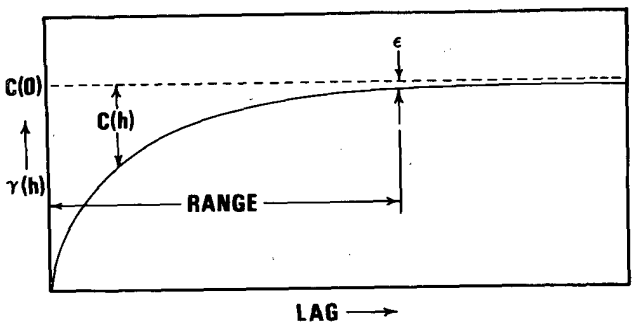


Figure 2. A variogram illustrating the range of distances over which pairs of observations are related. (Adapted from Hajrasuliha et al, 1980.)

The properties of a regionalized variable may be summarized as follows: A regionalized variable is a random function with a defined continuous distribution at each point in space. This random function has a degree of spatial continuity which may be expressed by the variance of the difference in values between observations. For practical application, the relationship between this spatial variance and distance h must be constant over the whole domain under consideration (Henley, 1981).

2. REGIONALIZED VARIABLE ANALYSIS FOR EVAPOTRANSPIRATION IN OREGON

For applications in water resource system design in the state of Oregon, it was decided to attempt to characterize evapotranspiration over the entire state using regionalized variable analysis. Because of data availability and the fact that initial analysis was to be made on a monthly basis, a temperature based evapotranspiration estimating method was chosen. The method chosen was the modification of the Blaney-Criddle method developed by the Food and Agricultural Organization (FAO) of the United Nations (Doorenbos and Pruitt, 1977). Various studies have shown this modification of the Blaney-Criddle method to be superior for monthly evapotranspiration estimates compared to other temperature based methods when local calibration using a lysimeter is not possible (Burman et al, 1983; Cuenca et al, 1981). In fact, Erpenbeck (1981) reports that the FAO modification of the Blaney-Criddle method was superior to seven other temperature based methods for monthly evapotranspiration estimates for the state of Washington which lies directly north of and adjacent to Oregon.

The FAO modification of the Blaney-Criddle method for a reference crop of grass is of the following form:

$$ET_r = C\,[P\,(0.46\,T + 8.13)] \quad (8)$$

where ET_r = monthly reference evapotranspiration for grass, mm/day

P = percentage of annual daytime hours during each day of month, percent/day

T = mean monthly temperature, °C

C = climatic calibration coefficient

Values for P are tabulated as a function of time of year and latitude. The main contribution of the FAO modification compared to earlier versions of the Blaney-Criddle methos is the refinement of the evapotranspiration estimate by the climatic calibration coefficient C. C is a function of the following parameters:

a) n/N, actual sunshine hours/maximum possible sunshine hours.

b) RH_{min}, minimum relative humidity, percent

c) U_{day}, daytime wind velocity, m/sec

Maximum sunshine hours, like the P factor, is a function of latitude and time of year and is tabulated. Daytime wind velocity required for the FAO modification is that measured at a two meter height between the arbitrary daytime hours of 0700 to 1900. A graphical solution of equation (8) can be applied with only estimated values of n/N, RH_{min}, and U_{day} in ranges of low, medium and high and is described in a number of publications (Doorenbos and Pruitt, 1977; Burman et al, 1983). A regression equation for C as a function of measured values of n/N, RH_{min} and U_{day} has also been developed (Frevert et al, 1983; Burman et al, 1983).

The difficulty in applying the FAO modification of the Blaney-Criddle method in Oregon is not having sufficient measured data to evaluate the C factor at all stations where temperature is measured. This is the case where numerical values for the variables which contribute to the C factor are to be used rather than more conveniently developed estimated values. It is probable that the problems encountered by the authors in this regard will be typical of others wanting to

apply this method. A number of assumptions were required to allow for the greater number of stations with only temperature data to be used in the analysis of regional evapotranspiration. If stations with temperature data were within relatively close proximity to other stations at which n/N, RH_{min} and U_{day} were recorded, and within the same climatic region, values of the parameters required for C were assumed to hold for the station with temperature data only. This assumption will undoubtedly cause an error in the evapotranspiration estimate at the station with only temperature data. Because of the crucial importance of having as many stations as possible for the regionalized variable analysis, it was decided this error would have to be accepted. This point will be further discussed in the Conclusions and Recommendations section. It should be mentioned that the decision of how to distribute data from stations at which n/N, RH_{min} and U_{day} were measured was made after extensive travel throughout Oregon to insure that measurements were applied to stations within similar climatic regions.

Oregon lies on the Pacific coast of the United States and is nearly rectangular in shape extending approximately 550 km (340 mi) in the East-West direction and 440 km (270 mi) in the North-South direction. The main topographical features which affect climate are two predominantly north-south mountain ranges, the Coastal Range and the Cascades, separated by approximately 150 km (90 mi) in the western part of the state, and the high plateau and highlands east of the Cascades. West of the Cascades, the climate is generally humid with annual precipitation in the range of 1000 mm (40 in.) in the inter-mountain valley. East of the Cascades, semi-arid to arid conditions prevail with annual precipitation in the range of 250 mm (10 in.) or less. Topographic features and local meteorological conditions have led to Oregon being divided into the 10 climatic regions indicated in Figure 3. For the purposes of this study, these regions were further grouped into five major climatic regions. These were a) coastal, b) north inter-mountain valley, c) south inter-mountain valley, d) north high plateau, and e) south high plateau. These regions will be referred to throughout the remainder of this paper.



Figure 3. Topographic features and climatic regions for Oregon.

Based on the assumptions indicated for application of meteorological data, there were 175 stations randomly distributed over the state of Oregon for which evapotranspiration estimates could be made. A computer program for semivariogram analysis was designed based on previous work (Journel and Huijbregts, 1978). The state was divided into a square grid 12.87 km (8.00 mi) on a side. A tolerance of 3.22 km (2.00 mi) was accepted so that data from a station within 3.22 km (2.00 mi) of a grid corner was transferred to the corner. If two or more stations were within the tolerance interval for the same grid corner, the evapotranspiration estimate for that corner was taken as the average of that computed from the individual stations.

A semivariogram function was initially computed for the North-South and East-West directions using data from the entire state as a check of the procedure and computer program. Due to the relative uniformity of climatic parameters in a North-South direction and relative lack of uniformity in an East-West direction, it was expected that for each lag the value of semivariogram function in the East-West direction would be greater than that obtained for the North-South direction. This proved to be the case. Although the magnitude of the results are not significant because of violation of the stationarity assumptions, the general shape of the functions for the two directions gave confidence that the computerized semivariogram analysis was functioning as expected. Using data from the grid resulted in between 9 and 22 data pairs per lag being used in equation (6) for lags from 1 to 15 [i.e. distances from 12.86 km (8.00 mi) to 193.05 km (120.00 mi)]. This was assumed to be a sufficient number of data pairs for calculation of a representative semivariogram function within the anticipated range.

Calculation of the drift (first-order moment of evapotranspiration increments) as an indicator of the validity of the assumption of first-order stationarity was performed. This was done using the method suggested by David (1977) and previously applied by Vauclin et al (1982). The formula for the drift is given by:

$$D(k) = \frac{1}{N(k)} \sum_{i=1}^{N(k)} (ET_{r_i} - ET_{r_{i+k}}) \qquad (9)$$

where D(k) is the computed value of the drift for lag k and other parameters are as previously defined. A plot of the drift in the East-West direction for evapotranspiration estimates made for the month of August, 1979, is indicated in Figure 4. Although there is no apparent trend in the drift, there is a certain amount of bias since all of the computed values except one is positive. If there was no bias in the drift, the value of D(k) would be expected to oscillate around zero. This result indicates that the first-order stationarity assumption could not be applied to data analyzed in the East-West direction.

To attempt to work within the constraints of first-order and weak second-order stationarity, i.e. the intrinsic hypothesis, it was initially felt that it would be possible to perform the semivariogram analysis separately for each of the five climatic regions. Proceeding with such an

analysis, it was found that the resulting semivariogram functions had a poorly defined shape and that the maximum number of data pairs for the first 15 lags was 7 with most lags having fewer. This few number of data pairs per lag gave a semivariogram function which could not be relied upon.

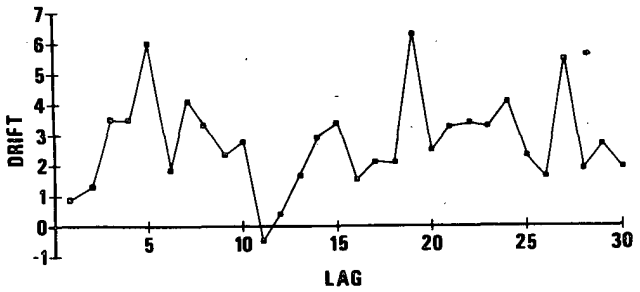


Figure 4. Drift computed for mean monthly evapotranspiration in East-West direction for August, 1979.

Although other alternatives which will be discussed in the Conclusions and Recommendations section could have been chosen, it was decided to normalize the evapotranspiration data from the five separate climatic regions. This normalization procedure would allow all of the data to be grouped together for the semivariogram analysis. The data were normalized by computing the mean reference evapotranspiration for each of the five climatic regions. A dummy variable of 10 mm/day reference evapotranspiration was added to each of the computed values to insure the final results were all positive, and the mean for the climatic region was subtracted from the value at each occupied grid corner. This resulted in a set of evapotranspiration data which was similar in magnitude, due to the normalization process, but still retained the variation of the individual climatic regions. The drift computed for these data using equation (9) showed no systematic bias with D(k) oscillating about zero.

3. RESULTS

The resulting semivariogram function computed using equation (6) on the normalized data is plotted in Figure 5. The dashed line shown in the figure is the sample variance, or lag zero covariance, about which the semivariance is seen to oscillate after approximately lag 10 [equivalent to 129 km (80 mi)]. The solid line shown through the data is a parabolic regression equation chosen to represent the semivariogram function. Any number of types of mathematical representations of the semivariogram function have been applied (Henley, 1981; Gambolati and Volpi, 1979a). Vauclin et al (1982), among others, plot a semivariogram function which has a discontinuous slope at lag k equivalent to the range. Beyond the range, the semivariogram function is modeled as a horizontal straight line. The parabolic regression line in Figure 5 indicates that the data exhibit a range of approximately lag 10 [equivalent to 129 km (80 mi)] and that beyond this distance application of classical statistics rather than regionalized variable analysis is recommended. Within the range, the curve drawn through the data in Figure 5 gives an indication of the spatial structure of the variation of evapotranspiration in an East-West direction for the month of August, 1979. The semivariogram function for the North-South data exhibited a similar characteristic shape and range.

The parabolic regression equation plotted in Figure 5 does not indicate a value of zero semivariance at lag zero. It instead exhibits a positive value of the semivariance which is termed the "nugget effect". This effect corresponds to the variability that occurs within distances shorter than the sampling interval and is typical of hydrologic variables. In fact, the semivariogram function depicted in Figure 5 is similar in shape to those derived from much better, "well-behaved" data, that of soil temperature taken using infrared techniques along two transects over bare soil with a distance of 1.0 m between measurement sites (Vauclin et al, 1982). Although considerable experimentation and refinement is necessary, the shape of the semivariogram function in Figure 5 compared with previously published work indicates that there is a distinct potential for application of semivariogram analysis to evapotranspiration.

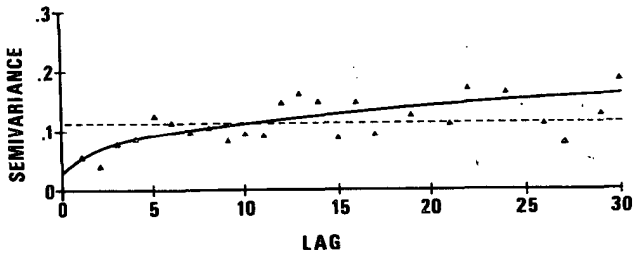


Figure 5. Variogram for mean monthly evapotranspiration in East-West direction for August, 1979.

4. KRIGING METHOD

Theoretical Development

One of the principal aims of geostatistical methods is the estimation of values at given locations from observations made at other locations while accounting for the spatial structure of the variable. The kriging method performs interpolation subject to the following three conditions:

a) The kriged value of the variable at experimental points is the same as the measured value at those points. That is, a surface which passes through the kriged data also passes through the measured data.
b) The mean of the kriged data is equal to the mean of the measured data. Kriging is therefore said to result in a non-bias condition.
c) The variance of the estimated values is minimized.

Kriging techniques are refined versions of weighted moving average techniques. Let Z(x) be

the randon function under study, Z_α represent measured data, and Z_k represent kriged estimates for unoccupied grid corners. The various kriging techniques are based on the simple linear model given by:

$$Z_k = \lambda_1 Z_1 + \lambda_2 Z_2 + \lambda_3 Z_3 + \ldots + \lambda_n Z_n \quad (10)$$

or $$Z_k = \sum_{\alpha=1}^{n} \lambda_\alpha Z_\alpha \quad (11)$$

To obtain a zero error in expectation, termed the non-bias condition, the condition is imposed that,

$$\sum_{\alpha=1}^{n} \lambda_\alpha = 1 \quad (12)$$

so that,

$$E\{Z_k\} = m \sum_{\alpha=1}^{n} \lambda_\alpha = m = E\{Z_\alpha\} \quad (13)$$

where m is the mean or expectation of the random function Z(x).

The difference between Z_k and Z_α at any point p for any set of weights λ_α is the estimation error. As a heuristic example, if the weight on the closest observation at point j is one and all the other weights are zero, then the value of Z_{α_j} would be used as the estimator of Z_{α_p} (Henley, 1981). The estimation error, ε, is given by,

$$\varepsilon = Z_{\alpha_p} - Z_{\alpha_j} \quad (14)$$

If there is no trend, Z_{α_j} is an unbiased estimate of Z_{α_p} and the average error is zero over all possible estimates. The variance of the difference of the estimation error is not zero, however. It is the mean or expectation of the squared difference given by,

$$\delta_\varepsilon^2 = E\left[(Z_{\alpha_p} - Z_{\alpha_j})^2\right] \quad (15)$$

Since Z_{α_p} is the value at location p and Z_{α_j} is the value at location j, this variance is in fact a spatial variance. It is the value of twice the semivariance at distance h between locations p and j. If instead of using just one weighting on one observation j, a number of observations were given different weights, each observation would contribute a different proportion to the total estimation variance of Z_{α_p}. There are in fact an infinite number of ways in which the weighting factors can be allocated, each producing a different estimation variance. The kriging method seeks the one combination of weights which produces the minimum estimation variance (Henley, 1981).

The minimum weights are determined by differentiating δ_ε^2 with respect to the weights λ_α and setting the resulting partial derivatives equal to zero. Full derivation of the system of equations is given in other publications (Clark, 1979; David, 1977; Matheron, 1963). A summary is given here.

For n observations, there will be n unknown weights and n equations. Since the sum of the weights must be equal to one for an unbiased estimate, the additional equation for the sum of the weights equal to 1.0 needs to be included. There are now n + 1 equations. To obtain a well behaved system it is necessary to add one additional unknown, a Lagrangian multiplier. The resulting set of simultaneous linear equations is given below and is called the kriging system.

$$\begin{aligned}
&\lambda_1\gamma(h_{11}) + \lambda_2\gamma(h_{12}) + \lambda_3\gamma(h_{13}) + \ldots + \lambda_n\gamma(h_{1n}) + \mu = \gamma(h_{1p}) \\
&\lambda_1\gamma(h_{21}) + \lambda_2\gamma(h_{22}) + \lambda_3\gamma(h_{23}) + \ldots + \lambda_n\gamma(h_{2n}) + \mu = \gamma(h_{2p}) \\
&\quad\vdots \\
&\lambda_1\gamma(h_{n1}) + \lambda_2\gamma(h_{n2}) + \lambda_3\gamma(h_{n3}) + \ldots + \lambda_n\gamma(h_{nn}) + \mu = \gamma(h_{np}) \\
&\lambda_1 + \lambda_2 + \lambda_3 + \ldots + \lambda_n + 0 = 1.0
\end{aligned} \quad (16)$$

In equation (16), $\gamma(h_{ij})$ is the value of the semivariogram function for distance h_{ij} between observations i and j and μ is the Lagrangian multiplier.

The following assumptions are required for the simple kriging method given by the above series of equations:

a) The data exhibit at least weak second-order stationarity, i.e. the intrinsic hypothesis holds.
b) There is no regional trend or drift in the data.
c) The semivariogram is independent of location.

It should be noted that the kriging estimator is an ordinary weighted average constrained to produce a smooth surface which will not contain values which lie beyond the local and lower bounds of the observed data. The kriging system gives a linear unbiased estimator which results in the minimum estimation variance.

Applications of Kriging

Numerous publications indicate application of the kriging method for analysis of hydrologic variables. Current articles point out that the major advantage of the kriging method may not be as an interpolating method, but may be its

advantage in being able to quantify the interpolation error. Gambolati and Volpi (1979b) point this out and further indicate that any claim to produce more accurate estimates at unoccupied grid corners using kriging compared to traditional interpolation techniques, e.g. least squares polynomial, fails to be proven. This is because a degree of subjectivity is introduced by the decision process of selecting the operative model to represent the semivariogram. It should be pointed out that although choice of the semivariogram function to account for the spatial structure of the variable is subjective, it is not arbitrary as is application of a mathematical function which is not based on any measurements of the variable to be interpolated. The modeler is presented with the choice between an imperfect model, i.e. a model based on incomplete information with respect to the shape of the variogram, and an arbitrary model. Gambolati and Volpi (1979b), with reference to their application of kriging to groundwater aquifers, emphasize that the unique advantage of kriging is its ability to provide an assessment of the interpolation error.

A recent publication which perhaps has the greatest relevance to the current research with respect to application of kriging in data network analysis was done by Villeneuve et al (1979) in the design of streamflow sampling networks. Villeneuve et al (1979) concluded that application of kriging offers the possibility to estimate the variance reduction gained by the addition of data collection stations in region of high variance. Kriging gives an estimation variance for any point in the kriged domain which has significant applications in network design. Continued work in analysis of regional evapotranspiration will attempt to apply these same principles to the evapotranspiration variable.

5. CONCLUSIONS AND RECOMMENDATIONS

Application of regionalized variable analysis and the kriging method has been indicated in studies of a number of meteorological and hydrological variables including precipitation (Delhomme and Delfiner, 1973), streamflow (Villeneuve et al, 1979) and groundwater aquifers (Gambolati and Volpi, 1979a, Volpi et al, 1979). In the work reported in this paper, an attempt has been made to apply geostatistics to regional evapotranspiration, specifically for development of the semivariogram. Success of other researchers in applications of these methods to hydrologic data analysis leads to the conclusion that work in evapotranspiration should also be successful and the variogram indicated in Figure 5 points towards this conclusion. The shape of the semivariogram function indicated in Figure 5 is not unlike that indicated for transects of soil temperature measurements spaced 1.0 meter apart (Vauclin et al, 1982) which could be expected to present a reliable variogram because of the relative simplicity of the variable analyzed.

The major problem in the current evapotranspiration analysis is having a sufficient number of data points to develop representative variograms separately for the various climatic regions of Oregon. This problem led to the normalization and grouping of data from all the regions in development of the variogram indicated in Figure 5.

Two other approaches may be applicable to overcome this problem of paucity of data, which will be experienced whenever evapotranspiration estimates are to be made over large regions. First is application of the technique of cokriging. By considering the spatial correlation that exists between the evapotranspiration variable as estimated by the FAO modified Blaney-Criddle method and another more frequently observed variable such as temperature, cokriging may improve the precision of estimating evapotranspiration (Nielsen et al, 1983). As in general worldwide, there are many more stations with temperature measurements in Oregon than with data required to evaluate the C factor of the FAO modified Blaney-Criddle method. In addition, for the data analyzed in this study, an average of 70 percent of the difference in evapotranspiration was described by differences in temperature. Application of the cokriging technique which uses both the semivariogram function and cross-semivariogram function (Nielsen et al, 1983) may allow for more stations to be brought into the analysis of each climatic region in Oregon and to strengthen the description of spatial variation within the region.

A second possible solution to the problem of having an adequate number of data pairs to confidently perform regionalized variable analysis is application of nonparametric geostatistics. This technique is explained in a recent book by Henley (1981) which also has one of the clearer descriptions of basic geostatistic techniques from which material was taken for this paper. Nonparametric geostatistics was developed for application with limited data or in cases in which the data do not appear to satisfy the criteria required. Classical statistics usually assumes a normal or derivative of the normal distribution for natural phenomenon. Geostatistics makes additional assumptions about the permissible types of spatial continuity relationships, specifically that the mean and variance are independent of spatial location or described by a simple spatial trend. There is, however, a broad class of nonparametric and distribution-free statistical methods which do not require assumptions necessary in classical statistics. In his book, Henley (1981) indicates application of these nonparametric techniques to the analysis of regionalized variables. Possible applications of this method to evapotranspiration analysis will be tested as this research progresses.

The ultimate goal of this research is efficient, timely, computerized plotting of evapotranspiration contours over large regions. Such contours could be drawn for different probabilities based on long historical records of meteorological variables for water resource system and irrigation system design. Contours could also be drawn for data collected in the previous 24 to 48 hours for irrigation scheduling in the western United States where timely information on crop water consumption is crucial to maintaining the highly productive irrigated agricultural sector with maximum efficiency. It is felt that the application of a method which accounts for the spatial structure of evapotranspiration over a

region will prove to be very beneficial in such an analysis.

6. REFERENCES

Burman, R. D., R. H. Cuenca and A. Weiss. 1983. Techniques for Estimating Irrigation Water Requirements, in Advances In Irrigation, Vol. 2, D. Hillel (editor). Academic Press, New York.

Clark, I. 1979. Practical Geostatistics. Applied Science Publishers, London.

Cuenca, R. H., J. M. Erpenbeck and W. O. Pruitt. 1981. Advances In Computation of Regional Evapotranspiration. Proceedings Water Forum '81, Vol. I, American Society of Civil Engineers Specialty Conference, San Francisco, California, August. pp. 73-80.

David, M. 1977. Geostatistical Ore Reserve Estimation. Elsevier, New York. 364 pp.

Delhomme, J. P. and P. Delfiner. 1973. Application du Krigeage a L'Optimisation D'Une Campagne Pluviometrique En Zone Arid. UNESCO-AIHS-OMM, Madrid. 22 pp.

Doorenbos, J. and W. O. Pruitt. 1977. Crop Water Requirements, FAO Irrigation and Drainage Paper 24 Revised, Rome. 144 pp.

Erpenbeck, J. M. 1981. A Methodology To Estimate Crop Water Requirements In Washington State. Thesis presented to Department of Agricultural Engineering, Washington State University, Pullman, Washington. 217 pp.

Frevert, D. K., R. W. Hill, and B. C. Braaten. 1983. Estimation of FAO Evapotranspiration Coefficients. Journal of Irrigation and Drainage Division of American Society of Civil Engineers. (In press.)

Gambolati, G. and G. Volpi. 1979a. Groundwater Contour Mapping In Venice by Stochastic Interpolators, 1. Theory. Water Resources Research, Vol. 15, no. 2, pp. 281-290. April.

Gambolati, G. and G. Volpi. 1979b. A Conceptual Deterministic Analysis of the Kriging Technique in Hydrology. Water Resources Research, Vol. 15, no. 3, pp. 625-629. June.

Hajrasuliha, S., N. Baniabbassi, J. Metthey and D. R. Nielsen. 1980. Spatial Variability of Soil Sampling for Salinity Studies In Southwest Iran. Irrigation Science, Vol. 2, pp. 1-12.

Henley, S. 1981. Nonparametric Geostatistics. Applied Science Publishers, Halsted Press Division, John Wiley and Sons, New York. 145 pp.

Journel, A. G. and Ch. J. Huijbregts. 1978. Mining Geostatistics. Academic Press, New York, 600 pp.

Krige, D. G. 1951. A Statistical Approach To Some Basic Mine Valuation Problems On The Witwatersrand. Journal of Chemical and Metallurgical Mining Society of South Africa, Vol. 52, pp. 119-139.

Matheron, G. 1963. Principles of Geostatistics. Economic Geology, Vol. 58, pp. 1246-1266.

Nielsen, D. R., P. M. Tillotson, and S. R. Vieira. 1983. Analyzing Field-Measured Soil-Water Properties. Agricultural Water Management, Elsevier. (In press.)

Vauclin, M., S. R. Vieira, R. Bernard and J. L. Hatfield. 1982. Spatial Variability of Surface Temperature Along Two Transects of a Bare Soil. Water Resources Research, Vol. 18, no. 6, pp. 1677-1686. Dec.

Villeneuve, J. P., G. Morin, B. Bobee, D. Leblanc, and J. P. Delhomme. 1979. Kriging in the Design of Streamflow Sampling Networks. Water Resources Research, Vol. 15, no. 6, pp. 1833-1840. Dec.

Volpi, G., G. Gambolati, L. Garbognin, P. Gatto, and G. Mozzi. 1979. Groundwater Contour Mapping In Venice by Stochastic Interpolators 2. Results. Water Resources Research, Vol. 5 no. 2, pp. 291-297. April.

ACKNOWLEDGMENTS

This research has been supported by the Agricultural Experiment Station of Oregon State University and the African-American Institute. This paper has been designated as Agricultural Experiment Station Technical Paper No. 6856.

ASPECTS OF OCEANIC WHITECAP COVERAGE DEPENDENCE ON WIND SPEED: HETEROSCEDASTICITY IN THE DATA, AND THE ESTIMATION OF THE BEAUFORT VELOCITY*

Iognaid G. O Muircheartaigh[1] and Edward C. Monahan[2]

[1]Department of Statistics,
University College,
Galway, Ireland.

[2]Department of Oceanography
University College,
Galway, Ireland.

1. INTRODUCTION

Several diverse considerations have recently focused attention on the need for a precise expression describing the dependence of oceanic whitecap coverage, W, on the 10_m-elevation wind speed, U. Whitecap coverage, because of the effect of each whitecap on the apparent microwave brightness temperature of the sea surface (Webster, et al, 1976; Wentz, 1983) is amenable to remote measurement with such satellite instruments as the Scanning Multichannel Microwave Radiometer (Gloersen and Barath, 1977). If the relationship between wind speed and whitecap coverage, the measured quantity, is known, then the output of the satellite SMMR and similar instruments can be made to yield up the near-sea-surface wind speed (Monahan and O'Muircheartaigh, 1981). Attention has recently been focused on the role of whitecaps in the production of marine aerosols (e.g., Cipriano and Blanchard, 1981; Monahan, et al, 1982) and in the sea-air exchange of gases (e.g. Kerman, 1983; Monahan and Spillane, 1983). In developing formal models and budgets for such properties of the marine atmospheric boundary layer as its aerosol population, it is necessary to have a detailed expression relating whitecap coverage, and hence aerosol production, to the near-surface wind speed (Monahan, et al, 1983).

The expressions for the U-dependence of W published to date fall almost without exception into the category of simple power-laws (Monahan, 1971; Wu, 1979; Monahan and O'Muircheartaigh, 1980). But it has long been recognized that there is a wind speed below which whitecapping does not occur, a wind speed which we refer to as the Beaufort velocity, U_B. Munk (1947 suggested that the wind speed associated with the onset of whitecapping was about 6 ms^{-1}, but more recent observers (Blanchard, 1963; Monahan, 1971) have concluded that U_B is more like 3 or 4 ms^{-1}. Nevertheless, workers in the field of remote sensing (e.g. Gloersen and Barath (1977) still favour relatively high, i.e. 7 ms^{-1}, values for U_B. With the correct value of U_B still in doubt, and hence the range of applicability of various modeling equations and the range of utility of several remote sensing systems uncertain, the need for an accurate determination of U_B from the various W, U-data sets (Monahan, 1971; Toba and Chaen, 1973; Monahan, et al, 1981) is manifest.

It is clear that a power-law expression for the U-dependence of W is inadequate from the point of view of estimating U_B, as such an expression implies $U_B = 0$. In this paper we describe a more general model which we fitted to our data sets; several other models were also fitted to the data, but the best model, in the sense of minimum mean squared error (cf. Monahan and O Muircheartaigh, 1980) turned out to be a simple modification of the power-law.

2. STATISTICAL METHODOLOGY

In a previous paper (Monahan and O Muircheartaigh, 1980), we have fitted a power law expression

$$W = \alpha U^{\lambda} \qquad (1)$$

to the data sets described in the introduction. The parameters of equation (1) were estimated by (a) a combination of ordinary least squares (OLS) analysis and a functional minimisation technique and (b) robust biweight fitting (Mosteller and Tukey, 1977), a method of estimation which consists of an iterative application of OLS, with successive calculation of weights which reduce the influence of outliers in the original data set, thereby yielding an estimate which better

*This work was supported by ONR Grant N00014-82-G-0024.

characterises the main body of the data set. This latter method is appropriate in situations where the data suggest departures from certain basic assumptions.

In this paper, we describe the fitting of the modified power-law expression

$$W = aU^b - c \qquad (2)$$

to the various data sets, described in detail below. This model allows for a non-zero value of U_B. In fact, if $\hat{a},\hat{b},\hat{c}$, are our estimates of a,b,c respectively, then U_B is estimated by

$$\hat{U}_B = (\frac{c}{a})^{1/b}. \qquad (3)$$

The estimation procedure for the parameters of equation (3) was two-fold. Firstly, the parameters of the model (a,b,c,) were estimated by non-linear least squares. Essentially this involved the minimisation, with respect to (a,b,c,) of the quantity

$$R(a,b,c) = \Sigma_i (W_i - aU_i^b - c)^2 \qquad (4)$$

where W_i denotes the i^{th} observation on W, and U_i the i^{th} observation on U. This estimation involved the assumption of homoscedasticity in the data, i.e. essentially that all the observations were equally reliable. An analysis of the residuals

$$r_i = (W_i - \hat{a}U_i^{\hat{b}} - \hat{c}) \qquad (5)$$

(where $\hat{a}$ $\hat{b}$ $\hat{c}$ are the estimates of a,b,c, respectively) revealed that this certainly was not the case in this instance, and that the variance of the value of W about its predicted value ($aU^b - c$) was an increasing function of U. Hence it was clear that heteroscedasticity was present in the data. For a full description of the treatment of heteroscedasticity in regression analysis see (Kmenta (1971). Essentially the problem is to determine the form of the heteroscedasticity. When no assumptions are made about the nature of the heteroscedasticity, we have to rely entirely on the sample information and estimate the variances of the residuals from the data. Since for each specific value of U the value of the variance of the residuals may be different, we need several observations on the dependent variable W for each U_i, in order to estimate this variance. Since our data consisted of single observations at different values of U, it was not possible to determine the form of the dependence of the variance of the residuals on U from the data. However, an analysis of the dependence of the absolute values of these residuals on the independent variable U revealed that this (absolute value) was strongly dependent on U, and that the form of this dependence was given by

$$E\{|r_i|\} = \alpha + \beta U_i^2 \qquad (6)$$

Clearly, therefore, the expected value of the absolute value of the residuals was proportional to U^2, and therefore since $E\{|r|\}$ is proportional to the standard deviation of the residuals, the variance of the residuals would be proportional to U^4.

The general principles of weighted least squares therefore suggest that each of the observations be weighted by a factor of $\frac{1}{U^4}$ (the inverse of the variance of the residuals) in the calculation of the least squares estimates. The effect of this weighting was to attach relatively high importance to points with low wind speed, (and small variability) and less importance to those with high wind speed (and large variability) in the estimation of the parameters a,b, and c of equation (2), and in the consequential estimation of U_B.

3. THE ANALYSIS OF THE DATA

The analysis described above was applied to

(a) a set of (W,U) observations obtained during the 1978 Joint Air-Sea Interaction experiment (JASIN) conducted in the vicinity of Rockall in the North Eastern Atlantic (Monahan et al, 1981).
(b) the data set described in (a) combined with those of Monahan (1971) and Toba and Chaen (1973).

The analysis of the data with respect to (a) is described in detail, and that with respect to (b) in summary form.

3.1. Initial analysis of JASIN data

In fitting the model described in equation (2) to the data, it was deemed appropriate to scale the wind speed by dividing the U values by 10. Hence the model actually fitted to the 55 data points in the JASIN set was

$$W = a(\frac{U}{10})^b - c \qquad (7)$$

Table 1 gives the initial estimates of the parameters and their standard errors.

Table 1. Initial Estimates of parameters of modified power law for JASIN data.

Parameter	Estimate	Standard Error
$\hat{a}$	.0237	.006
$\hat{b}$	2.197	.607
$\hat{c}$	.00358	.00384

An analysis of the residuals for this model revealed, as indicated previously, that there was a strong positive dependence on U. Exploratory analysis with different forms indicated that a very good description

of this dependence was given by equation (6), or rather the scaled version

$$E\{|r|\} = \alpha + \beta(\frac{U}{10})^2 \qquad (8)$$

Table 2 gives the parameter estimates of this model based on the JASIN data.

Table 2. Parameter estimates and standard errors for equation (8) for JASIN data.

Parameter	Estimate	Standard Error
$\hat{\alpha}$	- .00149	- .00121
$\hat{\beta}$	.0142	.00177

It is clear from Tabel 2 that $E\{|r|\}$ is essentially proportional to U^2. Hence, as already described, we then performed a weighted least squares estimation of the parameters of equation (2), with weights proportional to $\frac{1}{U^4}$. The results of this estimation as applied to the JASIN data set are presented in Table 3.

Table 3. Weighted estimates of parameters and standard errors for JASIN data

Parameter	Estimate	Standard Error
$\hat{a}$	0.0207	0.00419
$\hat{b}$	2.550	0.526
$\hat{c}$	0.00177	0.00217

The final estimated model for the JASIN data set is therefore given by

$$W = 0.0207(\frac{U}{10})^{2.55} - 0.00177 \qquad (9)$$

The corresponding estimate $\hat{U}_B$ of U_B is obtained by putting $W = 0$ in equation (9) and solving for U, i.e.

$$\hat{U}_B = 10(\frac{\hat{c}}{\hat{a}})^{1/\hat{b}} = 3.80 \text{ ms}^{-1}.$$

It should be noted that the corresponding estimate of U_B derived from the unweighted least squares estimation presented in Table 1 was 4.23 ms^{-1}.

3.2. The combined data sets

In this section we present the results of an analysis similar to that described in Section 3.1 when applied to a data set consisting of a combination of the JASIN set, and the data from Monahan (1971) and Toba and Chaen (1973) involving a total of 145 W,U observations. The initial analysis again revealed that the dependence of $|r|$ on U was well described by a model of the form given in equation (8). The subsequent weighted least squares analysis yielded the results given in Table 4.

Table 4. Weighted estimates of parameters and standard errors for combined data sets.

Parameter	Estimate	Standard Error
$\hat{a}$	0.00657	0.00249
$\hat{b}$	2.176	0.8370
$\hat{c}$	0.00088	0.00167

The corresponding estimate $\hat{U}_B$ of U_B was $\hat{U}_B = 3.98 \text{ ms}^{-1}$.

4. SUMMARY

The paper describes the fitting of a modified power law expression to a number of different whitecap cover, wind speed data sets. Evidence is presented of heteroscedasticity in the data, and the form of this heteroscedasticity is determined. The consequential modifications in the least squares analysis of the original model are carried out via a weighted least squares analysis, and the corresponding estimates of the Beaufort velocity U_B are obtained. These estimates fall in the range 3.8 - 4.0 ms^{-1}.

5. REFERENCES

Blanchard, D. C., 1963: The electrification of the atmosphere by particles from bubbles in the sea. Progress in Oceanography, 1, 71-202.

Cipriano, R. J., and D. C. Blanchard, 1981: Bubble and aerosol spectra produced by a laboratory 'breaking wave'. Journal of Geophysical Research, 86, 8085-8092.

Gloersen, P. and F. T. Barath, 1977: A scanning multichannel microwave radiometer for Nimbus-G and Seasat-A, I.E.E.E. Journal of Oceanic Engineering, OE-2, 172-178.

Kerman, B. R., 1983: Distribution of bubbles near the ocean surface, Proceedings, International Symposium on Gas Transfer at Water Surfaces, Cornell U. (in press)

Kmenta, J., 1971: Elements of Econometrics Macmillan.

Monahan, E. C., 1971: Oceanic Whitecaps, Journal of Physical Oceanography, 1, 139-144.

Monahan, E. C., and I. G. O'Muircheartaigh, 1980: Optimal Power-Law Description of Oceanic Whitecap Coverage Dependence on Wind Speed. Journal of Physical Oceanography, 10, 2094-2099.

Monahan, E. C. and I. G. O'Muircheartaigh, 1981: Improved Statement of the relationship between surface wind speed and oceanic whitecap coverage as required for the interpretation of satellite data, 751-755, in Oceanography from Space, J.F.R. Gower, Ed., Plenum Pub.

Monahan, E. C., and M. C. Spillane, 1983: The role of oceanic whitecaps in air-sea gas exchange, Proceedings, International Symposium on Gas Transfer at Water Surfaces, Cornell U., (in press).

Monahan, E. C., K. L. Davidson, and D. E. Spiel, 1982: Whitecap Aerosol Productivity Deduced from Simulation Tank Measurements, Journal of Geophysical Research, 87, 8898-8904.

Monahan, E. C., I. G. O'Muircheartaigh and M. P. FitzGerald, 1981: Determination of surface wind speed from remotely measured whitecap coverage, a feasibility assessment, Proceedings of an EARSel-ESA Symposium, Application of Remote Sensing Data on the Continental Shelf, Voss, Norway, 19-20 May 1981, European Space Agency SP-167, 103-109.

Monahan, E. C., D. E. Spiel, and K. L. Davidson, 1983: Model of marine aerosol generation via whitecaps and wave disruption, Preprint Volume, Ninth Conference on Aerospace and Aeronautical Meteorology, American Meteorological Society, (in press)

Mosteller, F. and J. W. Tukey, 1977: Data analysis and regression, Addison Wesley.

Munk, W. H., 1947: A critical wind speed for air-sea boundary processes, Journal of Marine Research, 6, 203-218.

Toba, Y., and M. Chaen, 1973: Quantitative expression of the breaking of wind waves on the sea surface, Records of Oceanographic Works in Japan, 12, 1-11.

Webster, W. J., T. T. Wilheit, D. B. Ross, and P. Gloersen, 1976: Spectral characteristics of the microwave emission from a wind-driven foan-covered sea, Journal of Geophysical Research, 81, 3095-3099.

Wentz, F. J., 1983: A model function for ocean microwave brightness temperatures, Journal of Geophysical Research, 88, 1892-1908.

Wu, J., 1979: Oceanic whitecaps and sea state, Journal of Physical Oceanography 9, 1064-1068.

A GENERAL STOCHASTIC MODEL OF THE PRECIPITATION PROCESS WITH APPLICATIONS ON INTEGRATED PRECIPITATION, MAXIMUM DAILY PRECIPITATION, ETC.

Hans Alexandersson
Department of Meteorology
University of Uppsala
Uppsala, Sweden

1. INTRODUCTION

In statistical climatology precipitation data have often been analysed using empirical relations. Using the concepts and theories developed for stochastic processes a more general and, to some extent, physical approach can be adopted. Theoretically there is an almost infinite number of ways to formulate more or less complicated stochastic models of the precipitation process. Here something like a middle course has been chosen as the following requirements should be fulfilled

1) The stochastic process should have some physical significance.
2) It should be possible to reduce it to an analytical expression for distributions of integrated amounts.
3) It should be possible to estimate the parameters just using series of monthly integrated amounts.
4) It should be as general as possible so it can be applied to other interesting quantities as maximum daily precipitation, and so it can be modified to cope with more complex problems.

The second requirement considerably reduces the number of possible models to be chosen between. As argued below the compound Poisson-exponential (Cpe) model seems to be a good compromise between simplicity and complexity. The two parameters of this model can then be physically interpreted as a number parameter (mean number of independent precipitation events within e.g. a month) and an amount parameter (mean amount of precipitation at each event during e.g. a month). It should be mentioned that the application of a compound Poisson distribution on precipitation data was probably first proposed by W. Feller (1968).

2. FUNDAMENTAL ASSUMPTIONS AND RELATIONSHIPS

The two fundamental assumptions of the Cpe model can be formulated as follows

1) A rain event is assumed to start within an infinitesimal time interval with a probability proportional to the length of this interval. Denote the proportionality constant by ν.
2) The rain amount at each event is assumed to have an exponential distribution. Denote the argument of this distribution by λ which means that the mean value at each event is $1/\lambda$ mm.

Let us discuss these two assumptions in some detail. The first one is also called the Poisson postulate and can also be interpreted as an assumption that the rain events of this model are independent. This can certainly be criticized but one must be aware that the "events" of this model usually comprise several individual "incidents" of rain as shown by the pluviograph below. The bars included are the reduced form of the rain intensity process and which should be described by the Cpe model. Note that the three distinguistable incidents of rain are subjectively comprised to two events.

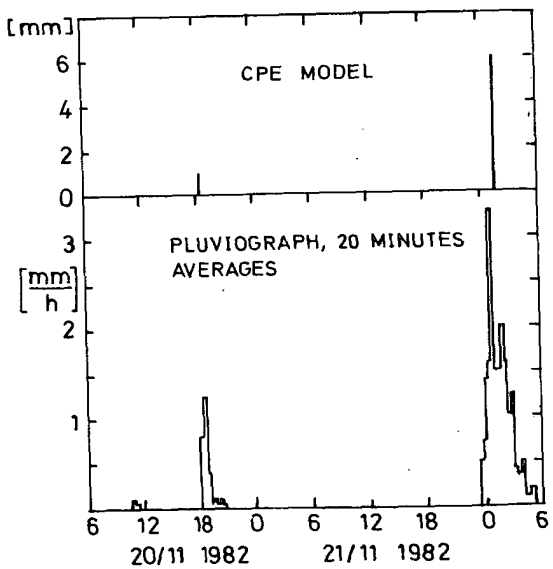


Fig. 1. Pluviograph record with rain intensities as 20 minutes averages from Uppsala, Sweden (59°52'N, 17°38'E) and a reduced set of bars with amounts at each "event".

To assume that the events are independent cannot be strictly true unless they occur at really long time intervals due to some tendency of a long memory of the atmosphere-ocean system in middle latitudes. The most striking and discussed long memory is related to blockings and long sequences of dry days. It is, however, found that these sequences are indeed not far from geometrically distributed, which means that from a stochastic point of view they are almost and merely a random clustering of dry days (Neumann, 1955; Gabriel and Neumann, 1957). A short memory will also be reflected in almost zero correlation between nearby months. This is also generally found for middle latitudes except within some very maritime

western slopes, see e.g. Corona (1979) and Alexandersson (1983). The latter report will henceforth be abbreviated A (1983). Also the autocorrelation of series of daily precipitation should drop off rapidly. This is shown in Section 6 and in Fig. 9 where a modified form of the Cpe model is discussed.

For a finite time interval the Poisson postulate consequently gives a Poisson distribution for the number of rain events with argument and mean value equal to $m=\nu t_0$. This can be shown quite easy (Feller, 1968). Summing up this passage it is felt the first assumption above is quite good to be as simple as it really is.

The second postulate can to some extent be justified empirically but also with some theoretical arguments. Suppose first that a rain event with an average rain intensity r (mm/h) ends within a time interval τ with a probability proportional to the length of this interval. Denote the proportionality constant by μ. Then from elementary statistical theory the integrated amount at this event will have an exponential distribution with mean value r/μ. Then, what is important is only the ratio r/μ so an event with large intensity r, like a shower, will also have a large μ and the ratio r/μ may well be about the same for heavy rain of short duration as for long-lasting moderate frontal rain. We can also analyse an observed distribution of daily amounts which, although not identical to amounts of independent events as interpreted by the Cpe model, clearly should resemble an exponential distribution. This is made in Fig. 2 for Särna, Sweden (61°41'N, 13°8'E) and shows a rather typical exponential decay. Days with less than about 1 mm, which occur more frequently than expected by the exponential curve, could certainly often be naturally included in precipitation events of a nearby day, especially since the consecutive 24 hour periods are an artificial division of the continuum of time.

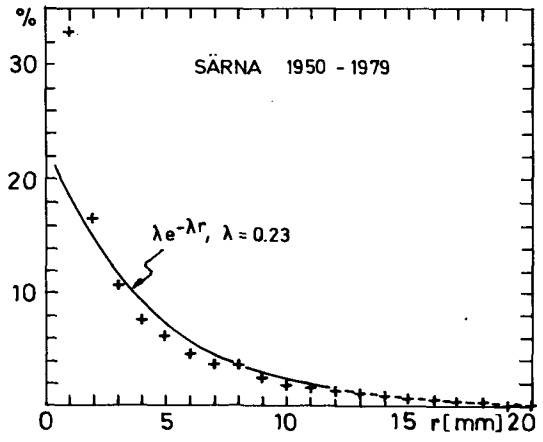


Fig. 2. Observed distribution for the daily amount of precipitation at Särna in central Sweden. Also included is an exponential curve.

As noted in the second postulate we simplify the notations by using $1/\lambda$ for the mean amount at each event instead of r/μ. Thus the Cpe model contains two parameters only, the constants ν and λ. For the integrated amount during a fixed time interval t_0 we use m (νt_0) and λ as parameters. For this integrated amount, denoted by Y, the stochastic process is reduced to a Cpe distribution for which we can write

$$Y = X_1 + \ldots + X_N \tag{1}$$

where, for N and X respectively, we have

$$p(n) = m^n e^{-m}/n! \quad , \quad n \in \{0,1,2\ldots\} \tag{2}$$

and

$$f(x) = \lambda e^{-\lambda x} \quad , \quad x \geq 0 \tag{3}$$

This is quite a unique two parameter distribution to use in applications since there is a finite amount of probability of obtaining no precipitation at all during the time interval considered. Other widely used distributions, as the gamma and the Weibull distributions, must be completed and complicated by a third parameter to handle series of say monthly precipitation as these may contain rare months of no precipitation at all even in northern Europe.

The frequency function for the Cpe distribution is unfortunately not so simple but can at least be concentrated to the form

$$\begin{cases} f(y) = e^{-(\lambda y+m)} \sum_{n=1}^{\infty} (\lambda m)^n y^{n-1}/(n!(n-1)!), & y>0 \\ p(y) = e^{-m} & , y=0 \end{cases} \tag{4}$$

This a bit awkward sum must be handled by computer which, however, is not a great drawback in these days. More important is that the model can be reduced to an analytical expression for integrated amounts. Then we are able to use monthly series of precipitation to estimate the parameters of the model. Especially we may use the maximum likelihood method.

For n larger than 12 the product n! is approximated by the Stirling formula written as

$$n! \simeq \sqrt{2\pi}\, n^{n+\frac{1}{2}} e^{-n} \cdot (1+1/(12n)+1/(288n^2)) \tag{5}$$

3. MOMENTS AND ESTIMATION PROCEDURES

The Cpe distribution is more complex than several other types of distributions used in climatology. It turns out to be a short cut to make use of the moment-generating function to derive the moments of this compound distribution. Denoting this function by $\psi_Y(t)$ the general expression for a compound Poisson distribution is given by (Feller, 1968)

$$\psi_Y(t) = e^{m(\psi_X(t)-1)} \tag{6}$$

Here X is assumed exponential so

$$\psi_X(t) = \lambda/(\lambda-t) \tag{7}$$

and then

$$\psi_Y(t) = e^{mt/(\lambda-t)} \tag{8}$$

From this expression all central moments can be derived from successive differentiations at $t = 0$. For mean value, variance and skewness we get respectively

$$\begin{cases} E(Y) = m/\lambda \\ V(Y) = 2m/\lambda^2 \\ Sk(Y) = 3/\sqrt{2m} \end{cases} \tag{9}$$

Thus the mean value of the integrated amount is obtained as the product of mean number and mean amount. We also see that the skewness is always positive and tends to zero as m tends to infinity.

The coefficient of variation is, like the skewness, independent of λ and given by $\sqrt{2/m}$.

Using the first two relations in (9) we can estimate the parameters m and λ from the sample mean $\overline{y}$ and the sample variance s^2. This technique, the method of moments, gives

$$\begin{cases} m = 2\,\overline{y}^2/s^2 \\ \lambda = 2\,\overline{y}/s^2 \end{cases} \qquad (10)$$

Of course we can as well use mean value and coefficient of variation, two statistical quantities quite often found on plotted maps. Consequently the method of moments points out a fast way to establish e.g. monthly distributions. However, as the Cpe distribution is significantly skewed for precipitation amounts for months and shorter periods it can be expected that the more satisfactory maximum likelihood method gives clearly better estimates.

To find the maximum likelihood solution is equal to find the maximum of the product of probabilities for the observed series, i.e.

$$L = \prod_{i=1}^{k} f(y_i) \qquad (11)$$

Only a partial solution can be obtained when maximizing this product and that is (A, 1983)

$$m \cdot 1/\lambda = \overline{y} \qquad (12)$$

It turns out be easier to obtain the complete solution by using Eq (12) in Eq (11) to get an equation with one independent variable. It is perhaps worthwhile to show how this has been done more precisely. First three values of m is chosen as the solution of Eq (10) and two adjacent values. Then Eq (12) is used as this relation is valid for the maximum likelihood solution. Then the product L, or rather the logarithm of L, is calculated for each set (m, λ). A second order polynomial is then fitted to the points in the (m, ln L)-plane and then maximized and a new pair $(m_x, \ln L_x)$ is obtained. Then the pair with lowest ln L-value in the original set is replaced by the new one and so on. This technique is illustrated by Fig. 3 below

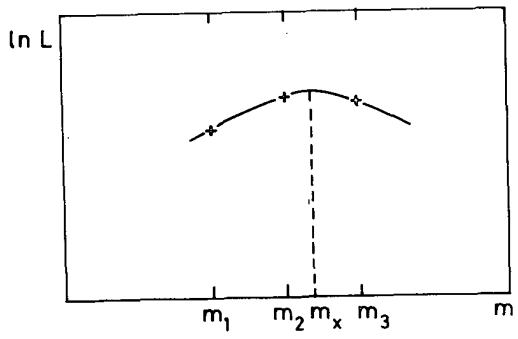


Fig. 3. A sketch of the iteration technique used to find the maximum likelihood estimates of the Cpe distribution.

The differences between the estimates for the two methods are generally within 10%, the largest deviations occurring for the driest months and months with very low, sometimes zero, values in the observed series. In subsequent examples the maximum likelihood estimates are always used.

4. APPLICATIONS TO MONTHLY PRECIPITATION SERIES AT TWO STATIONS

The annual variation of the number and the amount parameter on a monthly basis will in principle be sufficient to describe the precipitation climate at a single point as far as monthly precipitation is concerned. Here two stations are used representing quite different climates of the southern peninsula of Sweden (Götaland and parts of Svealand). One is Uppsala (59°52'N, 17°38'E) at the northeastern part of the peninsula, the other is Gångarebo (57°01'N, 12°55'E) in the more exposed southwestern area. Uppsala is situated on flat plains surrounding the lake Mälaren, an area which receives 500-550 mm/year on average. Gångarebo on the other hand is situated on the western slopes of the hilly inland area of Götaland and receives about 1100 mm/year on average. The gradients within this latter region are large from about 700 mm at the coast to 1100-1150 mm at altitudes of 150-200 m and 15 to 20 km from the coast. According to studies made at other windward slopes the increase of 415 mm/175 m (or rather 60%/175 m) is clearly larger than at other but steeper slopes where the orographic control is more clear-cut (Chuan & Lockwood, 1974; Deljanic, 1972). This indicates that the friction discontinuity (sea/land) together with e.g. sea breeze fronts may be about as important as the ordinary orographic uplift for these gentle slopes and at this distance from the coast.

In Fig. 4a and 4b the Cpe parameters are plotted as functions of month. Note that for Uppsala 147 years of reliable measurements were available while for Gångarebo a series of 69 years can be obtained.

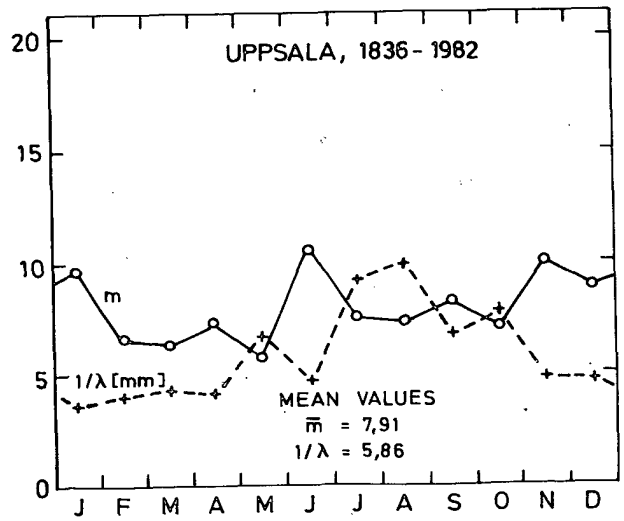


Fig. 4a. Annual variation of Cpe parameters, Uppsala 1836-1982.

From these plots we find that the number of events are almost the same and the difference in precipitation amounts is reflected mainly in the mean amount at each event. Even 147 years of measurements give jerks in these curves that partly could be explained by an insufficiently large sample.

We can also note that the total annual number of independent rain events are 95 and 89 for Uppsala and Gångarebo respectively. This is in a way

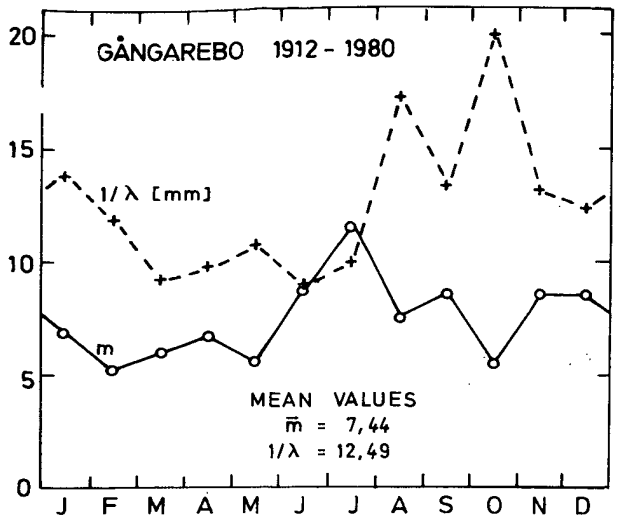


Fig. 4b. Annual variation of Cpe parameters, Gångarebo 1912-1980.

an estimate of the degrees of freedom for the precipitation climate, i.e. on the average the memory of the atmosphere concerning precipitation falls off to near zero after $365/90 \simeq 4$ days. This is in reasonable accordance with more direct studies of this (Neumann, 1955) and also with the autocorrelation of daily precipitation discussed in Section 7. Furthermore the number of rain events, 95 and 89, can be compared with the number of days with precipitation at the two stations which are for ≥ 0.1 mm 163 and 201 and for ≥ 1.0 mm 97 and 142 respectively (1962-79). Thus several of the just slightly wet days do not constitute independent rain events in the sense of the Cpe model.

To use the Cpe distribution in making one-point or spatial analyses of percentiles (e.g. the value exceeded by 5% of the population), probabilities, return periods etc. the parameters must of course be inserted in the expression for the frequency distribution. Once this parent distribution is established one may also get distributions for extreme years, (p. 404 ff, Lindgren, 1968), e.g. for the largest precipitation during an August month within 20 years. One example is shown in Fig. 5. August has been the wettest month in large parts of Sweden during the 20:th century.

From this plot one may read any probabilities or percentiles and classify into very dry, dry, normal, wet and very wet according to a chosen scheme of probability into each class.

The most crucial parts of the distributions are the ends representing unusually dry or wet conditions. Here a bias of the theoretical model chosen is very critical. In the table below theoretical and observed values of cumulative distributions are compared for F(10), i.e. the probability of obtaining less than 10 mm. At Uppsala this probability reaches a maximum of about 10% in February and March but is as low as about 1% in the summer months. Especially for Uppsala with 147 observations the agreement should be good between observations and the Cpe distribution of these more or less rare events.

Table 1. Comparison between theoretical, Cpe (F(10)) and observed (F*(10)) probabilities of obtaining less than 10.0 mm during a month.

	Uppsala, 1836-1982		Gångarebo, 1912-1980	
Month	F(10)	F*(10)	F(10)	F*(10)
Jan	0.030	0.020	0.011	0.000
Feb	0.106	0.109	0.047	0.072
Mar	0.113	0.116	0.039	0.036
Apr	0.076	0.082	0.023	0.014
May	0.072	0.061	0.041	0.065
Jun	0.009	0.007	0.006	0.000
Jul	0.014	0.027	0.001	0.000
Aug	0.015	0.014	0.005	0.000
Sep	0.017	0.000	0.003	0.000
Okt	0.026	0.034	0.021	0.029
Nov	0.012	0.007	0.004	0.000
Dec	0.025	0.007	0.004	0.000
Mean value	0.043	0.040	0.017	0.018

No alarming deviations and no systematic bias seem to be present here which is also supported by other investigations of the extreme tails of the Cpe distribution.

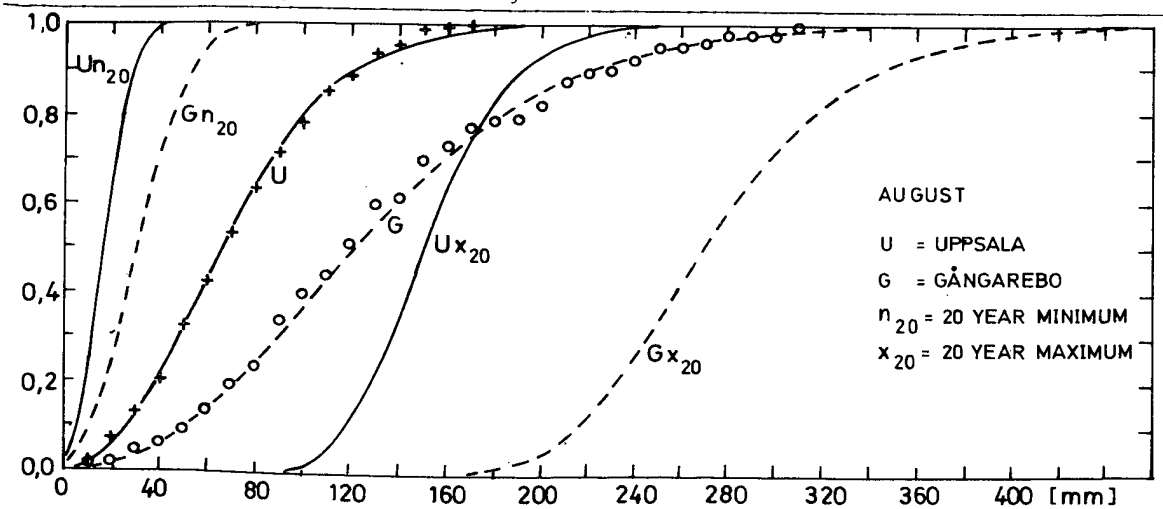


Fig. 5. Cumulative Cpe distributions and extreme value distributions at Uppsala and Gångarebo.

5. COMPARISONS WITH OTHER DISTRIBUTIONS

Frequently the χ^2-distribution is used to compare the goodness of fit of distributions, but here we have instead used the likelihood function to compare three distributions, the Cpe, the Γ and the Normal distribution, although the last one is defined below zero. All three are two parameter distributions, thus it is fair to make a direct comparison of the value of the likelihood function when the maximum likelihood estimates are derived. Some Swedish stations were used and as a few months with zero precipitation existed the ordinary Γ distribution sometimes failed. In such cases the incomplete Γ distribution, including a third parameter which represents the probability at zero, must be used. The maximum likelihood estimates of the Γ distribution have been derived by Thom (1958) and for the Normal distribution they are simply equal to the estimates obtained from the method of moments.

Only in 5 out of 60 cases was the Normal distribution best. Thus we hereafter only compare the test results of the Cpe and the Γ distribution. Including the three months with zero values the Cpe distribution was better in 41 out of 60 cases. This is highly significant according to a sign test. Excluding the three months with zero values the Cpe distribution is better in 38 out of 57 cases which still is a significant difference. Thus even the incomplete Γ distribution cannot compete with the Cpe distribution although the latter one only has two parameters. Comparing distributions with different numbers of parameters should include a loss function and is far from straightforward. A more complete description of this comparison can be found in A (1983).

In an extensive study of the Swedish precipitation climate it was found that the Γ distribution was better than e.g. the Lognormal and the Weibull distribution (Eriksson, 1980). Thus it is felt that the Cpe distribution with its natural parameters is superior in describing monthly and annual precipitation probabilities, at least as far as no more than two parameters are used.

6. THE MAXIMUM CPE DISTRIBUTION

Let us make another application of the general Cpe model. Instead of considering the sum in Eq. (1) we may be interested in the largest of the individual rains which we can denote by Z.

$$Z = \text{Max}\ \{X_1, .., X_N\} \tag{13}$$

Here, as before, X is assumed exponential with mean value $1/\lambda$ and N is assumed of Poisson type with mean value m. To find the frequency distribution of Z one has to use the concept of conditional probability and the theory for maximum values. The result is a type of a double exponential distribution and more precisely

$$\begin{cases} f(z) = \lambda m\, e^{-\lambda z}\, e^{-m\, e^{-\lambda z}} & , \ z > 0 \\ p(z) = e^{-m} & , \ z = 0 \end{cases} \tag{14}$$

and for the cumulative distribution

$$F(z) = e^{-m\, e^{-\lambda z}} \quad , \ z \geq 0 \tag{15}$$

Henceforth we use Max Cpe for this distribution in the text. Mean value and variance are approximately given by

$$\begin{cases} E(Z) \simeq (\ln m + \gamma)/\lambda, \ \gamma = 0.5772 \ \text{(Eulers constant)} \\ V(Z) \simeq \pi^2/(6\lambda^2) \end{cases} \tag{16}$$

The derivations of these formulas as well as the maximum likelihood estimates of the Max Cpe distribution can be found in A (1983).

One way of obtaining m and λ, e.g. monthly, would be to use the values determined by application of the Cpe distribution to the series of integrated amounts. However, in climatology one is more interested in 24 hour maxima than in maxima of amounts at certain events as defined by the Cpe model. These quantities are generally not identical. As a rule an independent event comprises more than 24 hours of precipitation. This can e.g. be seen from the values of m, the number of events, which is clearly smaller than days with precipitation (≥ 0.1 mm), see Section 4. This is also to be expected from the fact that a rain may fall at both sides of the shift of the day, or rather at both sides of some special reading time which is 6 GMT in Sweden.

To adapt the Max Cpe distribution to describe 24 hour maxima we use observed series of these maxima and the maximum likelihood technique to find m and λ. This technique is easier to apply on the Max Cpe than on the Cpe distribution.

We can make a little special trick to get some more information out of Max Cpe distributions. Then we transform Z according to

$$Z' = \begin{cases} Z-c & , \ z > c \\ 0 & , \ z \leq c \end{cases} \tag{17}$$

where c is a lower "heavy rainfall" limit, i.e. if this value is not exceeded during a specific month or year we consider that month or year to be without any truly heavy rainfall. In Sweden 10 mm could be an appropriate value to choose as a lower limit. One effect of this transformation is that the number m of the Max Cpe model now will be the number of really heavy rainfalls above c mm in 24 hours. The variable Z' may also be better than Z to describe very rare heavy rainfalls using a relatively high value on c. Then very large amounts (outliers) will be more important for the distribution. This technique is possible to use if we have very long series and may be better than using the variable Z if the very heaviest rains fall somewhat aside of the Max Cpe curve.

In Fig. 6 two months, May and August, are analysed and the corresponding observed cumulative distributions are also marked. Also included is the number of days with more than 10 mm (c), a Poisson variable according the Cpe model.

In Fig. 7 the 24 hour maximum of the whole year is plotted together with 20 year minima and maxima (cf the previous chapter) for Uppsala.

Comparing August with the whole year we find that they give almost the same probabilities above 50 mm while at 10 mm, the heavy rainfall limit chosen here, the cumulative distribution for August is 0.16 but 0.00 for the year. This is not surprising since the individual month August contributes with many of the largest values (11 out of 19 cases with > 40.0 mm during 1866-1982) without contributing with its lowest values.

It can be added that the 24 hour maxima of Uppsala are indeed low for being in southern Swe-

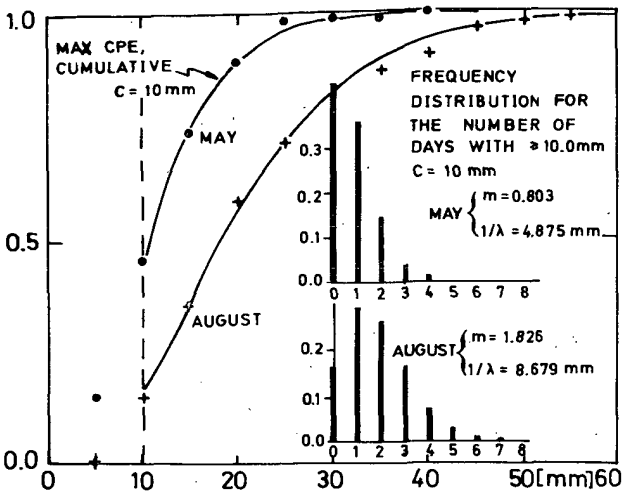


Fig. 6. Theoretical (Max Cpe, c=10) and observed distributions for the 24 hour maximum precipitation in May and August, Uppsala 1866-1982.

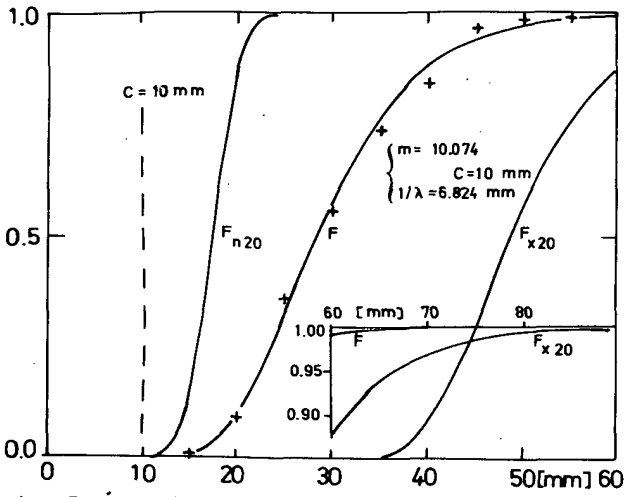


Fig. 7. The 24 hour maximum of the whole year (Max Cpe, c=10 mm) for Uppsala 1866-1982.

den. Generally stations in the more coastal regions (within about 25 km from the coast) and in regions with definitely higher average precipitation as at Gångarebo in the south-west region have clearly higher 24 hour maxima, both mean values and extreme values. Although the daily precipitation series from Uppsala has been analysed from 1836 and onwards no value above 60 mm within a consecutive day (6 GMT-6 GMT) has been observed during 147 years! This must be quite rare but on the other hand only few Swedish stations have observed above 100 mm in 24 hours during the 20:th century. But as an example the rather nearby Bromma airport just outside Stockholm received 82,7 mm during a few night-time hours in the beginning of June 1980.

7. SIMULATIONS OF HOURLY AND DAILY VALUES

For time intervals shorter than a month but still above about 15 days we can make use of the Cpe model thanks to its natural structure. Denoting the new period length by n' and the length of the month by n we can use the new primed estimates

$$\begin{cases} m' = m n'/n \\ \lambda' = \lambda \end{cases} \qquad (18)$$

Thus we can rapidly establish distributions for some other length intervals. For longer periods, like two months, we can of course make a convolution of two Cpe distributions. But for really short periods the simplification using instantaneous amounts becomes too unrealistic. A rather straightforward way to generalize the Cpe model is then to introduce a constant rain intensity r and the parameter μ as already introduced in Section 2. Although a far from realistic assumption in the real world we may use it to see if the time structure, or more precisely the autocorrelation of series of daily precipitation, of this modified Cpe model will be comparable to real data. This would also support the assumptions made, especially the one about independent events forming a Poisson process. Then the observed autocorrelations would merely result from the fact that a rain event lasts long enough to make nearby days correlated.

In Fig. 8 the Cpe model used so far and this slightly extended version are compared using two congruent realizations in order to clarify the difference.

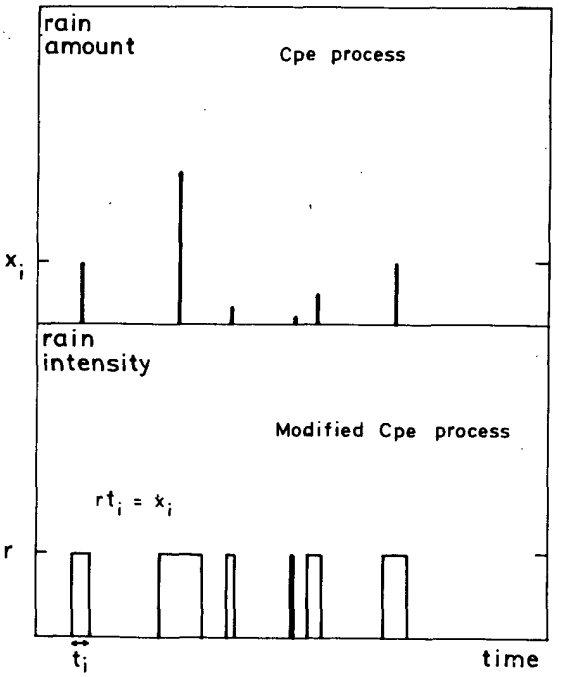


Fig. 8. Congruent realizations of the Cpe and the modified Cpe model.

The modified Cpe process used here is equivalent to a birth-death process with intensity matrix A defined as

$$A = \begin{bmatrix} -\nu & \nu & 0 & 0 & \dots \\ \mu & -(\nu+\mu) & \nu & 0 & \dots \\ 0 & 2\mu & -(\nu+2\mu) & \nu & \dots \\ 0 & 0 & 3\mu & -(\nu+3\mu) & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix} \qquad (19)$$

We have then allowed two events (or more) to go on simultaneously and then also the intensity of death is doubled. The average probability π_i of

the process of being in state i (i events going on simultaneously) then becomes

$$\pi_i = (\nu/\mu)^i \, e^{-\nu/\mu}/i \; ! \qquad (20)$$

To obtain an average precipitation intensity the mean monthly amount can be divided by the fraction of the time with observations of precipitation. This was done in A(1983) for Uppsala and this time fraction was also used for some other stations.

The average time of precipitation for winter months was about 21% and for summer months 11%. Then for Uppsala average precipitation intensities of 0.22 and 0.83 mm/hour are obtained for January and July respectively. For Gångarebo, and using 21 and 11% which is only approximately correct, the same intensities will be 0.61 and 1.39 mm/h. These rather low values arise from the quite large number of observations when just light precipitation falls, often not detectable by recording gauges.

Now we are ready to simulate e.g. hourly values from the m and λ estimates just using monthly series and the constant rain intensity value. Here hourly values were simulated for a number of Swedish stations and then mean values of autocorrelations of daily values were obtained for the two months. An observed series of daily values from Särna ($61^{o}41'$N, $13^{o}8'$E) was available on data and used in the comparison. The theoretically derived autocorrelations are listed in Table 2 while Fig. 9 depicts the observed correlations of the Särna series. In the simulations 250 years were constructed which reduced the uncertainty of the autocorrelations to a minimum. For the much shorter observed series the months were grouped in three classes to get more stable values. Nevertheless we can see that the slightly modified or extended Cpe model gives very reasonable time series of daily precipitation as far as autocorrelations and the gross time structure are concerned.

Table 2. Daily autocorrelations (ρ_i, i = time lag in days) as mean values for several stations and simulated years using the modified Cpe model.

	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
January, Sweden:	0.50	0.16	0.06	0.02	0.01
July, Sweden:	0.23	0.02	0.01	0.00	0.00

If we are interested in simulations of monthly values only a set of random numbers (preferrably uniformly distributed) and the Cpe distribution is sufficient in generating Cpe numbers. However, instead of using the complex Cpe frequency function, Eq. (4), it is easier to make the simulation in two steps beginning with the number of events (a Poisson variable with parameter m) and then going on with the amount at each event (an exponential variable with parameter λ). The amounts are simulated as integer values. Then sequences of independent Cpe variables can be simulated for hundreds of years. It may be quite instructive to make spectral analyses or calculate running averages on these independent numbers. Such analyses give surprisingly well defined long-term fluctuations indicating that many of the climatic fluctuations existing in real climate data may well be accidental and not caused by a true change of the basic physical parameters and boundary conditions of the atmosphere-ocean system.

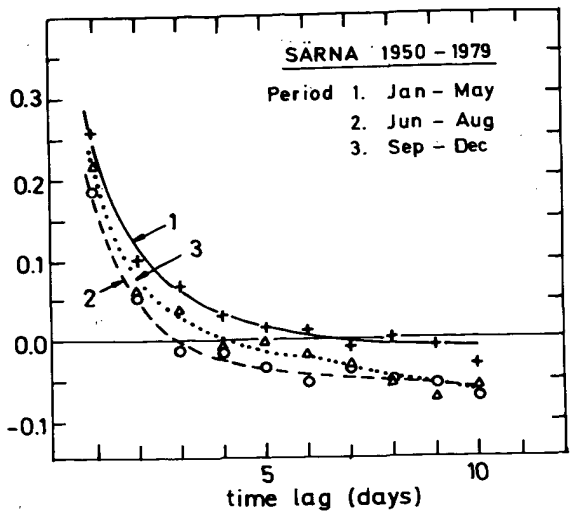


Fig. 9. Observed autocorrelations of daily precipitation, Särna 1951-1980.

Finally, let us look at the extreme months in sets of 500 year long series of Cpe distributions for Uppsala and Gångarebo. Also included is the observed minima and maxima during 147 and 69 years respectively (integer values).

Table 3. Extreme values (mm) in series of 500 years of simulated monthly precipitation amounts. Also included is the observed minima (n*) and maxima (x*) for the two stations, Uppsala (147 years) and Gångarebo (69 years).

	Uppsala				Gångarebo			
	n_{500}	n^*_{147}	x_{500}	x^*_{147}	n_{500}	n^*_{69}	x_{500}	x^*_{69}
Jan	2	4	102	79	6	12	281	205
Feb	0	1	90	85	0	6	262	205
Mar	0	0	107	69	0	2	197	157
Apr	0	3	85	79	0	0	209	197
May	0	3	143	120	0	0	237	153
Jun	4	10	130	110	0	13	219	210
Jul	5	5	224	200	19	30	341	231
Aug	0	3	220	167	0	16	421	309
Sep	4	11	165	154	10	25	320	268
Oct	1	5	191	135	0	1	389	289
Nov	4	4	138	112	7	21	301	251
Dec	0	8	115	130	5	18	274	260
Year	305	311	920	811	687	793	1609	1496

8. SUMMARY

The flexibility and vigour of the compound Poisson-exponential model makes it well suited for several applications in precipitation climatology. The basic assumptions are simple enough to get an analytical expression for integrated precipitation amounts, which was considered very important since the main amount of climatological data is published, and handled with, in form of monthly values. Then it also becomes rather easy to use the maximum likelihood technique to estimate the two parameters of the model, the mean number of precipitation events and the mean amount at each event. The Cpe model is also easy to use for deriving an extreme value distribution for the maximum amount at an individual event. The same dis-

tribution is then also used for the more concrete 24 hour maximum of a year or a month for which observed series often are available.

The Cpe model links together models simulating time series of precipitation, hitherto as a rule using Markov chains applied to sequences of dry and wet days, and empirical distributions used for integrated amounts or extreme daily amounts. The model is based on a Poisson process. The assumption of independent events in this process is clearly an idealization. On the other hand the mean number of independent precipitation events is a sensible parameter closely related to the degrees of freedom of the true precipitation process at a station.

So far we have exclusively been dealing with univariate statistics at one point. Some extentions of these analyses to precipitation fields, especially the spatial correlation, have been presented elsewhere (A, 1983).

An extended work along these terms is planned. Then it is hoped that it will be possible to use a few effective assumptions to get a realistic description not only of the gross features of the precipitation process but also of the frequencies of instantaneous rainfall intensities. Then it will be possible to derive different rainfall intensity statistics for which the interest has grown rapidly, see e.g. Jones and Sims (1978) and Tattelman and Graham (1982). Then it will be necessary to extend the input data to somewhat more special climate elements as perhaps the frequency of different types of precipitation using the idea of a mixed exponential distribution for the amount (Woolhiser and Pegram, 1979). It could then be possible to increase the number of stations useful for such studies beyond the very few ones equipped with high resolution precipitation recorders.

For practical applications it is important to know more exactly the uncertainty inherent in climatological data or theoretical distributions and parameters. A more thorough analysis of this is also planned to get a more precise knowledge of the uncertainty of different probabilities, return periods etc. based on the Cpe model.

REFERENCES

Alexandersson, H., 1983: A stochastic model of precipitation with applications, extensions and examples. *Reports No* 69, Dept. of Meteorology, University of Uppsala, Sweden, 102 pp.

Chuan, G.D, and J.G. Lockwood, 1974: An assessment of topographical controls on the distribution of rainfall in the central Pennines. *Met. Mag.*, 103, 275-287.

Corona, T.J., 1979: Further investigations of the interannual variability of Northern Hemisphere continental precipitation. *Env. Res. Paper No* 20, Colorado State University, Fort Collins, USA.

Delijanic, I., 1972: The characteristic changes of the amount of precipitation in connexion with altitude in the mountainous areas of Yugoslavia. Contribution in *Distribution of precipitation in mountainous areas, volume II, Technical papers, WHO/OMM No* 326, 278-289.

Eriksson, B., 1980: Statistical analysis of monthly precipitation data. Part II. Frequency analysis of monthly precipitation data. *RMK No* 17, SMHI, Sweden, 152 pp.

Feller, W., 1968: *An introduction to probability theory and its applications, volume I,* third edition. John Wiley & Sons Inc., New York, 509 pp.

Gabriel, K.R. and J. Neumann, 1957: On a distribution of weather cycles by length. *Quart. J. Roy. Meteor. Soc.*, 83, 375-380.

Jones, D.M.A. and A.L. Sims, 1978: Climatology of instantaneous rainfall rates. *J. Appl. Meteorol.*, 17, 1135-1140.

Lindgren, B.W., 1968: *Statistical theory,* second edition. The Mac-Millan Company, London, 521 pp.

Neumann, J., 1955: On the frequency of dry and wet spells at Tel Aviv. *Bull. Res. Counc. of Israel,* 5A, 28-39.

Tattelman, P. and D.D. Granthman, 1982: A survey of techniques for determining short-duration precipitation rate statistics. *AFGL-TR*-82-0357, Air Force Geophysics Laboratory, Mass., USA, 33 pp.

Thom, H.C,S., 1958: A note on the gamma distribution. *Mon. Wea. Rev,* 86, 117-122.

Woolhiser, D.A. and G.G.S. Pegram, 1979: Maximum likelihood estimation of Fourier coefficients to describe seasonal variations of parameters in stochastic daily precipitation models. *J. Appl. Meteorol.*, 18, 34-42.

EXPLORATORY ANALYSIS OF PRECIPITATION EVENTS IN WESTERN OREGON

Barbara G. Brown, Richard W. Katz, and Allan H. Murphy

Department of Atmospheric Sciences
Oregon State University
Corvallis, Oregon 97331, U.S.A.

1. INTRODUCTION

Statistical models of precipitation play an integral role in the analysis and forecasting of precipitation-sensitive phenomena. For example, long-term forecasts of soil erosion require specific types of information regarding the local characteristics of precipitation (e.g., precipitation intensity and duration). This information may be obtained through statistical analysis of the important precipitation characteristics for a large set of precipitation "events." In this context, a precipitation event may be defined in a manner that is appropriate for the application of interest, or simply in a way that is convenient. Of course, the results of the analysis must be interpreted in terms of the definition selected.

This paper describes an exploratory analysis of the characteristics of precipitation events in western Oregon. Section 2 contains a brief description of the data used in the study, and the precipitation event definition and analysis approach are discussed in Section 3. The results of an exploratory analysis of some characteristics of the precipitation events are described in Sections 4 and 5, including marginal distributions of the precipitation event characteristics (Section 4) and conditional distributions of some pairs of characteristics (Section 5). Section 6 contains a discussion of the results and their implications, and a comparison of these results with some results for other sites and for precipitation events created using other definitions. Finally, a brief summary is presented in Section 7.

2. PRECIPITATION DATA

The basic precipitation data used in this study consist of over 31 years (July 1948 to December 1979) of hourly precipitation amounts. These data were collected by the U.S. National Weather Service (NWS) at Salem, Oregon. The resolution of the hourly precipitation amounts is to the nearest 0.01 inches (0.25 mm), with a minimum measurable value of 0.01 inches (i.e., trace values are excluded).

Salem is located in the Willamette Valley in western Oregon between the coastal and Cascade mountain ranges. Ninety percent of the average annual precipitation at Salem (41 inches/104 cm) falls between 1 October and 31 May, with very little (none in some years) falling as snow. The analyses described in this paper were restricted to data from the wet season (October to May).

3. ANALYSIS OF PRECIPITATION DATA

For this study, a precipitation event is defined to be a period of one or more consecutive hours of precipitation followed by one or more hours of no precipitation. That is, two adjacent "wet" hours are considered to be part of the same precipitation event, whereas two wet hours that are separated by one or more dry hours are considered to belong to separate precipitation events. Precipitation events defined in this way can be termed "primitive" events. They are based on the simplest and most logical approach to formulating precipitation events without the use of additional information (e.g., regarding the synoptic weather situation). Moreover, such events can easily be combined to create composite events that may be more appropriate for a particular application (e.g., for modeling the processes leading to soil erosion).

Several characteristics of precipitation events are of interest. These characteristics include the event duration, magnitude, average intensity, and maximum intensity, and the time between events. The duration of an event is the total number of consecutive hours of precipitation. The magnitude is the total amount of precipitation that occurs during the event (i.e., the sum of the hourly amounts). The average intensity of an event is the average rate of precipitation per hour during the event. Thus, magnitude equals average intensity times duration. The maximum intensity of an event is the maximum hourly amount of precipitation within the event. Time between events is the number of dry hours separating the event from the previous event.

The hourly precipitation amounts for Salem were summarized by combining them into primitive precipitation events. A total of 8523 precipitation events were thus formulated from the record of hourly precipitation amounts. Each of the five

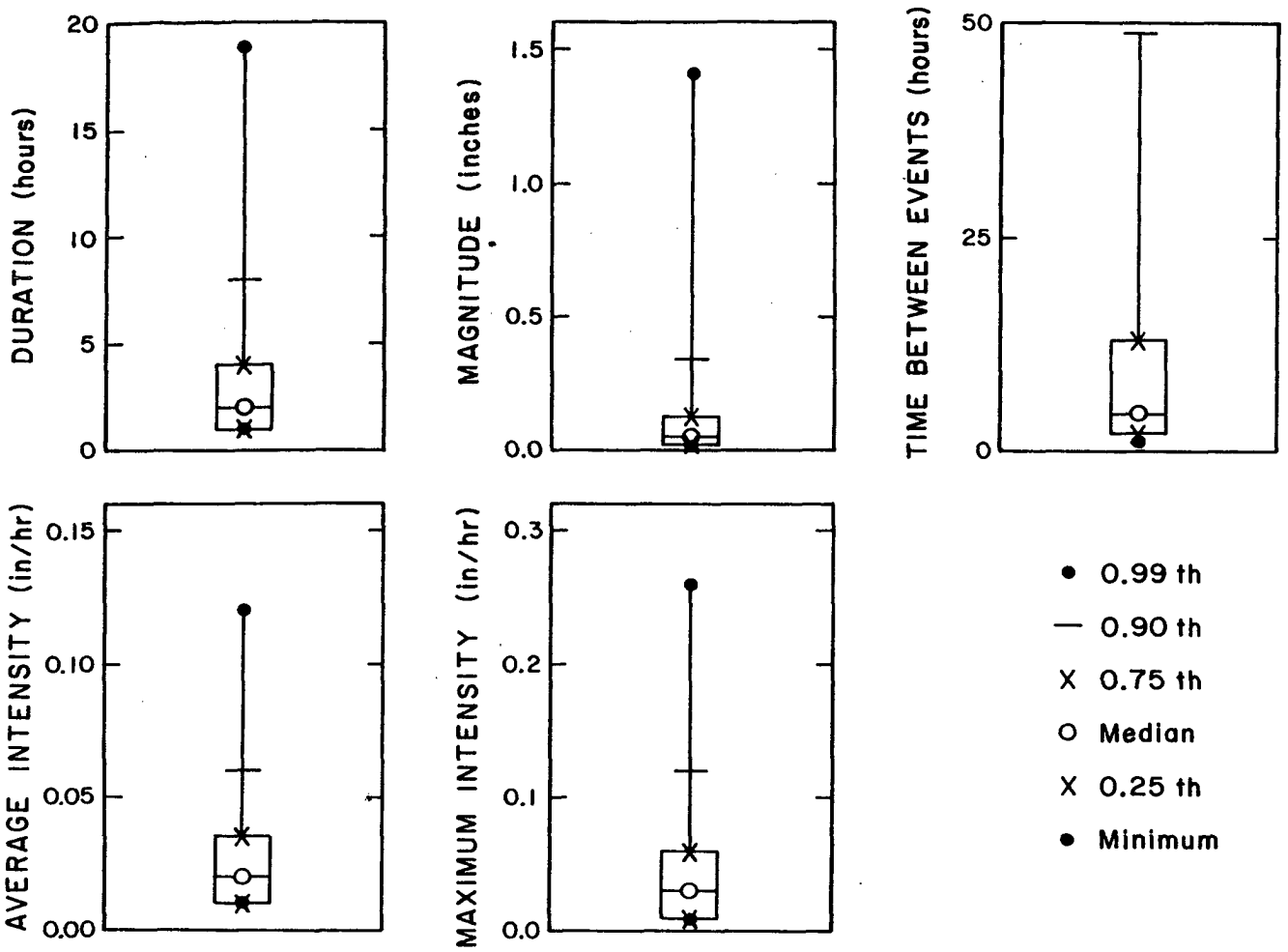


Figure 1. Box plots of the distributions of the precipitation event characteristics. The key for identifying the quantile values shown is given in the lower right corner.

precipitation event characteristics (duration, magnitude, average intensity, maximum intensity, and time between events) was calculated for each event.

The results of an exploratory analysis of the characteristics of the primitive precipitation events are described in the following two sections. This analysis consisted of examination of individual characteristics and of pairs of characteristics. First, in Section 4 some properties of the individual characteristics and their distributions are considered. Second, in Section 5 the relationships between some pairs of the characteristics are examined through the evaluation of conditional distributions of the characteristics.

4. MARGINAL DISTRIBUTIONS

The marginal distributions of the values of the precipitation event characteristics can be conveniently displayed using box plots (Tukey, 1977). These plots provide a concise method of presenting several properties of the distributions at the same time. Box plots of the distributions of the five precipitation event characteristics are shown in Figure 1.

A major feature that is illustrated by the box plots in Figure 1 is the skewed nature of the distributions of the five characteristics. That is, most of the precipitation events had relatively small values of the characteristics, whereas a few events had relatively large values. For example, the upper quartile value of magnitude is 0.13 inches and the 0.99th quantile value is 1.41 inches; in other words, 75 percent of the events produced less than 0.13 inches of precipitation, whereas 1 percent produced more than 1.41 inches. This result is not unexpected since the skewed nature of the distributions of precipitation variables has been noted in previous studies (e.g., Biondini, 1976).

The box plots in Figure 1 can be used to obtain other general information about the distributions of the precipitation event characteristics. For example, note that the quantile values of the distribution of maximum intensity are approximately twice as large as the corresponding quantile values of the distribution of average intensity. This result suggests that, on the average, the largest hourly rainfall amount during a precipitation event is approximately twice as large as the average hourly amount. Moreover, specific probability estimates may be taken directly from the box plots (e.g., the estimated probability of

occurrence of a precipitation event with a maximum intensity exceeding 0.12 in/hr is 0.10).

The extreme values of the characteristics that can be expected to occur in a 31-year period (the period of record) and in a single year can be estimated using order and return statistics, respectively. The order statistics are simply the values of a particular characteristic ranked from largest to smallest. The return statistics consist of every 31st order statistic. Thus, the first n order statistics for a particular characteristic provide estimates of the n largest values of the characteristic that could be expected to occur in a 31-year period, whereas the first n return statistics provide estimates of the n largest values that could be expected in a single year. For example, a precipitation event that lasts as long or longer than the tenth duration order statistic would be expected to occur approximately ten times in 31 years. A precipitation event that lasts as long or longer than the tenth duration return statistic would be expected to occur about ten times in a single year. The order and return statistics are useful for determining estimates of the probabilities of occurrence of specific values (particularly extreme values) of the precipitation event characteristics.

The first 100 maximum intensity order and return statistics are shown in Figure 2. As would be expected based on the maximum intensity box plot in Figure 1, the first few order statistics of maximum intensity are very large, and later values are much smaller. Moreover, the slope of the order statistic curve is quite steep for the first few values and much flatter for later values. This result is due to the skewness of the distribution of maximum intensity illustrated in Figure 1. The return statistic curve has a similar shape. However, since a larger portion of the range of maximum intensity is covered by the return statistics, the slope is even steeper for the early return statistics than for the early order statistics. The curves in Figure 2 provide much information about the distribution of extreme values of maximum intensity. For example, based on these statistics, an event with a maximum intensity of at least 0.40 in/hr can be expected to occur about ten times in 31 years; an event with a maximum intensity of at least 0.20 in/hr can be expected seven times in a single year. The curves of the order and return statistics for the other precipitation event characteristics are quite similar to the maximum intensity curves shown in Figure 2 and thus are not presented here.

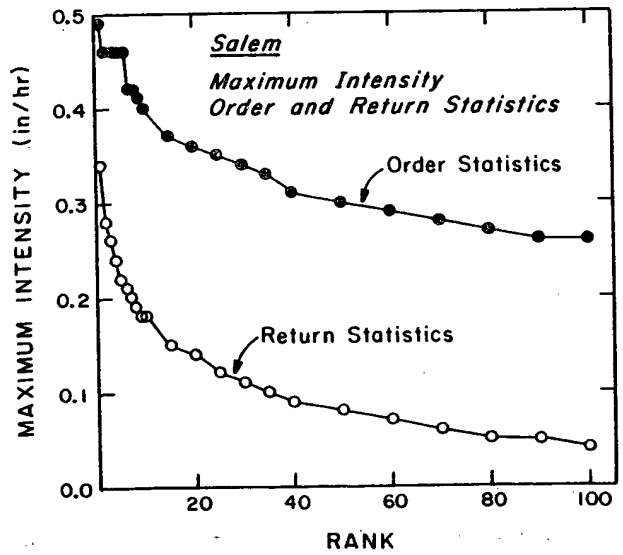


Figure 2. Maximum intensity order and return statistics.

5. CONDITIONAL DISTRIBUTIONS

It is of interest to consider the distributions of pairs of precipitation event characteristics in order to investigate the types of relationships that exist among the characteristics. In particular, the relationships between the precipitation event duration and the characteristics magnitude, average intensity, and maximum intensity are examined in this section. Another type of dependence that could be investigated is the relationship between the characteristics and the time between events. However, previous evaluations of these data have provided no indication that this type of dependence exists.

The conditional distributions of the characteristics given many different values of duration can provide some insights into the types of relationships that exist between the precipitation event characteristics and the duration of the event. These distributions can most easily be displayed by means of "moving" statistics as in Figure 3. Similar displays have been employed by Cleveland and Kleiner (1975) and Katz (1978). Figure 3 contains the quantile values of the conditional distributions of magnitude given values of duration ranging between 1 and approximately 30 hours. For example, the median magnitude for events lasting 10 hours is about 0.45 inches. The corresponding quantile values for the individual distributions are connected by lines so that four curves are presented: the 0.25th, 0.50th, 0.75th, and 0.90th quantile values of the conditional distributions. The curves were smoothed using

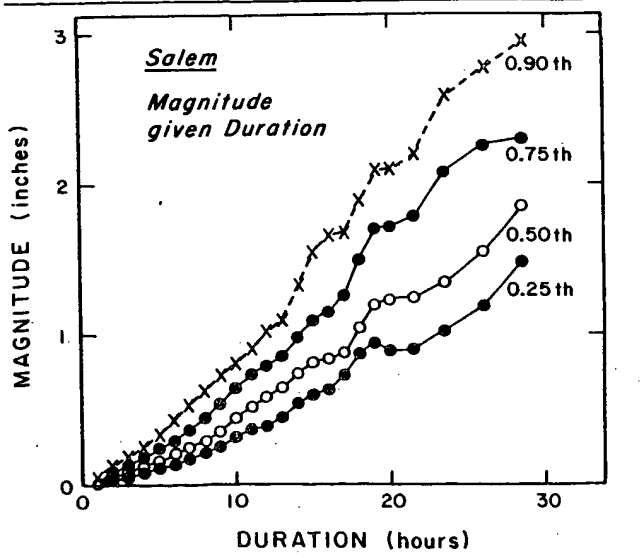


Figure 3. Quantile values of the conditional distributions of magnitude given duration.

hanning (Tukey, 1977) in order to make the trends with increasing duration more visible. Note that due to the smaller numbers of events with large durations, the concomitant quantiles are for conditional distributions given an interval of duration values rather than a single value.

The curves in Figure 3 indicate that there is a strong relationship between precipitation event duration and magnitude: the longer it rains, the greater the amount of precipitation accumulated. This relationship, which is not unexpected, is evidenced by the increasing values of the conditional magnitude quantiles with increasing values of duration. Moreover, the variability of magnitude appears to depend on the duration, as indicated by the increase in the distance between the upper and lower quartiles (i.e., the interquartile range) as the duration becomes larger.

The quantile values of the conditional distributions of average intensity given duration are displayed in Figure 4. The trend in these

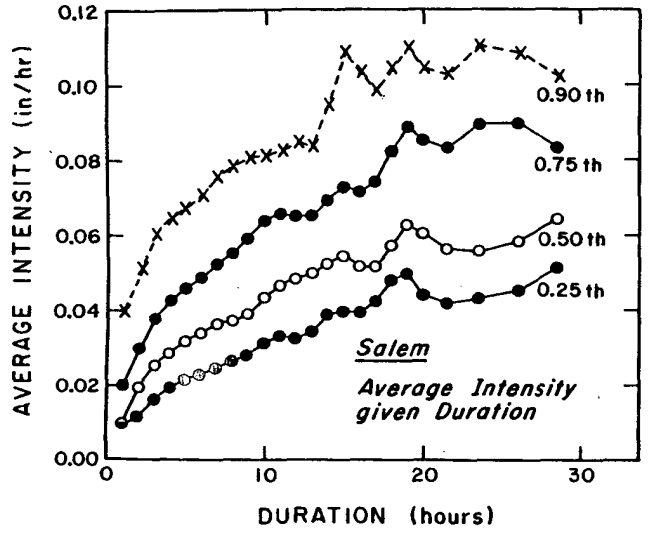


Figure 4. Quantile values of the conditional distributions of average intensity given duration.

quantile values is not as strong as the trend for the quantile values of magnitude given duration. Nevertheless, there clearly is a relationship between these two characteristics, as indicated by the gradual increase in the conditional average intensity quantiles with increasing values of duration. For example, the median value of average intensity for events lasting 5 hours is only slightly more than 0.03 in/hr, whereas the median value of average intensity for events lasting 20 hours is about 0.06 in/hr.

The existence of a relationship between average intensity and duration could not have been anticipated. Although average intensity is functionally related to duration (magnitude equals average intensity times duration), it also is functionally related to magnitude. Without knowledge of the exact nature of the relationship between magnitude and duration, it is not possible to assume a priori the form of the relationship between average intensity and duration.

The relationship between maximum intensity and duration also is quite strong, as indicated by the curves in Figure 5. The conditional quantile

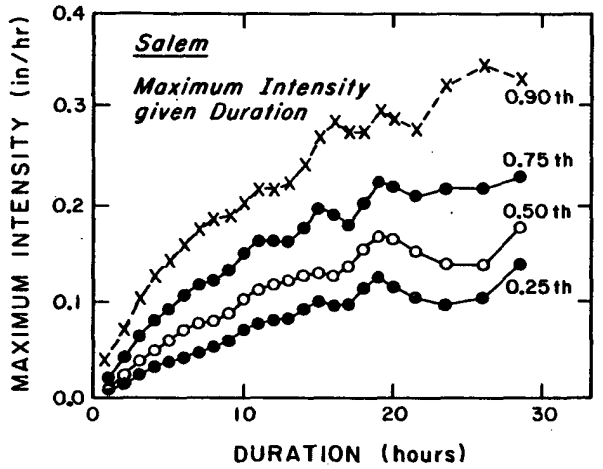


Figure 5. Quantile values of the conditional distributions of maximum intensity given duration.

values of maximum intensity and the interquartile range are increasing functions of duration, at least for events lasting 20 hours or less. Apparently, the probability of a large maximum intensity value is greater for precipitation events that last a long time than it is for short-duration events. For example, less than 10 percent of the events lasting 5 hours had a maximum intensity of at least 0.10 inches, whereas this percentage is greater than 50 for events lasting 10 hours. The implications of this result and the others described in this section are considered in Section 6.

6. RELATED RESULTS AND DISCUSSION

In a related study (Brown et al., 1983a,b), the results presented in Sections 4 and 5 were found also to qualitatively characterize a more complex definition of a precipitation event at Salem and at several other locations as well. In that study, similar analyses to those described in this paper were applied to precipitation data from five sites using a precipitation event definition that is appropriate for estimating the effects of precipitation in causing soil erosion. The precipitation events were derived from the primitive precipitation event data base by taking the union of any events separated by less than six hours. The resulting precipitation events will be denoted here as "composite events."

The five sites included Salem as well as Portland and Eugene (also in western Oregon), and two NWS stations, Pendleton and Walla Walla, located in eastern Oregon and Washington, respectively. Portland and Eugene are situated in the Willamette Valley and have climates similar to that at Salem. However, Pendleton and Walla Walla are on the eastern side of the Cascade mountains in a region that is both drier and colder in the winter than the Willamette Valley. The basic data for

the additional stations are of the same type as the basic data for Salem described in Section 2 (i.e., hourly precipitation amounts from mid-1948 to the end of 1979).

The results of analyses of the characteristics of the composite events at the five stations are generally similar to the results for the primitive events at Salem. In particular, the marginal and conditional distributions of the characteristics for the Willamette Valley sites (Portland, Salem, and Eugene) are all qualitatively similar to the results presented in Sections 4 and 5. Moreover, many features of the results for the Pendleton and Walla Walla composite events are similar to the primitive event results for Salem. However, one major difference in the results is that there is apparently no relationship between average intensity and duration at Pendleton and Walla Walla, as indicated by the conditional distributions of average intensity given duration (e.g., see Figure 6). Such a relationship is

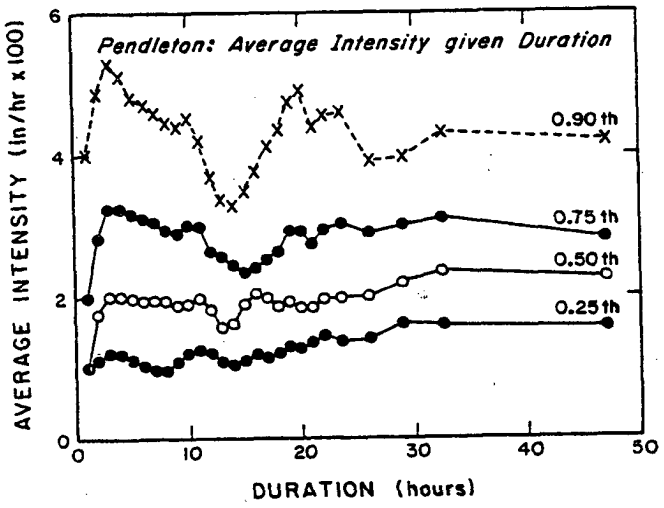


Figure 6. Quantile values of the conditional distributions of average intensity given duration for Pendleton composite events.

evident for the Portland, Salem, and Eugene composite events.

It is possible that the existence of a relationship between average intensity and duration for the Willamette Valley events, as indicated by the conditional distributions, could have resulted from a seasonal dependence of both characteristics. For this reason, the relationship also was evaluated for a small subset of the Salem primitive events (those events ending in January). The results of this analysis again indicated that the average intensity of a precipitation event is dependent upon the duration of the event. This relationship and the strong relationship between the maximum intensity and duration of precipitation events in western Oregon suggest that the hourly rainfall amounts within the events have a particular structure and relationship to one another. The same type of structure apparently does not exist within the Pendleton and Walla Walla precipitation events. Nevertheless, these results indicate that the assumption that hourly rainfall intensities are independent and identically distributed given the precipitation event duration, made in some stochastic models of precipitation (e.g., Nguyen and Rousselle, 1981), is not always valid.

7. SUMMARY AND CONCLUDING REMARKS

The results presented in the preceding sections provide a general statistical description of primitive precipitation events in western Oregon. Some conclusions that can be drawn regarding the characteristics of these precipitation events include the following:

(a) The marginal distributions of the characteristics are strongly positively skewed. That is, precipitation events with relatively small values of the characteristics occur most frequently and events with large values of the characteristics occur relatively infrequently.

(b) The most extreme values of the precipitation event characteristics generally are much larger than the values for less extreme events, as evidenced by curves of order and return statistics.

(c) The values of some pairs of precipitation event characteristics are not independent. For example, the values of precipitation event magnitude, average intensity, and maximum intensity are all related to the duration of the precipitation event.

These conclusions also apparently characterize, in general, other types of precipitation events as well as precipitation events at some other locations, as described in Section 6. The relationship between average intensity and duration which was found to be a feature of precipitation events in western Oregon, however, is apparently not a characteristic of precipitation events at all locations. This result suggests that the structure and relationships of hourly rainfall amounts within precipitation events would be an interesting topic for further investigation.

The tools of exploratory data analysis have produced a large amount of information in this study about the statistical properties of precipitation events. The results of the analyses have provided general descriptions of the precipitation event characteristics and insights into the relationships between the characteristics. Clearly, the use of exploratory data analysis is advantageous and rewarding in this type of study where the variables possess both high variability and asymmetric distributions, and where the objective is to obtain a general description of the statistical properties of a phenomenon. Moreover, the analyses described in this paper could easily be used to study other types of climatological events (e.g., wind events) that may be of interest, or other types of precipitation events that may be appropriate for different applications (e.g., for forest fire forecasting).

REFERENCES

Biondini, P., 1976: Cloud motion and rainfall statistics. Journal of Applied Meteorology, 15, 205-224.

Brown, B. G., R. W. Katz, A. H. Murphy, and J. D. Istok, 1983a: Precipitation events and frequency distributions of precipitation event characteristics in western Oregon. Corvallis, Oregon State University, Department of Atmospheric Sciences, Report (in preparation).

Brown, B. G., R. W. Katz, A. H. Murphy, and J. D. Istok, 1983b: Precipitation events and frequency distributions of precipitation event characteristics in eastern Oregon and Washington. Corvallis, Oregon State University, Department of Atmospheric Sciences, Report (in preparation).

Cleveland, W. S., and B. Kleiner, 1975: A graphical technique for enhancing scatterplots with moving statistics. Technometrics, 17, 447-454.

Katz, R. W., 1978: Persistence of subtropical African droughts. Monthly Weather Review, 106, 1017-1021.

Nguyen, V.-T.-V., and J. Rousselle, 1981: A stochastic model for the time distribution of hourly rainfall depth. Water Resources Research, 17, 399-409.

Tukey, J. W., 1977: Exploratory Data Analysis. Reading, MA, Addison-Wesley, 688 pp.

A STATISTICAL MODEL FOR PRECIPITATION DATA - WITH APPLICATION TO INTERPOLATION PROBLEMS

PETER ALLERUP AND HENNING MADSEN

METEOROLOGISK INSTITUT
KØBENHAVN, DENMARK

1. INTRODUCTION

The present research work is part of an investigation started in consequence of a WMO draft concerned with climatic atlases and standard normals 1931-60. It gave rise to the preparation of new standard normals, and in this connection it was necessary to develop statistical methods for homogeneization and interpolation of missing values primarily concentrated on precipitation.

However, first of all an extensive analysis was undertaken leading to a comprehensive statistical model for recorded variations in monthly precipitation values (Allerup (1981)). This model is then applied to interpolations, and it is demonstrated that the results depend on regional and temporal variations of the precipitation conditions.

Many attempts to build adequate statistical models for the purpose of analyzing precipitation data may be found in the literature. Some authors have attached importance especially to the problem of estimating the correct shape of precipitation distribution (Shenton and Skees(1970), Shenton and Bowman (1973), Buishand(1977)). The stochastification problem has been in focus in these analyses. Another major part of statistical model building is concerned with methods from the theory of stochastic processes (Gabriel and Neumann (1962), Green (1970), Gates and Tong(1976), Matalas and Langbein (1962),Iyengar(1981)). In the ligh of such methods, concepts of autocorrelation and crosscorrelation between various series are the analytic main tools. Furthermore, the use of process theory has been much increased by introduction of the random field theory into meteorological subjects (Meija and Iturbe (1974 a,b,c), Gandin (1970), Cliff and Ord (1975), Lenton and Iturbe (1977), Niper and Streit (1982).

However, many of these methods do not provide a satisfactory explanation to the empirical facts found in monthly point precipitation data, and they are not specifically directed towards interpolation purposes. We shall, therefore, emphasize empirical structures valid for monthly point precipitation values and take advantage of these findings in theoretical formalizations leading to proposals for a relevant statistical model. It turns out, that some simplifications of the theoretical stuctures proposed are needed to bring about workable methods in practical interpolation problems. It is, however the intention also to present some theoretical consequenses of various formalizations of the empirical findings.

2. DATA FOR THE ANALYSIS

The fundamental data used for this investigation were selected according to the following main criteria:

a. Both coast stations and inland stations are present in the sample.
b. The stations are selected so as to provide a broad spectrum of distances between the stations.
c. The stations have measurements from a 30-year period without interruptions (i.e. having no interpolated values at all).

44 precipitation stations in the period 1931-60 were selected in Denmark, see Fig. 1. Data consist of 30 repetitions of monthly recorded precipitation values, where every calendar month has been analyzed separately. The analysis has confirmed the gain from this sub-division of the year into homogeneous periods. In Jutland two areas are encircled calling attention to stations especially used for illustration of interpolations: Seven stations with very short interdistances representing what could be termed an "adequate"

choice for interpolation aims and six stations un-evenly spread at inland and on different coasts representing what will be called an "inadequate" choice.

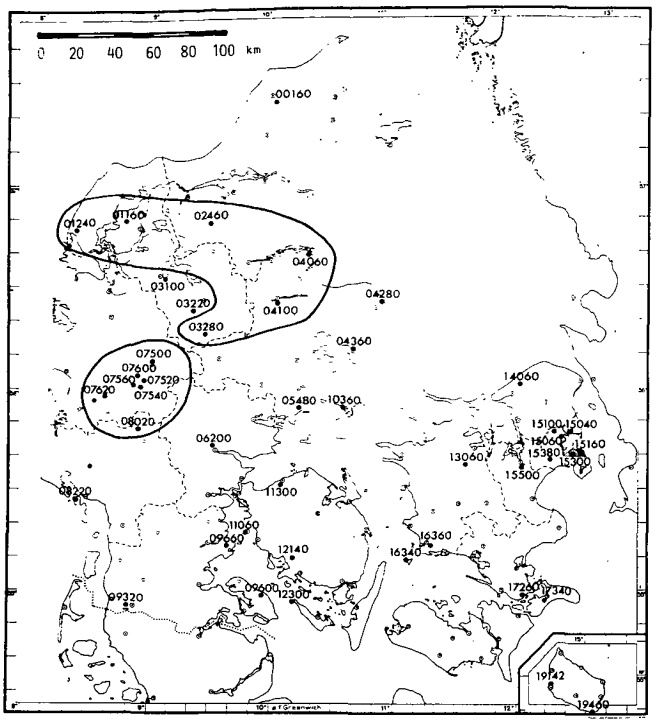


Fig. 1. 44 precipitation stations in Denmark. Two subsets of stations for interpolations are encircled.

3. EMPIRICAL STRUCTURES IN MONTHLY PRECIPITATION

Relation between monthly precipitation values at two stations is illustrated in Fig.2. This figure represents but one example of pairwise combination of values from stations in a specific month over 30 years but highlights structures typical for all such relations:

1. The points are distributed around a straight line through origo with slope determined by the pair of stations plotted.

2. The variance of the points around the straight line increases with increasing values of precipitation.

Characteristic 1 is well known to meteorologists; it means that the ratio of precipitation remains nearly constant for two neighbouring stations over a period of time.

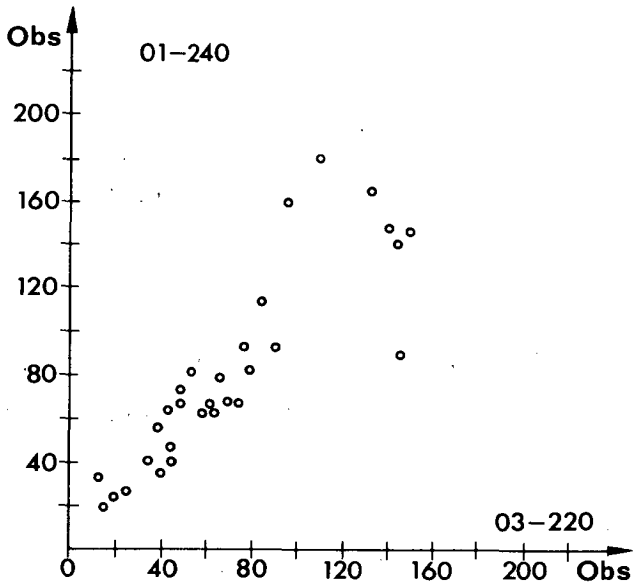


Fig. 2. Relation between observed precipitation values.

A logarithmic transformation of precipitation values bring about the following characteristics, which partly explains the structural content of 1. and 2. above. Two examples of logarithmic relations for a fixed month over 30 years are given in Fig. 3 and Fig. 4.

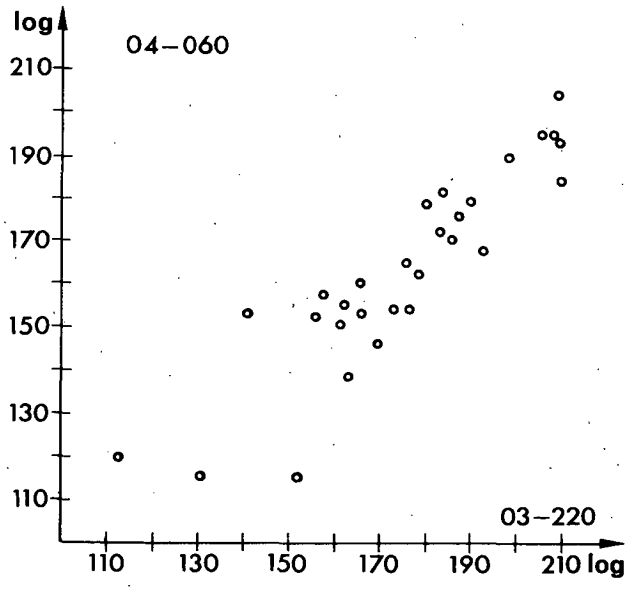


Fig. 3. Relation between logarithmic precipitation values.

3. The logarithmic difference of the precipitation remains on average constant, i.e. the points are scattered around a straight line with slope 1 and intersection determined by the pair of stations selected.

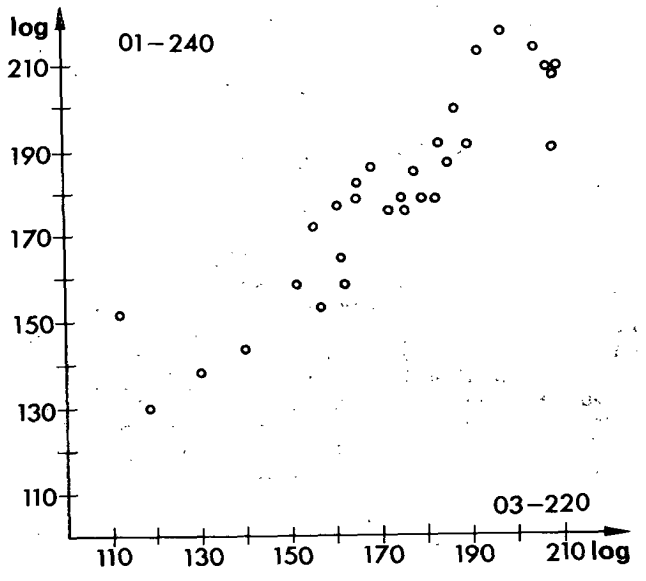


Fig. 4 Relation between logarithmic precipitation values.

4. The range of variation for the points around the line is fairly constant for large values.

5. Total variance increases with increasing distance between the stations.

6. For small values of precipitation the variation of logarithmic difference is increasing, i.e. the graph is more "open" towards origo.

Again the two examples in Fig. 3 and Fig.4 are just two visualizations of generally valid relations among logarithmic precipitation measurements - the only specification omitted in 3.-6. being the influence of the selected month of the year.

4. FORMALIZATION OF EMPIRICAL STRUCTURES

Let x_{it} denote logarithmic precipitation at station No. i on year No. t for a specific month. For the sake of clearness, notation for the selected month will be omitted in the following. Characteristic 1 and 3 implies for the expected values.

$$E(x_{it}) = \xi + \alpha_i + \beta_t \qquad (1)$$

stating that $E(x_{it})$ can be separated, additively into α_i - specific for the site, β_t - specific for the year and a common level parameter ξ.
Since

$$V(x_{it}-x_{jt}) = V(x_{it}) + V(x_{jt}) - 2\mathrm{cov}(x_{it},x_{jt}) \qquad (2)$$

is a measure of the strip width around the straight line (cf.4), it is evident that influence from distance (cf.5) between stations can be formalized as changes in correlation, $\mathrm{cov}(x_{it},x_{jt})$. On the other hand it seems plausible to assume independence between measurements taken from different years t_1 and t_2.

The additive structure (1) suggests investigation of the residuals

$$R_{it} = x_{it} - x_{.t} - x_{i.} + x_{..} \qquad (3)$$

to look for the distribution of stochastic errors. Inspection of R_{it}'s reveal symmetrical distributions well fitted by the normal law. Hence, summarizing the above we could make the following model:

$$x_{it} \sim Nf(\xi + \alpha_i + \beta_t, \sigma^2_{it})$$

$$\mathrm{cov}(x_{it},x_{jt}) = \sigma_{ijt} \qquad (4)$$

$$\mathrm{cov}(x_{it_1},x_{jt_2}) = 0, t_1 \neq t_2, \text{all } i,j$$

This model specification contains too many parameters to be identifiable, but since the additive mean value structure seems inevitable the reduction ought to take place with the σ_{ijt}'s. Allerup (1981) considers various simplifications of the variance covariance structure in order to make the parametric structure identifiable and estimable. Evidently, the distance between stations will enter into these attempts as an independent, external variable to describe changes in covariances. It turns out, that a key problem is a proper formalization of the "opening" towards origo, mentioned in 6.(Fig.3 and Fig.4).

If we choose to make the following approximation

$$\sigma_{it}^2 = \sigma_i^2$$

$$\sigma_{ijt} = \sigma_{ij} \qquad (5)$$

i.e. independence of time variation t, we have chosen to simply ignore the opening, leaving the variation of points around the straight lines in Fig.3 and Fig.4 to be bounded by parallel lines. With S_{ij} and S_{ii} denoting ordinary covariances and variances among the residuals R_{it} (over time t) Allerup (1981) demonstrates that

$$S_{ij} \approx \sigma_{ij} - \sigma_{i.} - \sigma_{.j} + \sigma_{..}$$

$$S_{ii} \approx \sigma_i^2 - \sigma_{i.} + \sigma_{..} \qquad (6)$$

which makes residuals of the variances and covariances estimable - unbiased. Letting $\Sigma = ((\sigma_{ij}))$ and $M = (\mu_1,\ldots,\mu_N)$,

M being the characteristic roots of $((S_{ij}))$ it appears that a complete solution to (6) can be written as

$$\Sigma = CMC' + CQC' \qquad (7)$$

with C being an orthogonal transformation to be determined. Indeterminances in (6) have thereby been fixed to indeterminances in the Q-matrix, which is symmetric and contains arbitrarily chosen values in i'th row and i'th column (i=1,...,N) and zero's elsewhere. One way of removing arbitrariness in (7) is to assume "constant dependency on average" between a station and all neighbours, i.e. assuming Σ to have the 1-vector as characteristic vector. Under such circumstances it can be shown that the simple covariance S_{ij} between residuals estimates conditional, partial correlation between station No.i and NO.j:

$$S_{ij} \approx \mathrm{corr}(x_{it}, x_{jt} | x_{.t}) \qquad (8)$$

For network with interdistance - and therefor correlations - fulfilling this rather restrictive assumptions, (8) may serve as basis for further work. However, the 44 stations in Denmark cannot meet this condition and we will have to leave such solutions to the general problem concerning identification of the σ_{ijt}'s. Allerup (1981) has shown the mathematical consequences of describing the "opening" towards origo (cf.6, Fig.3 and Fig.4) by the following relation

$$V(x_{it} - x_{jt}) = \varkappa_t \cdot \theta_{ij} \qquad (9)$$

formalizing this empirical feature in the log-log relation as a product of a time specific parameter and a parameter specific for the station pair (i,j). It is furthermore justified how θ_{ij} can be related to interdistance D(i,j) between stations. In fact, it seems in accordance with the structures found in data when specializing (9) to

$$V(x_{it}-x_{jt}) = \varkappa_t \quad [D(i,j)]^{\tau} \qquad (10)$$

This formalization is, however tied up with an assumption on the variances $V(x_{it})$ being t-dependent but independent - approximately - of the station (i) in question (see Allerup (1981) for details). The general distance independent formalization (9) in combination with the additive mean value structure (1) gives rise to less identification and estimation problems than the approximative relations (5), but, even the use of the stronger assumption (10) leaves great practical estimation problems to be solved. Besides, it seems meteorologically pressing to specify variances dependent on the site in question.

Allerup, Madsen and Riis (1982) and Allerup (1982) suggest a suitable compromise between (9), (10) and (5) when solving problems of estimating areal precipitation. We will utilize the same formalization for the present interpolation purposes. In fact our formalization of a statistical model will be

$$x_{it} \sim Nf(\xi+\alpha_i+\beta_t, \sigma_i^2)$$

$$\mathrm{cov}(x_{it_1}, x_{it_2}) = 0, t_1 \neq t_2 \qquad (11)$$

$$\mathrm{cov}(x_{it}, x_{jt}) = \gamma(\beta_t)[D(i,j)]^{\tau}$$

(11) allows for the basic multiplicative structure stated in (9) but specifies the covariance - or the so-called correlation function (e.g. Gandin (1970) - to be multidimensionally dependent on the level parameter β_t, the general "loading" of precipitation for year No. t. Estimation within the model (11) then takes place after grouping the years according to β_t values along well known lines from the normal distribution theory. The correlation function is estimated by means of

$$\log(1-\rho_{ij}) = \log\gamma(\beta_t) + \tau \log D(i,j) \qquad (12)$$

ρ_{ij} being the correlation between station No. i and No. j.

For the sake of simplicity the presented interpolation calculations have been based upon one correlation function, obtained by pooling together the various β_t-specific correlation functions.

5. INTERPOLATION

Interpolation of precipitation values - on the first time logarithmic values - can now be made utilizing model (11) as follows. Assume, that at time $t = t_0$ we want to interpolate an observation $x_{t_0 q}$ at a certain site q. As interpolated value we will calculate the expected value, E_q by means of (11). However, this expected value should be calculated conditional upon what is observed in a suitable distance from the station.

Hence, as "estimate" of the unobserved value $x_{t_0 q}$ we get

$$E_q = E(x_{qt_0} | x_{i_1 t_0}, \ldots, x_{i_p t_0}) \approx x_{t_0 q} \qquad (13)$$

where p neighbour observations $x_{i_1 t_0}, \ldots, x_{i_p t_0} = \underline{x}^{(p)}$ are observed.

In order to get the conditional distribution $p(x_{qt_0} | \underline{x}^{(p)})$ of x_{qt_0} for a given set of observations $\underline{x}^{(p)}$ we write (11) as a simultaneous distribution of x_{qt_0} and the neighbours $\underline{x}^{(p)}$.

$$P(\underline{x}^{(p)}, x_{qt_0}) = Nf\left\{\begin{vmatrix} m_1 \\ m_2 \end{vmatrix}, \begin{vmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{vmatrix}\right\} \qquad (14)$$

m_2, v_{22} are scalars, m_1, v_{12} and v_{21} are vectors and v_{11} a p x p matrix, all of which

are just re-organizations of the parameters in (11) for the smaller set:
$x_{qt_0}, x_{i_1t_0} \ldots x_{i_pt_0}$.

From the distribution (14) we calculate the conditional mean value E_q and the conditional variance V_q (e.g. T.W.Anderson (1958)):

$$E_q = m_2 + v_{21}\, v_{11}^{-1}\, (\underline{x}^{(p)} - m_1) \qquad (15)$$

$$V_q = v_{22} - v_{21}\, v_{11}^{-1}\, v_{12} \qquad (16)$$

Hence, the interpolated value is obtained from (15) and as a measure of interpolation accuracy the variance (16) can be used. This leads to the following 95% confidence interval for the unobserved original value $y_{qt_0} = \exp(x_{qt_0})$:

$$[\exp(E_q - 1.96\sqrt{V_q}), \exp(E_q + 1.96\sqrt{V_q})] \qquad (17)$$

6. EXAMPLES OF INTERPOLATION

As mentioned we have selected for illustrative purposes two sets of stations in Jutland (see Fig.1). One set consisting of seven stations with short interdistances - from approx. 4km to approx. 40km - situated in the interior of Jutland. These stations are first considered in the summer period, June and July together and, secondly in autumn, October and November together.

Another set, the so-called inadequate choice consist of six stations deliberately chosen within an area showing different precipitation conditions. The interdistances vary from approx. 25km to approx. 100km. These stations are considered in the summer period, June and July together.

Although we have stressed the importance of analyzing monthly precipitation into separate months, we have taken two months together, since they are very much alike from a meteorological point of view, and provide 60 instead of 30 consequtive values for the estimation of local covariances and variances. The effect of utilizing just one correlation function for all Denmark has not yet been studied, but e.g. obvious coast - inland contrasts makes it advantageous to estimate region-specified correlation functions.

In both sets one station is selected as "unknown", i.e. to be interpolated and the success of the interpolation is then evaluated by relating observed precipitation to interpolated values (Fig.8-10) over the 60 events. Note, that five, respectively six stations are used for estimation of local region specific correlation functions, while the actual interpolation calculations are based on just three stations in each group. In fact, the main formula (15) contains the structural term m_2 to be estimated as

$$m_2 = \left(\frac{\alpha_{i_1} + \cdots + \alpha_{i_p}}{p} + \beta_t\right) \qquad (18)$$

with p = 3 for the three most close by neighbours. $\underline{x}^{(p)} - m_1$ are residual values R_{it} for these three stations. V_{21}, V_{11}^{-1} are matrices containing a.o. covariances between the unknown station and the other stations to be estimated by means of a smoothed correlation function (11) (or (12) utilizing all five, respectively six stations in the region.

7. CORRELATION FUNCTIONS

Fig.5 - Fig.7 show the empirical relation (12) for the two regions. All three relations follow the expected linear structure, (12) but with obvious varying degree of precision. The two summer relations (Fig.6 and Fig.7) are both showing greater variation of points around the straight line than what is seen in the autumn relation (Fig.5). This may be justified from different structures of precipitation patterns - predominantly convective precipitation in the summer and frontal precipitation in autumn. Just one point in Fig. 5 representing a pair of stations which are more closely correlated than expected from

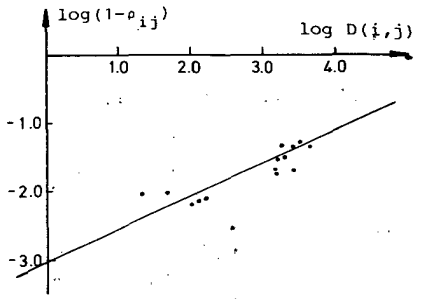


Fig. 5. Correlation - distance relation for region with 7 stations. Oct. and Nov. 1931 - 60.

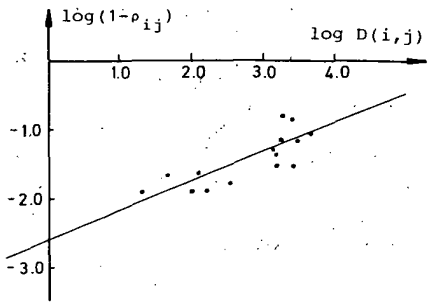


Fig. 6. Correlation - distance relation for region with 7 stations. June and July 1931 - 60.

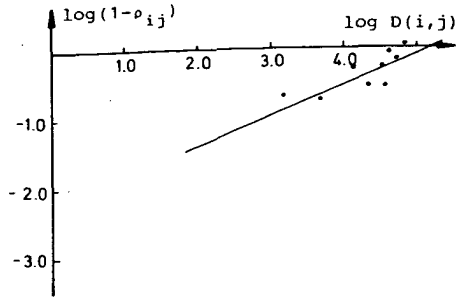


Fig. 7. Correlation - distance relation for region with 6 stations. June and July 1931 -60.

the model shows up. Note, that the relation for the "inadequate" region seems to be more inaccurate than the relation for the "adequate" region, confirming a less specific spatial correlation in this area as a whole - even considering the different distance available in the two regions.

Furthermore, reading off the slope, τ and interception, $\gamma(\beta_t)$ as mentioned estimated as one pooled value - we get

		$\gamma(\beta_t)$	τ
"adeqate" region	summer	-2.60	0.42
	autumn	-3.00	0.48
"inadequate" region"	summer	-2.25	0.43

(19)

If we re-write (12) into the mathematical form

$$\rho_{ij} = 1 - \exp(\gamma(\beta_t)) \cdot D(i,j)^{\tau} \quad (20)$$

we see that estimated γ, τ- parameters determining the correlation functions for the two regions in the summer period agree fairly well. However, the inadequacy of the one region might very well been illustrated by unacceptable γ and τ values (Allerup and Madsen (1983)).

Comparison between correlation functions within the "adequate" region shows a clear ordering of the two functions: For "all" distances (less than 785 km) correlations in autumn is found to be higher than correlation in summer. This is not surprising and it is in agreement with earlier results (Allerup (1981)). From (20) - or rather the form (13) concerning the covariances - we get immediately

$$\tau = \frac{\partial \mathrm{cov}(x_{it},x_{jt})}{\partial D(i,j)} / \mathrm{cov}(x_{it},x_{jt}) \cdot D(i,j) \quad (21)$$

which facilitates an interpretation of τ being, at a fixed distance the relative gain (or loss) of correlation by one spatial unit ∂D. It is further tempting to identify $\gamma(\beta_t)$ as a measure of general level of correlation between points at unit distance, but such interpretations should be seen in relation to ordinary constraints on the correlation function ($\rho \in [-1,1]$, see Allerup and Madsen (1983)).

8. RESULTS OF INTERPOLATIONS

In Fig. 8-10 the results of applying the interpolation procedure against one staion selected as "unknown" in each region are presented. In each region, three base stations, those closest to the unknown station have been used and we have tried to interpolate all 60 "missing" values.

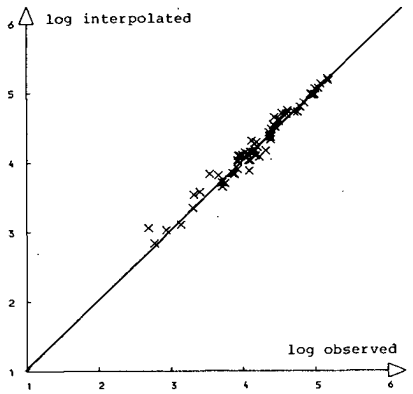


Fig. 8. Observed - interpolated values for region with 7 stations. Oct. and Nov. 1931 - 60.

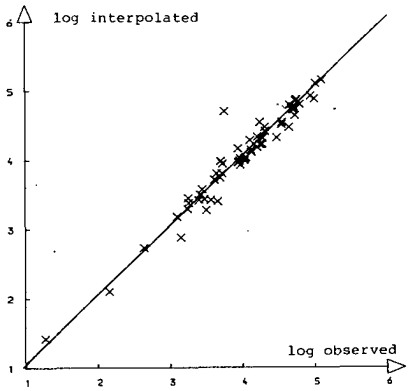


Fig. 9. Observed - interpolated values for region with 7 stations. June and July 1931 - 60.

The detailled steps in the interpolation procedure are as follows. First α_i and β_t parameters (1) and (11) are estimated for the region and α-parameter for the unknown site is estimated by averaging the three base stations' α parameters cf. (18). Adding β_t provides us with the structural part, m_2 (cf. (18)) of interpolated value E_q, cf. (15). Grouping of the observations according to β_t results in the pooled estimate of the variance matrix V_{11} for the base stations. This matrix is

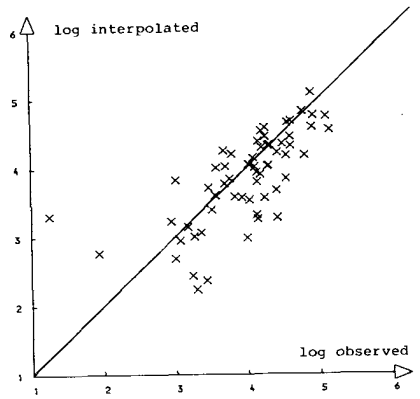


Fig. 10. Observed - interpolated values for region with 6 stations. June and July 1931 - 60.

then inverted, cf. (15). The correlation function for the region is estimated, cf. (12) or (20), and correlations between base stations and the unknown site is obtained from this estimated function by entering known distances. The V_{21} term in (15), however contains covariances instead of correlation coefficients. The transformation is obtained by estimating the variance, σ_q^2 at the unknown site as a simple average of the p base stations' variances. Finally, residuals R_{it} for the base stations provide the factor $\underline{x}^{(p)}-m_1$

The last part of the interpolation formula (15), $V_{21}V_{11}^{-1}(\underline{x}^{(p)}-m_1)$ is the stochastic part. The intention is, that the structural part, m_2 shall produce a first rough guess - a guess that is correct on average, within the mathematical framework of loadings β_t and the α_i's, cf. (1). This structural guess then shall be slightly modified taking into account known spatial correlations and local observed deviations $\underline{x}^{(p)}-m_1$ at a given time from the average.

Observed and interpolated values are distributed around the identity line (Fig. 8-9) for the "adequate" region indicating, for this selection of base stations, a proper, non-systematic deviating re-calculation of the log precipitation values at the unknown site. According to (17) we may further calculate 95% confidence intervals for the interpolated value. In fact, we find for this region $V_q = 0.0136$ valid for autumn and $V_q = 0.0796$ valid for the summer period as estimates of the conditional variance V_q, cf. (16). This lead to $\pm$ 0.22 limits in autumn and $\pm$ 0.55 limits in summer. Expectely, interpolation errors are found to be smaller in autumn than during the summer period. These calculated estimation errors fit well with Fig. 8-9.

For the "inadequate" region, Fig. 10 we first of all observe systematically underestimated log precipitation values. This kind of inadequacy mainly stems from a bad guessing on the structural part by means of the selected base stations. But it may be shown, that instead of correcting this bias, the stochastic part will introduce additional errors in the combined interpolation scheme due to different precipitation regimes in the region. The total accuracy of interpolation, therefore is very poor and this is emphasized by calculation of the variance V_q, cf. (16). V_q is found to be 0.1852, much greater than within the "adequate" region. Hence, the inadequacy of this region has proved to cause effects both on the structural and the stochastic part.

9. CONCLUSION

This investigation of monthly precipitation values from 1931-60 has revealed a simple additive structure in the log-values, stating that the expected logarithmic precipitation may be calculated by means of a site specific parameter and a time specific parameter. Furthermore, the variance structure was analysed resulting in a number of possible mathematical formalizations, the most promising attempt being one with a multiplicative structure. The final statistical model was applied to two regions in Denmark for interpolation purposes. It turned out that an acceptable degree of precision by the interpolations could be obtained in the one region where a homogeneous precipitation regime is prevailing. In the other region with heterogeneous precipitation conditions, we found inaccurate interpolation results.

Besides climatological purposes by interpolation of missing values, the statistical model has been successfully used within hydrological work. By calculating areal precipitation for an important catchment near Copenhagen the model thus has entered as a part of a operationally working method. Applications of the model to interpolations - both monthly and daily precipitation - have proved highly increased precision in comparison with traditional statistical methods.

10. REFERENCES

Allerup, P., 1981: A statistical analysis of monthly precipitation in Denmark. Climatological Papers No.7. Met. Inst.

Allerup, P.,1982: TTSI method for calculating areal precipitation. Proc. Int. Symp. on Hydrol. Bern.

Allerup, P., H. Madsen, J. Riis, 1982: Methods for calculating areal precipitation. Nordic Hydrol, 13.

Allerup, P., and H. Madsen, 1983: Interpolation of monthly precipitation. (To appear from Met. Inst.).

Anderson, T.W., 1958: An introduction to multivariate statistical analysis. Wiley, New York.

Buishand, T.A., 1978: Some remarks on the use of daily rainfall models. J. Hydrol., 36.

Cliff, A. D., and J. K. Ord, 1975: Model building and the analysis of spatial pattern in human geography. J. Roy. Stat. Soc., 3.

Gabriel, K. R., and J. Neumann, 1962: A markov chain model for daily rainfall occurence at Tel Aviv. Quart. J. Roy. Meteor. Soc., 88.

Gandin, L. S., 1970: The planning of meteorological station networks. WMO, tech. note No. 111.

Gates, P., and H. Tong, 1976: On markov chain modelling to some water data. J. Appl. Meteor., 15.

Green, J. R., 1970: A generalized probability model for sequences of wet and dry days. Mon. Wea. Rev., 98.

Iyengar, R., 1982: Stochastic modelling of monthly rainfall. J. Hydrol., 57.

Kruizinga, S., and G. J. Yperlaan, 1978: Spatial interpolation of daily totals of rainfall. J. Hydrol., 36.

Lenton, R. L., and I. R. Iturbe, 1977: A multidimensional model for the synthesis of processes of areal rainfall averages. Water Resources Res. 13,3.

Matalas, N. C., and W. B. Langbein, 1962: Information content of the mean. J. Geophys. Res. 67,9.

Meija, J. M., and I. R. Iturbe, 1974: a) On the synthesis of random field sampling from the spectrum: An application to the generation of hydrologic spatial processes. b) The design of rainfall networks in time and space. c) On the transformation of point rainfall to areal rainfall. Water Resources Res., 10,4.

Nipper, J., and U. Streit, 1982: A comparative study of some stochastic methods and autoprojective models for spatial processes. Envir. and Planning, 14.

Shenton, L. R., and K. O. Bowmann, 1973: Comments on the gamma distribution and uses in rainfall data. Paper read at the third conference on probability and statistics in atmospheric science. Boulder, Colorado.

Shenton, L. R., and P. Skees, 1970: Some statistical aspects of amounts and duration of rainfall. Random counts in physical science, Geo science and business. University Park, London.

A STUDY OF THE VARIABILITY IN SPACE AND TIME OF UPPER VOLTA RAINFALL

Tom A.B. Snijders

Econometric Institute
University of Groningen
Groningen, The Netherlands

1. INTRODUCTION AND SUMMARY

This study is a data-analytic investigation of the variability of Upper Volta rainfall. It constitutes a preparatory step for the construction of a statistical model of the daily rainfall process in Upper Volta, taking account of the possible persistence (over the years) and fluctuations of rainfall in this region. Such a rainfall model will be employed later in a study of variability of food production and agricultural risks in the Sahel-Sudan zone of West-Africa. These studies are part of the AGRISK research programme, outlined in Schweigman et al. (1981).

The yearly rainfall regime is relatively coherent over Upper Volta: according to 'common experience', years tend to be either wet for the country as a whole, or dry for the country as a whole. This is not reflected in high inter-station correlations for yearly rainfall totals, however. These are rather low, as is generally the case in (sub)tropical regions (see Section 2). Other indicators for the yearly rainfall regime can be found, however, which exhibit higher inter-station correlations. A detailed study of inter-station correlations for various indicators is presented in Section 3. This study reveals that inter-station correlations are considerably higher for the number of rainy days than for yearly totals. An interpretation of this result is given, based on the main atmospheric processes influencing rainfall in the Sudan-Sahel zone. It appears that in order to employ the number of rainy days as an indicator, it is important to exclude from the analysis those station-year combinations for which the recording of small rainfall quantities is doubtful. Section 4 is devoted to the construction of a regional index for the yearly rainfall regime of Upper-Volta, based on the number of rainy days and taking into account the irregular pattern of missing data. This index is used in Section 5 in a study of persistence and fluctuations of the Upper Volta rainfall regime. The null hypothesis that the rainfall process is, for different years, independent and identically distributed is clearly rejected by suitable non-parametric tests based on this index. This is due mainly to an apparent change of level around 1970. Tests of serial correlation over the periods 1940-1969 and 1970-1981 separately do not lead to significant results.

2. AIMS OF THE STUDY

In this section some initial reflections are mentioned which were the basis of the study reported in this paper. These reflections are based on the literature about the main atmospheric processes influencing rainfall in the Sudan-Sahel zone of West-Africa (e.g. Cochemé and Franquin (1967), Ojo (1977)) and on the literature about rainfall fluctuations in this region (see Nicholson (1980) and references cited there). The following conclusions were drawn from this literature.

(1) Almost all rain in Upper Volta falls during the rainy season, roughly contained in the months june-september. The main factor influencing the amount of yearly rainfall is the extent of the northward seasonal excursion of the Inter-tropical Convergence Zone. This factor affects a large part of West-Africa, including Upper Volta, for each year separately in roughly the same way. This suggests a considerable spatial coherence of the yearly rainfall regime, so that years will tend to be either wet for Upper Volta as a whole, or dry for the country as a whole.

(2) It is not to be expected that the coherence mentioned in (1) is expressed in high inter-station correlations for yearly rainfall totals. These tend to be low in the tropics and subtropics, as is documented, e.g. in Jackson (1977). This is caused by the fact that most of the rain falls in a few large showers, which are caused by disturbances leading to spatially extremely variable amounts of rainfall. E.g., in the central and northern parts of Upper Volta, the single day with the greatest amount of rainfall in a given year at a given rainfall station often accounts for 10 percent or more of the yearly rainfall total (see Lahaye (1980)), while the 8 days with the greatest amounts of rainfall

mostly account for 40 to 60 percent.
(3) The statistical and substantial significance of the persistence and fluctuations of West-Africa rainfall has been a matter of debate, e.g., in Bunting et al. (1976), Kraus (1977), Katz (1978) and Lamb (1982). Recent literature does suggest that the statistical significance of this persistence cannot be denied. A specific analysis for Upper Volta seems to have not been made yet, and agroclimatological studies for this country (such as, e.g., Franquin (1983)) tend to use the assumption that, for different years, the yearly rainfall process is independent and identically distributed.

These conclusions led to the following aims for this study.
A. A study of indicators for the yearly rainfall regime at a given rainfall station, which are less influenced by the size of occasional heavy showers than yearly totals. It can be expected that such indicators reflect, better than yearly totals, the atmospheric processes (notably the position of the Inter-tropical Convergence Zone) affecting the whole of Upper Volta in roughly the same way; while they reflect less than yearly totals the local atmospheric processes responsible for the exact amount of rainfall in heavy showers. Such indicators are also less affected by gross recording or transcription errors. If these expectations are correct, inter-station correlations for these indicators will tend to be higher than for yearly totals.
B. A study of the persistence and fluctuations of the Upper Volta yearly rainfall regime on the basis of a regional rainfall index, obtained by combining an indicator of the type mentioned in A for several rainfall stations.

3. INTER-STATION CORRELATIONS FOR VARIOUS INDICATORS

Daily rainfall data for the months june to september is used in this paper for 20 Upper Volta rainfall stations. Almost all rainfall stations for which the records start before 1930, and a number of stations for which the records start before 1950, were chosen. The stations were selected to obtain a reasonable geographic distribution. A list of these stations and more information about the data is contained in the appendix. For this section, only the years 1940-1981 are used because the selection procedure to be discussed later in this section left too few stations for the years before 1940. Table 1 gives some important points from the distribution of the 190 inter-station correlation coefficients for the total june-september rainfall recorded at these stations. The median correlation is .27, which is rather low (as expected).

Table 1. Extremes, p(.25) = first quartile, p(.5) = median and p(.75) = third quartile of inter-station correlation coefficients for the june-september rainfall total.

min	p(.25)	p(.5)	p(.75)	max
-.18	.16	.27	.39	.70

In order to limit the effect of incidental large showers, sums of truncated daily rainfall values were considered as indicators. Such an indicator is given by

$$s(t) = \Sigma_d \min\{r_d, t\} \qquad (1)$$

where r_d is rainfall for day d, the summation extends from june 1st to september 30th, and t is the truncation level. Units are mm., and rainfall is recorded as an integer multiple of .1 mm. Therefore s(.1) is just the number of rainy days, and (1) defines a continuous family of indicators, linking the rainfall total = $s(\infty)$ to the number of rainy days. The influence of the exact amount of positive daily rainfall values on s(t) is greater for greater values of the truncation parameter t. It was expected at the outset of this study that the indicator s(t) with the highest median inter-station correlation coefficient would have t somewhere between 5 and 50 mm.

A scrutiny of inter-station correlations for the indicators s(t) and of the basic rainfall data revealed that unconditional use of s(t) for t less than, say, 10 or 20 mm., leads to meaningless results because of the varying data quality. The proportion among rainy days of days with a small amount of rainfall, say, rainfall less than 5 mm., is very variable over station-year combinations. This suggests that small rainfall amounts have not always been well recorded: they may occasionally have been omitted, or small daily totals for several days may occasionally have been grouped into one positive total, with nil values for the other days. Another problem is that in quite a number of station-year combinations rainfall is mostly or always registrated in multiples of .5 or 1.0 mm., with the effect that a higher threshold for the registration of a rainy day is employed than when rainfall values .1, .2., etc. can be registrated. As rainfall values of .1 to a few mm. account for a considerable proportion of the rainy days (see Figure 1 below), these problems of observer reliability and instrument precision have the effect that s(t) values for t less than 10 or 20 mm. are not comparable over all station-year combinations. It appears from the data that these problems tend to affect most stations more in the beginning of the period 1940-1981 than later, thereby depressing s(t) for small t for many stations simultaneously in these early years. This results in artificially high correlations for s(t) for small values of t between many stations. Two approaches were considered in order to try to overcome this difficulty.
(i) Try to use the daily rainfall values for selecting those station-year combinations where small rainfall values seem to have been registrated to a sufficient extent, and use only those data in the analysis. This approach has the disadvantage that such a selection is never perfect.
(ii) Apply a threshold of a few mm., e.g., in the range from 3 to 10 mm., and set rainfall values below this threshold equal to 0. This approach has the disadvantage that the information contained in positive rainfall values below this threshold is lost, and that the remaining data for some station-year combinations still may suf-

fer from imperfect observer reliability.

The first approach turned out to yield the highest inter-station correlations. The results of this approach are discussed first, and a discussion of results of the second approach is given later in this section. The selection of station-year combinations for the first approach is based on the fraction of rainfall values recorded as multiples of .5 mm. among days with positive rainfall not exceeding 3.0 mm.,

$$p = \frac{\#\{d \mid r_d = k/2 \text{ for } k=1,2,\ldots,6\}}{\#\{d \mid 0 < r_d \leq 3.0\}},$$

#{d | P} denoting the number of days from june 1st to september 30th for which property P holds, and r_d denoting rainfall for day d in mm.; on the fraction of days with not more than 3.0 mm. rainfall among rainy days,

$$q = \frac{\#\{d \mid 0 < r_d \leq 3.0\}}{\#\{d \mid r_d > 0\}};$$

and on the number of days where positive rainfall less than 1.0 mm. was recorded,

$$n = \#\{d \mid .1 \leq r_d \leq .9\}.$$

Relatively high values for p and zero values for n suggest that for the station-year combination considered, a higher threshold for the registration of rain was employed than for station-year combinations with relatively low values for p and positive n. Relatively low values of q suggest that single days with small rainfall values were not always faithfully recorded. Those station-year combinations were selected for which

$$q \geq p-.2, \quad q \geq .1 \quad \text{and} \quad n \geq 1. \qquad (2)$$

This is illustrated by Figure 1.

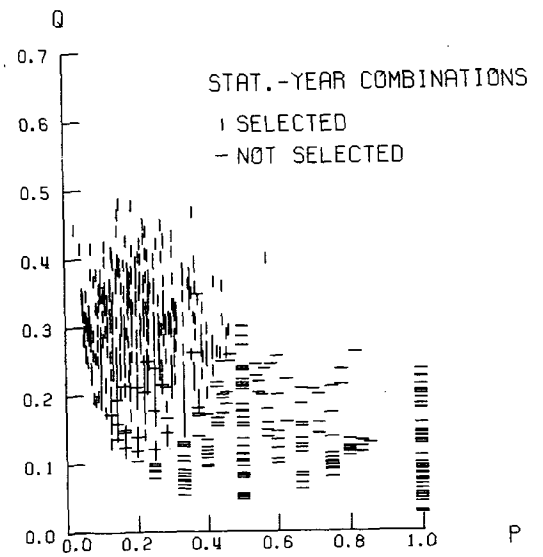


Figure 1. Plot of (p,q) coordinates of station-year combinations selected and not selected.

The results of this approach are not strongly sensitive to the choice of 3.0 mm. as an upper bound in the definition of p and q, nor to the precise definition of the selection criterion (2). Data-based selection procedures, like the one employed here, sometimes introduce artefactual effects; this seems, however, not to happen in this case. There seems to be no reason why this selection procedure should artificially tend to increase inter-station correlations, as it does not favor stations with high s(t) values to be selected together for the same year, with the exclusion of stations with low s(t) values for that year (or vice versa). The selection procedure will affect inter-station correlations mainly by the increased comparability and reliability of the remaining data, and by the decreased number of years on which most correlations are based.

Table 2 contains some important points from the distribution of the inter-station correlations for s(t) for various truncation levels t, using only the station-year combinations satisfying criterion (2), and using only correlation coefficients based on at least 8 years with data available for both stations. The number of such correlation coefficients is 169 (out of 190).

Table 2. Extremes and quartiles (indicated as in Table 1) of inter-station correlation coefficients for indicator s(t) for various truncation levels t, using selection procedure (2).

t	min	p(.25)	p(.5)	p(.75)	max
∞	-.46	.11	.30	.44	.91
50.0	-.39	.13	.32	.46	.90
20.0	-.18	.23	.41	.54	.86
10.0	-.17	.31	.47	.58	.88
5.0	-.18	.37	.48	.61	.88
.1	.01	.30	.45	.57	.88

The inter-station correlations for the rainfall total, or s(∞), with this selection procedure have a slightly higher median (.30) than without the selection procedure (.27). Their distribution is much more dispersed, which can be attributed to the smaller number of data on which almost all of these data are based. It is seen from Table 2 that the three quartiles of the distribution of the inter-station correlation coefficients for s(t) are maximal for truncation levels t about equal to 5 mm. The indicators s(t) for $.1 \leq t \leq 10.0$ do not lead to very different distributions for the inter-station correlation coefficients. The number of rainy days, or s(.1), was selected for further use in this study because of its conceptual simplicity. This indicator leads to considerably higher inter-station correlations than the rainfall total. It can be concluded that, although the median correlation coefficient of .45 shows that a large proportion of the number of rainy days cannot be accounted for by a 'factor' common to all stations, yet the number of rainy days reflects considerably better than the rainfall total the conditions which per year are common to all stations considered. Regarding the position of the Inter-Tropical Convergence Zone (ITCZ) as a main ele-

ment of these 'conditions common to all stations', this conclusion agrees quite well with the known atmospheric processes. The position of the ITCZ provides the potentiality for rain, but the factors determining the amount of rainfall are of a very local nature. This is confirmed in Figure 2, giving the inter-station correlation for the number of rainy days (where again the selection according to criterion (2) is used) as a function of distance. Distance is expressed in degrees; at these latitudes, one degree latitude and one degree longitude both are about equal to 115 km.

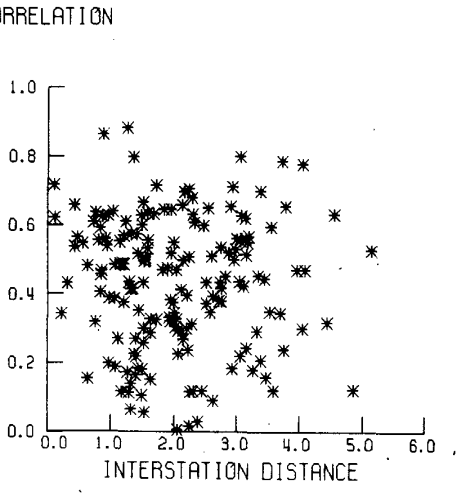


Figure 2. Plot of inter-station correlation coefficients for the number of rainy days against distance. The selection according to criterion (2) is used, and only correlations based on at least 8 years are shown.

It is seen that the correlations exhibit a very slight negative correlation with distance.

Calculating inter-station correlations for indicators s(t) for all available data and using a threshold of 5 mm. as proposed in approach (ii) above leads to a median inter-station correlation coefficient of .23 for the rainfall total, and .27 for the number of rainy days. The difference is not considerable. For truncation levels $.1 < t < \infty$ the situation is intermediate. When a threshold of 5 mm. is applied simultaneously with a selection of station-year combinations according to (2), the median inter-station correlation coefficient for the number of rainy days is .40. This figure is considerably higher than the median .27 obtained for the number of rainy days with a threshold of 5 mm. and without selection, but considerably lower than the medians .45 and .48 obtained for the number of rainy days and s(5), respectively, in approach (i) and reported in Table 2. This leads to the conclusion that approach (ii) is affected both by the imperfect reliability of the registration of the number of rainy days with rainfall exceeding the threshold in some of the station-year combinations, and by the loss of information contained in daily rainfall amounts not exceeding the threshold.

Some other indicators were tried besides s(t), namely, summed power tranforms

$$\Sigma_d (r_d)^p$$

for $0 \leq p \leq 1$ (which also is a family linking the number of rainy days, obtained for p=0, to the rainfall total, obtained for p=1), and simulated evapotranspiration measures based on simple water-balance models. These indicators did not lead to higher inter-station correlation coefficients than the number of rainy days. The results for the evapotranspiration measures will be reported elsewhere.

Arguments for the relevance of the number of rainy days as an indicator for the yearly rainfall regime at a given station are not restricted to the higher inter-station correlations obtained. For agricultural purposes, the number of rainy days seems a better indicator for the adequacy of rainfall to the plants than total rainfall, because total rainfall is dominated by a small number of heavy showers which disappear from the fields for a large part in the form of run-off.

4. CONSTRUCTION OF A REGIONAL RAINFALL INDEX

The rainfall index for the yearly rainfall regime in Upper Volta will be based on the number of rainy days in the months june to september for the 20 rainfall stations mentioned in the appendix. Station-year combinations are selected according to criterion (2) described in the previous section. The irregular pattern of missing and excluded data poses a problem for the combination of the values for individual stations into a single regional index. One must take into account the interaction between this pattern and the systematic differences in rainfall levels between stations and between years. Let s_{ij} be the number of rainy days recorded at station i during the months june to september of year j, for station-year combination (i,j) with available data and satisfying selection criterion (2). Other (i,j) combinations are considered as missing data. The sample mean $s_{.j}$ of the available s_{ij} for the fixed year j is an unsatisfactory index, as it will produce too high values for years where few stations with low average s_{ij} levels are included in the data, and vice versa. (It also has the well-known disadvantage of giving most influence to the stations having the highest variance for s_{ij}.) Kraus (1977) proposes the index

$$m_j^{-1} \Sigma_i (s_{ij} - s_{i.})/v_i \qquad (3)$$

where $s_{i.}$ and v_i are the sample mean and sample standard deviation, respectively, of the s_{ij} available for station i, while m_j is the number of stations available for year j. This index is used, e.g., by Lamb (1982). It is not a satisfactory combination of the s_{ij} into a series of

yearly indices, because for some of the stations i the set of years for which data are available, and on which the variables $s_{i.}$ and v_i used for the standardization are based, may be a subset of the complete set of years analyzed which is not sufficiently representative, as can be illustrated by the following example. Let a certain station i have data s_{ij} available only for a set J(i) of relatively dry years. As $(s_{ij}-s_{i.})/v_i$ has mean zero, it is positive for some years in J(i). For these years it provides a positive contribution to index (3), which is not balanced by positive contributions for years $j \notin J(i)$.

A different regional index is proposed here, based on an additive approximation to the data,

$$s_{ij} = \zeta + \eta_i + \xi_j + e_{ij} \qquad (4)$$

where the term e_{ij} is a residual which is hoped to be small. Approximation (4) is similar to an analysis of variance model for the s_{ij}, but the assumption is not formally made that such a model is satisfactory in a statistical sense. As the index for year j, we propose ξ_j produced by a least squares fit for approximation (4), i.e., ξ_j for which

$$\Sigma_{i,j}(s_{ij}-\zeta-\eta_i-\xi_j)^2 \qquad (5)$$

is minimal. This ξ_j is related to the sample mean $s_{.j}$ of the available data for year j by the formula

$$\xi_j = s_{.j} - \zeta - m_j^{-1}\Sigma_{i\in I(j)}\eta_i$$

where ζ and η_i are also defined as the values minimizing (5), and I(j) is the set of stations i for which data for year j are being used. This equation does not explicitly define ξ_j, but provides some insight into the way in which this index keeps account of the pattern of missing data. It was stated above that it is not assumed that (4) is a satisfactory model in a statistical sense. Such an assumption is not necessary as ξ_j is a descriptive statistic. The index ξ_j will, however, be more satisfactory according as the additive approximation (4) provides a better description of the data, with preferably more or less homoscedastic residuals e_{ij}. (If, e.g., rainfall totals instead of the number of rainy days were used as s_{ij}, then it might be wise to apply the construction of the index ξ_j not to the original data but to some concave transform as, e.g., $(s_{ij})^{\frac{1}{2}}$, in order to approximately stabilize the sample varainces of the residuals e_{ij}.)

The index ξ_j was computed for the period 1940-1981. The number of included stations is 4 to 8 for the period 1940-1950 and 9 to 19 for the period 1951-1981. Years prior to 1940 are not used because the number of data which remains after application of selection procedure (2) is too small. The results are given in Table 3.

Table 3. Index values ξ_j with the root mean square of the residuals (rms_j) and the number of included stations (m_j) for the years 1940 to 1981.

j	ξ_j	rms_j	m_j	j	ξ_j	rms_j	m_j
1940	-5.8	5.8	4	1961	1.8	5.5	12
1941	-1.1	4.5	5	1962	3.3	4.7	15
1942	-4.6	2.9	6	1963	2.12	3.7	14
1943	3.4	1.5	5	1964	8.5	4.8	17
1944	-4.9	4.4	8	1965	2.8	2.8	14
1945	1.25	3.9	8	1966	-1.7	3.8	13
1946	-2.95	4.3	7	1967	4.5	6.0	19
1947	-7.2	7.7	7	1968	2.5	6.0	14
1948	-2.3	3.0	7	1969	2.7	4.1	17
1949	-2.98	4.1	8	1970	8.12	3.2	18
1950	2.5	4.3	6	1971	5.6	4.1	17
1951	2.11	5.7	10	1972	-5.6	4.0	19
1952	6.9	5.5	9	1973	-6.9	4.3	17
1953	1.30	4.8	11	1974	9.5	3.6	18
1954	8.10	5.2	11	1975	-1.5	4.8	18
1955	2.0	4.7	12	1976	-5.1	5.0	18
1956	8.4	3.8	9	1977	-7.7	3.7	11
1957	6.1	5.7	15	1978	-5.0	4.9	10
1958	-3.4	4.7	9	1979	1.23	5.8	16
1959	4.9	7.5	13	1980	-6.57	4.2	15
1960	5.2	6.7	11	1981	-6.55	4.1	15

The least squares fit for (4) yields automatically the station indices η_i, which are presented in Table 4; and the value for ζ, which is equal to 52.0.

Table 4. Stations with their coordinates, index values η_i, root mean square of the residuals (rms_i) and the number of included years (n_i).

Station	N.lat.	W.long.	η_i	rms_j	n_i
Gaoua	10°20'	3°11'	7.3	5.3	39
Banfora	10°37'	4°46'	6.2	8.2	7
Diebougou	10°58'	3°15'	3.1	3.8	16
Leo	11°06'	2°06'	.5	5.3	16
Bobo-Dioulasso	11°10'	4°18'	12.2	5.5	39
Boromo	11°44'	2°55'	8.3	3.8	37
Tenkodogo	11°45'	0°23'	.1	4.7	25
Fada N'Gourma	12°04'	0°21'	6.0	6.0	37
Koupela	12°11'	0°21'	-4.2	4.9	23
Ouagadougou V.	12°22'	1°31'	2.8	4.5	39
Koudougou	12°16'	2°22'	-3.8	5.4	19
Dedougou	12°28'	3°28'	-2.5	3.1	27
Kantchari	12°28'	1°31'	-2.4	3.6	21
Yako	12°58'	1°16'	-5.4	5.2	18
Bogande	12°59'	0°08'	-12.3	1.9	11
Tougan	13°05'	3°04'	-3.8	4.0	22
Kaya	13°06'	1°05'	-9.2	5.3	24
Bam (Tourcoing)	13°20'	1°30'	-8.0	4.4	28
Ouahigouya	13°35	2°26'	-3.8	5.3	26
Dori	14°02'	0°02'	-9.4	4.4	34

It is seen from these tables that the residual root mean square is not too variable over years or stations nor strongly correlated with the yearly index ξ_j or station index η_i. It can therefore be concluded that the additive approximation (4) is adequate for the construction

of the regional index ξ_j for the yearly rainfall regime.

5. FLUCTUATIONS OF THE YEARLY RAINFALL REGIME

In this section the persistence and fluctuations of the yearly rainfall regime in Upper Volta are studied on the basis of the index ξ_j constructed in the preceding section. The autocorrelation at lag 1 for this series equals .40, and for lags 2 to 5 the autocorrelation remains well above .3. In order to apply statistical tests, these values are not very useful because well-founded distributional assumptions about the ξ_j cannot be specified and the sample size is rather small for the invocation of asymptotic approximations. Therefore some non-parametric tests are applied to the ξ_j series.

The climatological null hypothesis which will be tested in this section is that the rainfall regime for Upper Volta for the months june-september is, for different years, independent and identically distributed. In view of the use of selection procedure (2), it must be assumed that the instrument and observer characteristics at the stations considered (which characteristics, together with some characteristics of the rainfall regime, influence the outcome of selection procedure (2)) do not introduce, either because of inhomogeneity of the data passing the selection procedure, or because of the amount and pattern of missing data, serial dependence or level fluctuations into the series of indices ξ_j. This assumption seems not unreasonable when only years from 1951 or, better still, from 1961 are considered, for which the number of included stations (see Table 3) is fairly high. It is evident that, as the tests described below are based on ξ_j, attention is focussed on the detection of possible fluctuations or persistence in the number of rainy days. The null hypothesis and assumption formulated here jointly imply that the probability distribution of the vector $(\xi_j)_{j_1 \leq j \leq j_2}$ for, say, $j_1 = 1951$ and $j_2 = 1982$, is invariant under permutations of the j_2-j_1+1 indices ξ_j. (Because of the definition of the ξ_j as the least squares solution of (4) the indices ξ_j are not necessarily statistically independent.) The tests are based on the vector of ranks $R = (R_j)_{j_1 \leq j \leq j_2}$, R_j being the rank (in increasing order) of ξ_j among $\xi_{j_1}, \xi_{j_1+1}, \ldots, \xi_{j_2}$. The permutation invariance for the ξ_j implies that every permutation of the numbers $1,2,\ldots,T$ is equally probable as the outcome of R, where T is defined as $T = j_2-j_1+1$. Two tests are applied: Pettitt's change-point test, directed mainly at the detection of a level change, and the rank version of von Neumann's ratio, directed mainly at the detection of autocorrelation. These tests are described in Pettitt (1979) and Bartels (1982).

The test statistic for Pettitt's change-point test is

$$K = \max_{j_1 \leq t < j_2} |W_t|$$

where W_t is the Wilcoxon-Mann-Whitney test statistic for the two-sample problem for the samples $\xi_{j_1},\ldots,\xi_t$ and $\xi_{t+1},\ldots,\xi_{j_2}$, expressed as

$$W_t = \Sigma_{j=j_1}^{t} (2R_j - T - 1).$$

This test is designed for the detection of a level change after an unknown 'time' t. The asymptotic Kolmogorov-Smirnov null distribution for K is applicable only for very large sample sizes. Therefore the significance level was estimated by a Monte Carlo method, using a variance reduction technique by means of antithetic variates described in Snijders (1983). For $j_1=1961$, $j_2=1981$, the significance level attained was less than .01. It can be concluded that the null hypothesis of independence and identical distributions for the rainfall regime must be rejected. Table 3 seems to indicate a level change after 1971. In order to investigate this further, Pettitt's change-point test was also applied to some other periods j_1 to j_2, as indicated in the following table.

Table 5. Outcome k of test statistic K and associated significance level p for Pettitt's change-point test applied to the regional rainfall index ξ_j for $j_1 \leq j \leq j_2$.

j_1	j_2	k	p
1951	1981	200	<.01
1951	1971	32	.80
1961	1981	98	<.01
1961	1971	14	.65
1972	1981	8	.96

The results of these tests are consistent with a level change between 1971 and 1972, and constant levels both before 1971 and after 1972.

The rank version of van Neumann's ratio is given by

$$H = \Sigma_{j=j_1}^{j_2-1} (R_{j+1}-R_j)^2 / \Sigma_{j=j_1}^{j_2} (R_j-R_.)^2,$$

$R_.$ being the average rank (equal to $(T+1)/2$). This statistic is a function of the rank serial correlation coefficient. Small values of H point towards positive autocorrelation in the ξ_j series. Critical values can be found in Bartels (1982). On the basis of the significant result of Pettitt's change-point test for the period 1960-1981, it was decided to apply this test to the period 1951-1971 and 1972-1981 separately. The results were not significant. For the sake of completeness, the rank von Neumann's ratio test was also applied to some other periods j_1 to j_2. The attained significance levels are given in the following table.

Table 6. Outcomes h of test statistic H and associated significance levels p for the rank version of von Neumann's ratio applied to the regional rainfall index ξ_j for $j_1 \leqq j \leqq j_2$.

j_1	j_2	h	p
1951	1981	1.14	<.005
1951	1971	2.32	>.10
1961	1981	.95	<.005
1961	1971	2.27	>.10
1972	1981	2.22	>.10

The results of these tests agree with those of Pettitt's change-point test reported in Table 5, and are consistent with a level change between 1971 and 1972, while the null hypothesis of independence and identical distributions for the rainfall regime for the periods 1961 (or 1951) to 1971, and 1972 to 1981 separately, cannot be rejected on the basis of these tests. It can be reported here that when these tests are applied to an index computed as in Section 4 on the basis of rainfall totals (after a square root transformation), the same conclusion is obtained.

The results of these tests and the series of index values ξ_j reported in Table 3 show that the drought (expressed by the number of rainy days) which has affected Upper Volta from 1972 cannot be explained away as a chance effect, and that apart from the recovery years 1974 and 1979, the drought has been continuing at least until 1981. Based on data from 20 rainfall stations throughout West-Africa, Lamb (1982) finds a similar result, while Lamb (1983) reports that the drought continued in 1982.

APPENDIX. RAINFALL DATA USED

Daily rainfall data for Upper Volta rainfall stations, from the origin of the stations to 1973, were supplied by the Office de la Recherche Scientifique et Technique Outre-Mer (O.R.S.T.O.M.), Paris. Daily rainfall data for 1974-1981 were kindly supplied by the Direction de la Météorologie de la Haute Volta, Ouagadougou. The list of stations used in the present paper is given in Table 4. Station-year combinations for which data over the period june-september were incomplete, or for which the O.R.S.T.O.M. commentary document expressed any doubt concerning the reliability, were not used.

ACKNOWLEDGEMENTS

I am grateful to the Hydrological Service of O.R.S.T.O.M. and the Meteorological Service of Upper Volta for providing me with the Upper Volta rainfall data. I also like to thank Caspar Schweigman for his comments, which led to considerable improvements in the presentation of this paper.

REFERENCES

Bartels, Robert (1982), The rank version of von Neumann's ratio test for randomness, J. Am. Stat. Ass. 77, 40-46.

Bunting, A.H., M.D. Dennett, J. Elston and J.R. Milford (1976), Rainfall trends in the West African Sahel, Quart. J. Roy. Met. Soc. 102, 59-64.

Cochemé, J. and P. Franquin (1967), An agroclimatology survey of a semi-arid area in Africa south of the Sahara, Technical Note no. 86, World Meteorological Organisation, Geneva.

Franquin, P. (1983), Agroclimatological studies in sorghum and millet growing regions in Africa-climatic hazards. In: Agrometeorology of sorghum and millet in the semi-arid tropics, ICRISAT, Hyderabad.

Jackson, I.J. (1977), Climate, water and agriculture in the tropics, Longman, London.

Katz, E.B. (1977), Persistence of subtropical african droughts, Mon. Wea. Rev. 106, 1017-1020.

Kraus, E.B. (1977), Subtroptical droughts and cross-equatorial energy transports, Mon. Wea. Rev. 105, 1009-1018.

Lahaye, J.P. (1980), Etude des pluies journalières de fréquence rare en Haute Volta, C.I.E.H., Ouagadougou.

Lamb, Peter J. (1982), Persistence of subsaharan drougth, Nature 299, 46-48.

Lamb, Peter J. (1983), Subsaharan rainfall update for 1982: continued drought. Submitted for publication.

Nicholson, Sharon E., The nature of rainfall fluctuations in subtropical West Africa, Mon. Wea. Rev. 108, 473-487.

Ojo, Oyediran (1977), The climates of West Africa, Heinemann, London.

Pettitt, A.N. (1979), A non-parametric approach to the change-point problem, Applied Statistics 28, 126-135.

Schweigman, C., T.A.B. Snijders and F. Hagenzieker (1981), Risks in agriculture in developing countries caused by rainfall variability: Introduction to risk computations. Office for International Cooperation, University of Groningen.

Snijders, Tom A.B. (1983), Antithetic variates for Monte Carlo estimation of probabilities. Submitted for publication.

THE STANDARDIZED BIVARIATE LOGNORMAL PROBABILITY DISTRIBUTION SIMPLIFIES THE STORM-FLOOD RELATIONSHIP

Lourens A. V. Hiemstra

Department of Civil Engineering
University of Stellenbosch
Stellenbosch 7600, South Africa.

1. INTRODUCTION

For most practical engineering applications a flood is sufficiently defined if the complete flood hydrograph and its exceedance probability is known. Since the cross-correlation coefficient between flood peaks and their associated volumes is not unity, a family of floods exists for each exceedance probability. In an analysis of 1 172 individual floods on 43 selected catchments in South Africa by Hiemstra and Francis, 1981, it was found that the standardized bivariate lognormal probability distribution describes the flood peaks and their associated volumes quite well. Furthermore it seems as if the cross-correlation coefficient, the coefficient of variation of flood peaks and the coefficient of variation of flood volumes can be regionalized. For most of the catchments studied a cross-correlation coefficient of 0,85, a coefficient of variation for flood peaks of 1,30 and a coefficient of variation for flood volumes of 1,57 seem to be reasonably applicable. This means that if the geometric mean flood peak and its associated volume is known for a particular problem catchment, then the family of flood hydrographs for any desired exceedance probability can be found. For fully gauged catchments these values are easy to calculate but for ungauged catchments the problem is not yet satisfactorily solved. One possible way to find the geometric mean flood peak and its associated volume is to make use of records on rainstorms and to try and derive the desired flood from rainfall.

2. FLOODS FROM RAINSTORMS

For the calculation of the resulting flood a rainstorm can be defined in terms of it's areal distribution, it's direction and speed of travel, it's duration, it's depth of total rainfall, it's time distribution of rainfall intensities and it's exceedance probability. This gives about 12 necessary parameters. Furthermore, the response of a catchment to a given rainstorm depends on state variables of the catchment like wetness, state of vegetative cover, depth of water in the river, etc. The problem is therefore unwieldy and simplifying assumptions are necessary. It can be assumed, for design purposes, that the catchment is in an "average" state when the rainstorm occurs. It can also be assumed that the areal distribution throughout the storm is uniform. Likewise for the time distribution of rainfall intensiteis. The only remaining variables are therefore the duration of the storm, the total depth of rainfall and the exceedance probability of the storm. For this particular application, the exceedance probability can also be fixed at a constant. Only 2 variables remain for analysis, namely the duration of the rainstorm and the total depth of rainfall. The available data-base on rainfall, in most countries, can support such an analysis. For South Africa such an analysis gave satisfactory results for the calculation of design floods on ungauged catchments.

3. EXAMPLE

For the Vaal River at Standerton, a catchment with an area of 8 254 km^2, a few members of the floods with an exceedance probability of 0,02 were desired. A streamflow record stretching over 54 years was available and 5 members of the desired family of floods were extracted. Their relevant parameters appear in Table 1.

Table 1: Parameters of five members of the family of floods with exceedance probability of 0,02 from observed streamflow record

Volume ($m^3 * 10^6$)	Peak ($m^3 s^{-1}$)	Time to peak (hours)
308,9	2 231	22
436,0	2 129	32
583,9	1 878	59
692,4	1 507	95
737,1	1 164	149

By treating the catchment as ungauged and using rainfall to predict the corresponding members of the 0,02 exceedance probability family of floods the relevant parameters shown in Table 2 were obtained:

Table 2: Parameters of corresponding members of the family of floods predicted from rainfall

Volume ($m^3 * 10^6$)	Peak ($m^3 s^{-1}$)	Time to peak (hours)
279,1	2 467	17
403,3	2 302	27
557,3	1 929	52
669,2	1 465	95
723,4	1 055	165

The %-error in prediction of the relevant parameters appear in Table 3:

Table 3: %-error in prediction of relevant parameters from rainfall

Volume	Peak	Time to peak
- 10	+ 11	- 23
- 8	+ 8	- 16
- 4	+ 3	- 12
- 3	- 3	0
- 2	- 9	+ 11

The error in measurement of streamflow is usually about 10%, hence these results are satisfactory.

4. CONCLUSION

From the results shown in the 3 tables it seems as if the simplifying assumptions on the causative rainstorms made possible by the bivariate description of floods, may be useful in the prediction of design floods from rainfall on ungauged catchments.

5. REFERENCES

Hiemstra, L.A.V. and D.M. Francis, 1981: Run hydrographs for prediction of flood hydrographs. J. Hydr. Div., ASCE, V. 107, 759 - 775.

FIVE-DAY PERIODS (PENTADS) MAXIMUM RAINFALL: PROBABILITY TABLE FOR PELOTAS, RS, BRAZIL

João Baptista da Silva

Instituto de Física e Matemática
Universidade Federal de Pelotas
Pelotas, Rs, Brazil

1. INTRODUCTION

Rainfall, and specially maximum rainfall, has generaly been object of study as a consequence of its importance as a factor of human conditioning.

The knowledge of the characteristics of the maximum rainfall has great technical interest because of its frequent application in projects in hydraulics, i.e., projects of construction of dams and deviation of watercourses, where the accurrence of flood and its frequency is very important.

A usual approach is to determine the amount and the distribution of flood based on the rainfalls of great intensity from which it would originate.

Heavy rainfalls in short periods of time are of particular interest in hydraulic engineering. They contribute to the floods and originate erosion of the soil, damage to the crops and ditches of irrigation, rupture of dams, etc.

The purpose of the present work is the construction of a table of probabilities of maximum rainfall in pentads (periods of five days) in the city of Pelotas, Brazil. The probabilities are estimated using data corresponding to the period 1900/51 extracted from the records of the extinct weather station of the Escola de Agronomia Eliseu Maciel, located in downtown Pelotas.

The pentad is a convenient basic unit of observation, because it is a period suficiently short for agricultural purposes and allows comparison (when two successive pentads are pooled) with the decades (periods of ten days), very usual in works in Meteorology and Agricultural Ecology.

The method adopted in the construction of the table is essentially the same one used by Amaral & Baptista da Silva (1970) in the construction of a table of probabilities of rainfall and Amaral et alli (1975) in the construction of a table of probabilities of minimum temperature in Pelotas, Brazil. The method is based in a preliminar transformation of the data with the purpose of attaining homogeneity of variance and, presumably, the normality of the distribution, since these propieties usually do not hold for climatic elements in short periods.

The table was tested by comparing the estimated probabilities with the relative frequencies computed using data from the period 1951/78 collect at the agroclimatologic station of the Universidade Federal de Pelotas (UFPEL), 15 km from downtown Pelotas.

It should be pointed out that the proposed table is not intendend for forecast purpose. The table indicates the probabilities of different amounts of maximum rainfall in periods of five days in the city of Pelotas and in the area of the campus of the UFPEL and its neighborhood. It represents a kind of weather forecast for certain periods of an average year, but not for a particular year. Within this scope, it is expected that the table will be useful for those who wish to pay more attention in the planning of operations subject to the inclemency of the weather, taking advantage of the knowledge of the maximum rainfall.

2. CONSTRUCTION OF THE TABLE

It is well known that the distributions of climatic elements observed in short periods of time (i.e., pentads, weeks) generally do not satisfy the assumptions of homogeneity of variance, normality and additivity required by many statistical methods.

So, a preliminar step in the construction of the table is a transformation of the particulated form

* Supported by the agreement EMBRAPA-UFPEL

$$z = k\, x^{1-a} \qquad (1)$$

where x is the maximum rainfall in each pentad of each one of the 52 years (1900-51), a is the slop of the linear regression of the logarithm of the standard deviation with respect to the logarithm of the mean of the x's, and k is an arbitrary constant.

Using the mean $\bar{z}$ and the standard derivation s_z of the transformed values for each pentad, the values z_c corresponding to the levels of the conditional probability P_c are calculated by the formula

$$z_c = \bar{z} + s_z\, t \qquad (2)$$

where $t = t(P_c)$ is the lower point corresponding to the propability P_c of the unilateral t distribution for $N-1$ degrees of freedom (N is the number of years with rainfall of the 52 years of observations in the pentad).

It is well recognized the difficulty raised when periods without rainfall are considered in the distribution of the frequency of rainfall in short periods. The difficulty originates from the descontinuity of the contrast "rain on not rain" in oposition to the continuons variation of the amount of rainfall when it rains.

To circumvent this difficulty, the probability of rainfall (P_1) and the conditional probability (P_c) of the maximum rainfall being less than x milimeters given that it has rained in the pentad, are separately computed for each pentad, then, the probability that the maximum rainfall is less than x milimeters in the pentad is obtained through the formula

$$P = P_o + P_1\, P_c \qquad (3)$$

when $P_o = 1 - P_1 = n/52$ is the probability of not rain ($n = 52-N$ is the number of years without rain in the pentad in the 52 years of observation).

The regression equation of the logarithm of the standard deviation with respect to the logarithm of the mean of the maximum rainfall in each one of the 73 pentads of the year computed for the data of the extinct meteorological station was

$$Y = 0.3447 + 0.81005X,$$

where $X = \log x$ and $Y = \log s_x$; hence the transformation (1) turned out

$$z = 10x^{0.18995}$$

The logarithm of the standard deviation (Y) as a function of the logarithm of the mean (X) is represented in Figura 1.

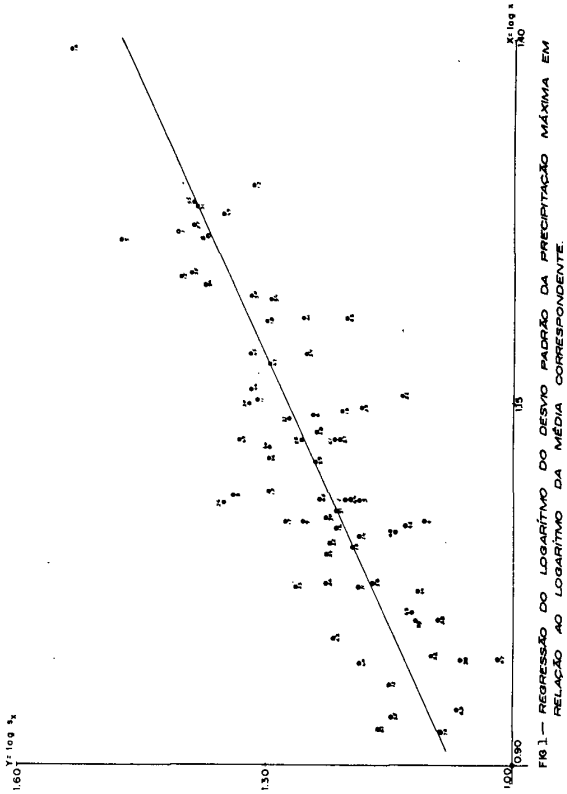


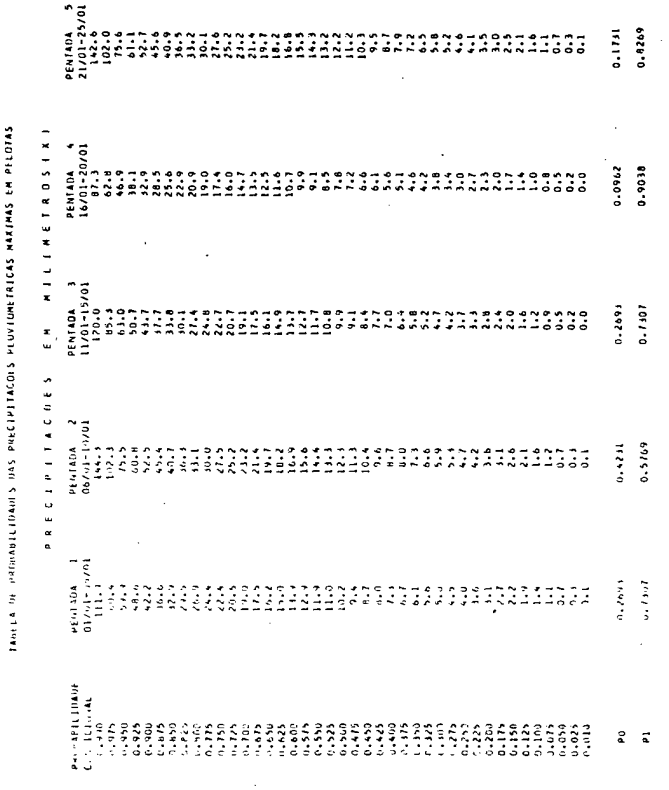
FIG 1 — REGRESSÃO DO LOGARÍTMO DO DESVIO PADRÃO DA PRECIPITAÇÃO MÁXIMA EM RELAÇÃO AO LOGARÍTMO DA MÉDIA CORRESPONDENTE.

For the first pentad of January the number of years with rainfall is $N_1 = 38$, and the mean and standard deviation of the transformed values are $\bar{z}_1 = 15.5579$ and $s_{z1} = 3.6791$. Since $t = -2.427$ for $P_c = 0.010$, expression (2) gives $z_c = 6.6287$, which corresponds, approximately, to x=0.1 mm.

Table I presents the values of the maximum rainfalls corresponding to the conditional probabilities P_c from 0,990 (99%) to 0,010 (1%) for the pentads 1,2,3,4 and 5 (January 1-5, 6-10, 11-15, 16-20 and 21-26, respectively). The complete table for the 73 pentads of the year is presented in Baptista da Silva (1979).

Bartlett's test (Snedecor, 1957) was applied to test the homogeneity of the variance of the 73 pentads for the original and transformed data. The value χ^2_1 obtained for the original data was 414.10 which is significant at the level of probability of 1% ($\chi^2_{0.01;72} = 102.8$), indicating that the variance of x is not homogeneous. For the transformed data the value of χ^2 was 41.82, which is not significant ($\chi^2_{0.05;72} = 92.79$) indicating that the transformation achieved the homogeneity of variance.

TABELA I



Tukey's test of additivity (Snedecor, 1957) was also applied to the original and transformed data. The values of the F statistic resulted F= 4.25 and F= 0.34, respectively, to be compared to $F_{0.05;1;3671} = 3.84$. This shows that the transformation achieved also the additivity of the effects (pentads and years) which did not hold for the original data.

The normality of the distribution of the original and transformed data tested for each pentad by the test presented in Fisher (1941), based in the statistics g_1 and g_2, measures of asymetry and kurtosis, respectively. The deviation of g_1 and g_2 from normality were tested by the bilateral Z test at the probability level of 0.05.

The deviation from normality for the original data were significant in 72 of the 73 pentads; it was not significant only in the pentad 52. On the other hand, the deviations from normality for the transformed data were significant in 53 of the 73 pentads, indicating that the transformation achieved normality of the distribution in about 30% of the 73 pentads.

So, the adopted transformation, besides of achieving the homogeneity of variance, produced the additivity of the effects of pentads and years (which could be important in other studies), and improved the normality of the distribution.

The complete results of the test are given in Baptista da Silva (1979).

3. USES OF THE TABLE

Table I contains the conditional probabilities P_c in decreasing order in the first column, from 0.990 to 0.010, and the corresponding maximum rainfalls in the following columns, one for each pentad, from 1 to 5. (P_c is the probability of the maximum rainfall being less than the amount indicated in the table for the considered pentad, given that it has rained). The probability of rain (P_1) and not rain (P_o) in each pentad are given at the bottom of the table.

The table allows the estimation of several probabilities of interest, as illustrated bellow.

3.1. Probability of rain in a given pentad (P_1)

This probability can be read out directly from the botton of the Table. Thus, for the first pentad of January, P_1=0.7307 This means that it should be expected to rain in 73 out of a set of 100 years.

3.2. Maximum rainfall corresponding to a given probability level

Consider, for example, the determination of the maximum rainfall in the first pentad of January corresponding to the probability level of 0.95. The probabilities of rain and not rain in this pentad are P_1= 0.7307 and P_o =0.2693, which, replaced in expression (3) give 0.95= 0.2693 + 0.7307 P_c; hence, P_c = 0.9316. This value is not in the Table, but is between 0.925 and 0.950, which correspond to 48.6 mm and 59.9mm, respec. Then, the maximum rainfall corresponding to the probability of rain 0.95 can be obtained by interpolation

3.3. Maximum rainfall corresponding to probability levels of 0.250 and 0.750 (25 to 75%)

In the first pentad of January, the maximum rainfalls corresponding to the levels of conditional probability 0.250 and 0.750 are respectively 4.0 mm and 22.4 mm. Thus, it should be expected that, in a set of 100 years with rainfall in this pentad, the maximum rainfall be between 4.0 and 22.4 mm in 50 of these years.

3.4. Probability level corresponding to a given maximum rainfall

Consider, for example, the determination of the probability of a maximum precipitation less than 50 mm in the first pentad of January. Since 50 mm is not in the table, one takes the closest inferior and superior values in the table, respectively 48.6 mm and 59.9 mm, which correspond to the conditional probabilities 0.925 and 0.950. Then, as $P_o = 0.2693$ and $P_1 = 0.7307$ the probability of a maximum rainfall, inferior to 50 mm, or of not rain, in the pentad is between $0.2693 + 0.7307\ (0.925) = 0.9452$ and $0.2693 + 0.7307\ (0.950) = 0.9634$, i.e., between 94,5% and 96,3%.

4. TEST OF THE TABLE

The table of the probabilities of maximum rainfall were tested using the observations recorded in 28 years (1950-71) at the agroclimatologic station of the Universidade Federal de Pelotas, distant 15 km apart from the place of the extinct meteorological station whose data were used to compute the table.

4.1. Interquartile Intervals

The maximum rainfalls corresponding to the probabilities 0.25, 0.50 and 0.75 represent, in any pentad, the 1st, 2nd and 3rd quartile, respectively. In each pentad, 25% of the maximum rainfalls, i.e., 7 out of the 28, should be in each of the interquartile intervals determined by those probabilities. The agreement between the observed and expectd values (from the table) in each pentad were tested by the χ^2 test with 3 degrees of freedom at the significance level $\alpha = 0.05$.

Thus, for the first pentad of January, the maximum rainfall corresponding to the conditional probabilities 0.25, 0.50 and 0.75 (obtained from Table I) are 4.0 mm, 10.2 mm and 22.4 mm, respectively. Since it rained in 22 of the 28 years, in five years and a half (i.e., in 25% of the 22 years) the maximum rainfall should be in each of the interquartile intervals determined by 4.0, 10.2 and 22.4. The observed numbers were 6, 3, 7 and 6; hence the computed value of $\chi^2 = 1.68$, wich is not significant ($\chi^2_{0.05;3} = 7.815$). The test showed significance for only 8 of the 73 pentads, i.e., in 11% of the pentads.

4.2. Analysis of contrasts

The three degrees of freedom among the four classes determined by the three quartiles can be decomposed into three contrasts in infinite ways. Consider the following contrasts:

$$c_1 = (n_4 + n_3) - (n_2 + n_1),$$

$$c_2 = (n_4 + n_1) - (n_2 + n_3),$$

$$c_3 = (n_4 + n_2) - (n_3 + n_1),$$

where n_i = numbers of years with rainfall in which the maximum rainfall is in the interquartile interval i ($i = 1, 2, 3, 4$).

The interpretation of the contrasts c_1, c_2 and c_3 can be seem in Fig. 2; the contrast c_1 increases with the median, c_2 increases with the dispersion and c_3 increases with the asymetry.

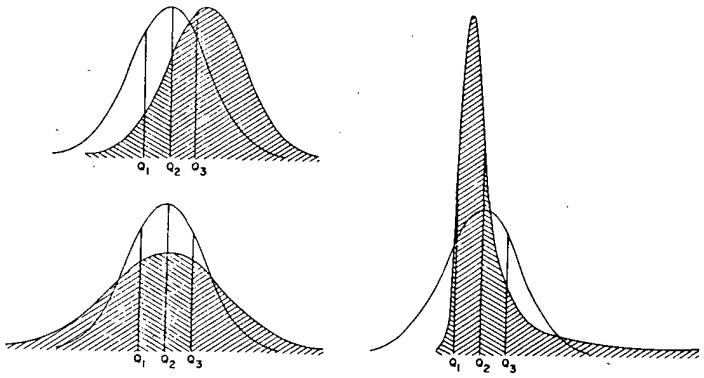


FIG 2 - INTERPRETAÇÃO DOS CONTRASTES c_1, c_2 e c_3

The agreement between the observed and expected values of the contrasts was tested by the χ^2 test with Yates correction for continuity (Fisher, 1941), with one degree of freedom, at the level of probability $\alpha = 0.05$.

The χ^2 test showed significance in one, ten and two of the 73 pentads for the contrasts c_1, c_2 and c_3, respectively.

4.3. Observed and expected values and confidence intervals for the largest maximum rainfall

A test of the table is provided by the agreement between the observed values of the largest of the maximum rainfalls and the corresponding confidence intervals computed from the table.

The following result is well known (Kendall & Stuart, 1958): Given a random sample of size k of any distribution, say $(x_1, x_2, \ldots, x_k)$, then

$$P\ (X \leq x_{(r)}) = \frac{r}{k+1},$$

where $x_{(r)}$ is the item of rank r in the array of the x's arranged in increasing order.

Thus, if P_{k_i} is the conditional probability corresponding to the largest of the maximum rainfalls in the pentad i, in k_i, years of rainfall in the total 28 years of observation, it follows that

$$E(P_{k_i}) = \frac{k_i}{k_i+1}, \quad i=1,2,\ldots,73. \qquad (4)$$

Let t is upper point of the distribution of t (unilateral) corresponding to this expected probability. Then, the expected value of the largest of the maximum rainfalls in a sucession of k_i years for the pentad i, can be determined by equation (2).

A 95% confidence interval for this expected value can be constructed as follows.

The conditional probability of the left end point of the confidence interval corresponding to 0.95 is $p_{k_i}(x_m)$ given by the expression

$$\{p_{k_i}(x_m)\}^{k_i} = 0.025 \quad ,$$

similarly, the conditional probability of the right end point of the confidence interval is $p_{k_i}(x_n)$, given by

$$\{p_{k_i}(x_n)\}^{k_i} = 0.975,$$

since all values of k_i should be less than x_n.

With the upper point of t (unilateral) corresponding to those probabilities $p_{k_i}(x_m)$ and $p_{k_i}(x_n)$, the expected values of the left and right end points of the confidence interval can be obtained using the same expression (4).

Consider, for instance, the first pentad of January, where it rained in 22 of the 28 years observation in the agroclimatologic station of UFPEL. In this case, $k_1 = 22$, hence

$$E(p_{k_1}) = 22/23 = 0.9565.$$

The corresponding t value for 37 degrees of freedom * is t= 1.7576.

Since for this pentad $\bar{z} = 15.5579$ and $z = 3.6791$, it follows that $z_c = 15.5579 + 3.6791\ (1.7576) = 22.0594$, which corresponds to the expected value of the largest maximum rainfall, x= 63.86 mm.

The conditional probabilities of the left and right end points of the 0.95 confidence interval for the first pentad of January are, respectively,

$$p_{k_1}(x_m) = \exp\ (\ln 0.025/22) = 0,8456,$$

$$p_{k_1}(x_n) = \exp\ (\ln 0.975/22) = 0.9988,$$

which correspond to the expected values of extremes of the confidence interval, x_m = 32.34mm and x_n = 206.44mm.

The confidence intervals were computed for all the 73 pentads of the year. In only seven of the 73 pentads the largest of the maximum rainfalls resulted out of the corresponding confidence interval.

The complete results of the test are presented in Baptista da Silva (1979).

*Since the estimated conditional probabilities in Table I were computed from 38 years of rainfall in the 52 years of observation.

5. REFERENCES

Amaral, E. & Baptista da Silva, J., 1970: Tabela de probabilidades das precipitações pluviométricas em Pelotas, Rs. Pelotas, Circular nº 44, Instituto de Pesquisa e Experimentação Agropecuária do Sul.

Amaral, E., Baptista da Silva, J, & Bassols, M.L., 1975: Tabelas de probabilidades da temperatura mínima em Pelotas, RS (media das mínimas e mínima absoluta). Brasília, Boletim Tecnico nº 9, Departamento Nacional de Meteorologia.

Baptista da Silva, J., 1979: Tabela de probabilidades das precipitações pluviometricas maximas pentadais em Pelotas, RS. Pelotas, Universidade Federal de Pelotas (Tese de Concurso para Professor Titular).

Fisher, R.A., 1941: Statistical methods for research workers. 8th ed., London, Oliver and Boyd.

Kendall, M.G. & Stuart, A., 1958: The advanced theory of Statistics. London, Charles Griffin & Company Limited.

Snedecor, G.W., 1957: Statistical methods. 5th ed., Ames, Iowa State College Press.

ANNUAL RAINFALL VARIABILITY IN THE NORTHEAST, BRAZIL

Ademilde de Seixas Valença
Instituto Nacional de Meteorologia

Jacques Ellis
Conselho Nacional do Desenvolvimento Científico e Tecnológico
Rio de Janeiro, Brazil

1. INTRODUCTION

The study of rainfall variability has been of worldwide special concern for many years and among those interested, Petterssen (1958) studied it by means of the coefficient of mean variability defined by the relation between the mean deviation and the mean, expressed as a percentage.

In more recent studies, the rainfall variability has been dealt by means of the coefficient of variation defined by the relation between the standard deviation and the mean, expressed as a percentage, Wilson (1971), Linacre (1977), although not taking into account the type of the frequency distribution of the variable.

In Brazil, the annual rainfall variability was studied by Torres and Mortera (1945) and Serra (1974), using the coefficient of mean variability.

In the northeast region of Brazil, annual rainfall shows a great variability which, according Petterssen (1958), is compared to the highest indices over the world, therefore to carry-out a more detailed study on the rainfall variability on that region. The high indices suggest that the studies based on probability models should be accomplished, since they provide a better characterization of the variability.

2. MATERIAL AND METHODS

The studies carried out were based on the analysis of 352 climatological series comprising the period of 1912/1958, which geographical distribution is shown in Figure 1. The data series were not up-dated, since it is not intended to obtain an analysis that makes possible a spatial comparison which period considered shows a large amount of simultaneous information consisting of a sample of approximately 45 years of observation.

NORTHEAST REGION, BRAZIL

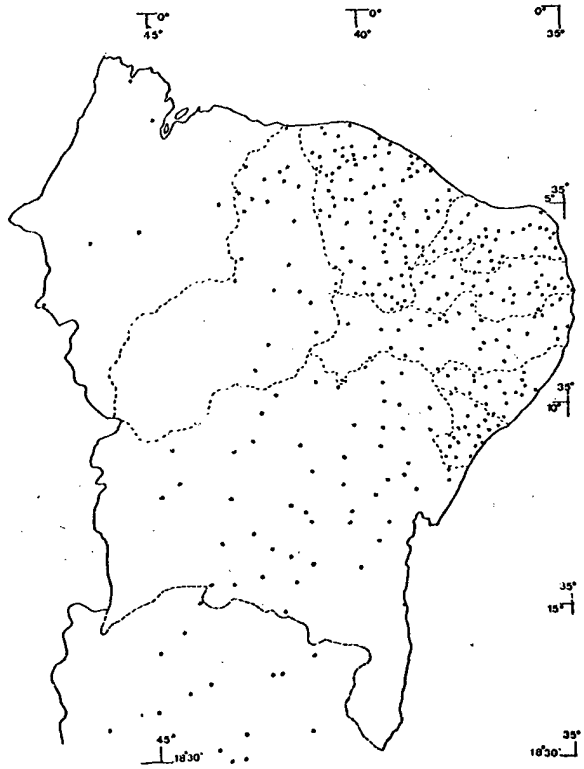


Fig. 1. Geographical distribution of the meteorological station.

As a first step, the coefficient of variation was estimated for each observation serie, plotted on a climatological map, 1:5.000.000 scale, and then analyses to provide a first hand in

formation on the rainfall variability over the region.

In studying the type of the frequency distribution which is the best representative of the annual rainfall over the northeast part of Brazil, it was accepted, in principle, the normality, since the normal distribution is the one that better fits to the annual rainfall totals, tested by means of the assimetry coefficient defined by the relation between the central moment third order and the cube of the standard deviation,

$$g_1 = \frac{\mu_3}{\sigma^3} \qquad (1)$$

and the ratio defined by the relation between the mean deviation and the standard deviation that in the normal distribution are related by the equation,

$$|e| / \sigma = \sqrt{\frac{2}{\pi}} \approx 0.80 \qquad (2)$$

The normality tests were applied to a level of 95% based on sampling distribution of the above mentioned characteristics. Cramér (1953) and Geary (1937).

In climatological series, which hypothesis of normality was not accepted, it was used the square root transformation aiming to obtain the normal distribution.

Since the distribution functions were known, the annual rainfall totals was estimated based on this functions corresponding to probabilities levels of 5%, 25%, 50%, 75% and 95%, plotted on climatological maps, 1:5.000.000 scale, and then analysed to obtain a best characterization of the annual rainfall variability.

A comparison between years considered as irregular was made analysing the anormalities which were identified by means of the probabilities of annual rainfall totals be equal or smaller than the values registered in that region in the years of 1919, 1921, 1924, 1931, 1932, 1934, 1940, 1942, 1947, 1951, 1953 and 1958.

3. RESULTS AND CONCLUSION

A first hand information on rainfall variability over the northeast of Brazil has been obtained from the coefficient of variation which isoline map is shown in Figure 2. As can be noted, the values vary from 20% to 70%, although only 2% of the area present a coefficient of variation lower than 20% whereas 16% of that area show coefficients higher than 40%. The high indices of variation can indicate that the region is submitted either to values exceptionally high or exceptionally low. Therefore, for any other reason, in these regions with great variation, it is desirable to know the lower and upper limits between which the annual rainfall totals can be expected with fixed probability level. The variability can be better characterized by those limits, which can be obtained from a probabilistic model.

It has been general practice for many years to assume that annual totals of all rainfall are normally distributed.

The tests applied at the level of 95%, did not allow reject the normality hipotesys in 200 series and the others were brought to normality by the square root transformation. Accepting the hipotesys of normality, this concept is utilized to estimate probabilities.

COEFFICIENT OF VARIATION

$$CV = 100 \times \frac{\text{standard deviation}}{\text{arithmetic mean}}$$

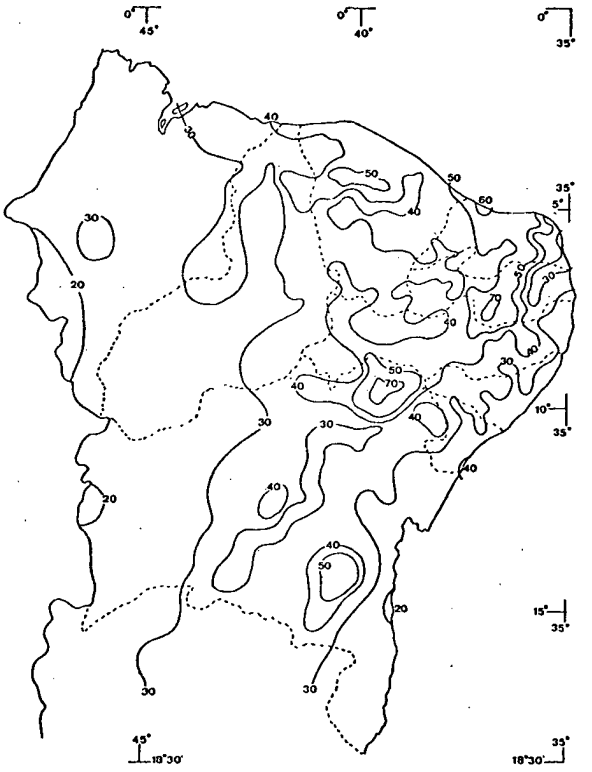


Fig. 2. Variability of the annual rainfall.

Thus, Figure 3, 4 and 5 are isohyets charts corresponding to probabilities levels of 5%, 95% and 50%, that is, the amount of precipitation which probability to be equall or smaller to those above referred.

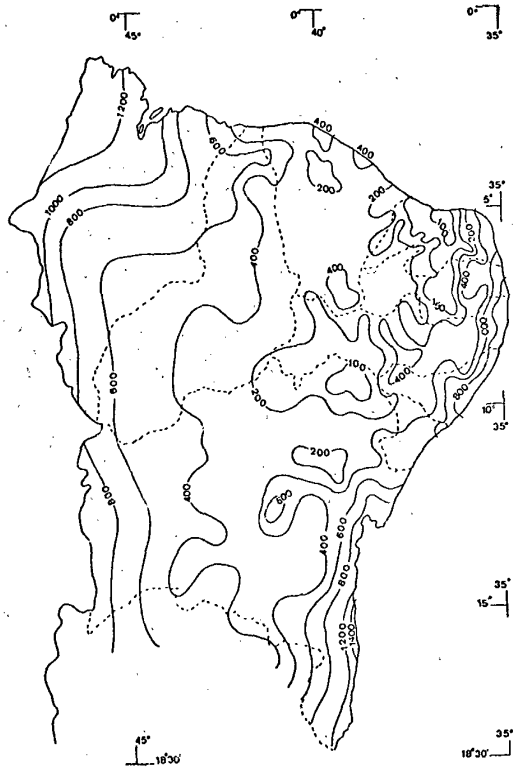


Fig. 3. Annual rainfall chart for the 5% probability level.

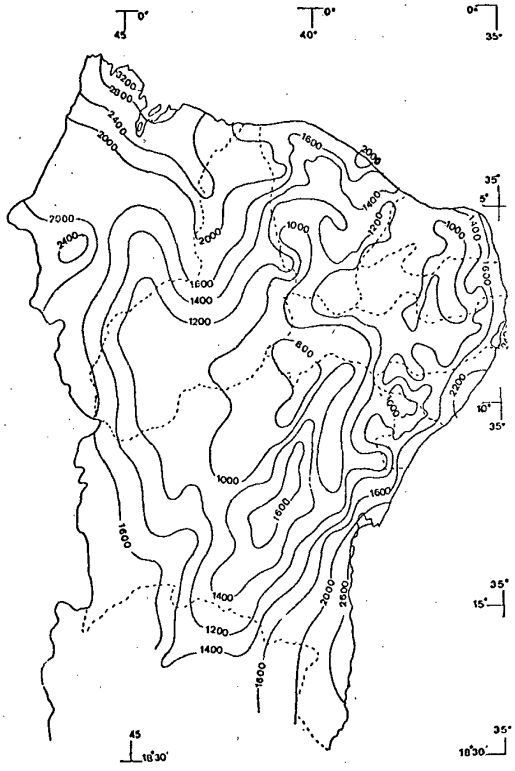


Fig. 4. Annual rainfall chart for the 95% probability level.

Annual rainfall limits, which can be expected with a probability of 90%, can be estimated through charts of 5% and 95% of probability, providing thus a detailed idea of annual rainfall variability over the region.

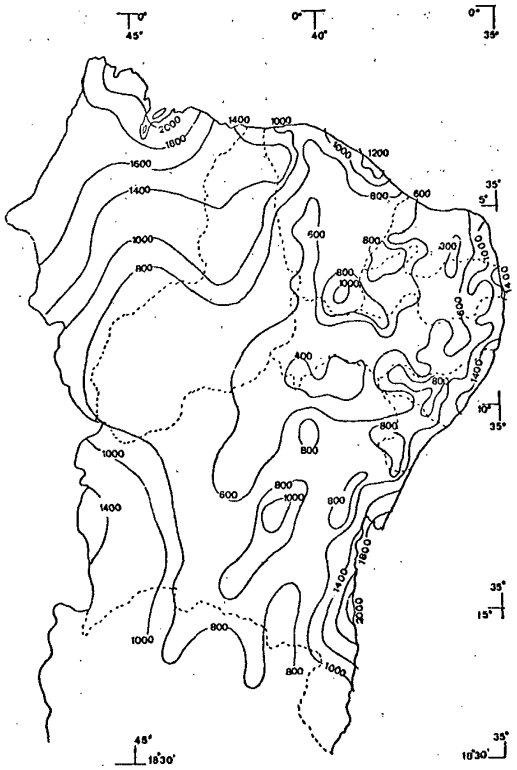


Fig. 5. Annual rainfall chart for the 50% probability level.

Thus, in those areas defined by the coefficient of variation higher than 50%, the annual rainfall can be expected for the range which the lower limit is between 100 and 200 mm, and the upper limit between 1000 and 1400 mm with a probability of 90%.

The isohyets chart for the 50% level shows that the lower values of annual rainfall, which are comprised between 300 and 600 mm, correspond to the highest values of the coefficient of variation in the region.

Another interesting exemple of application, based on the probabilistc model, is the anomalies analysis of the annual rainfall which are characterized by the occurrence probabilities of annual rainfall totals less than or equal to the recorded values.

The comparison between years is done not only by the degree of excep

tionality of rainfall which is characterized by the associated probability to the same rainfall, but also by the extension and location of areas presenting the same anomaly.

An example of application of this methodology is presented by a brief analysis of the annual rainfall totals recorded in 1924 and 1958, which presented apposite rainfall conditions.

In 1958 it can be pointed out an area characterized by the isoline of 5% probability, i.e., there is only one probability of 5% to occurrer annuals totals less than or equal to those recorded, which are comprising between 100 and 400 mm (see Fig. 6), whereas taking into account the area defined by the isoline of 20% probability it can be noted that the annual rainfall totals were less than 600 mm.

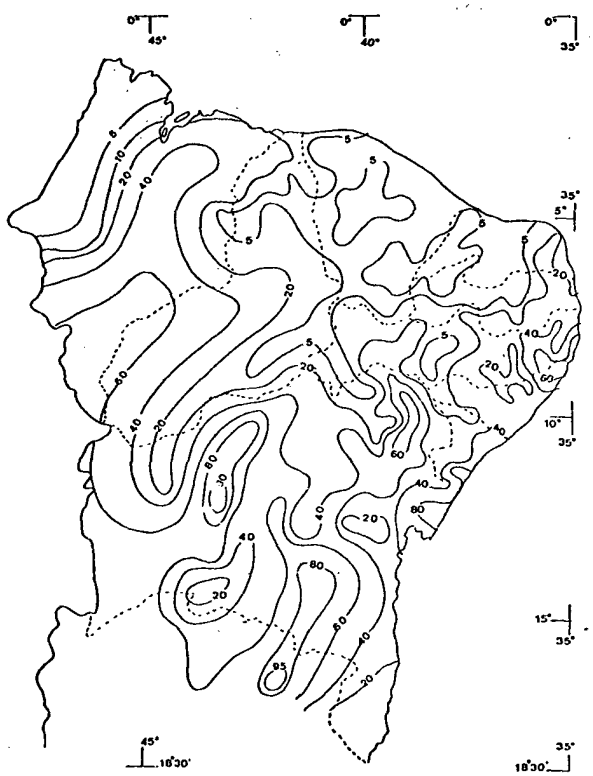


Fig. 6. Probability isolines of annual rainfall occurrence less than or equal to the recorded in 1958.

In 1924 considering the areas defined by the isoline of 95% probability, rainfall totals were recorded between 1000 and 2000 mm i.e., there is only one probability of 5% for these totals to be overpassed (see Fig. 7).

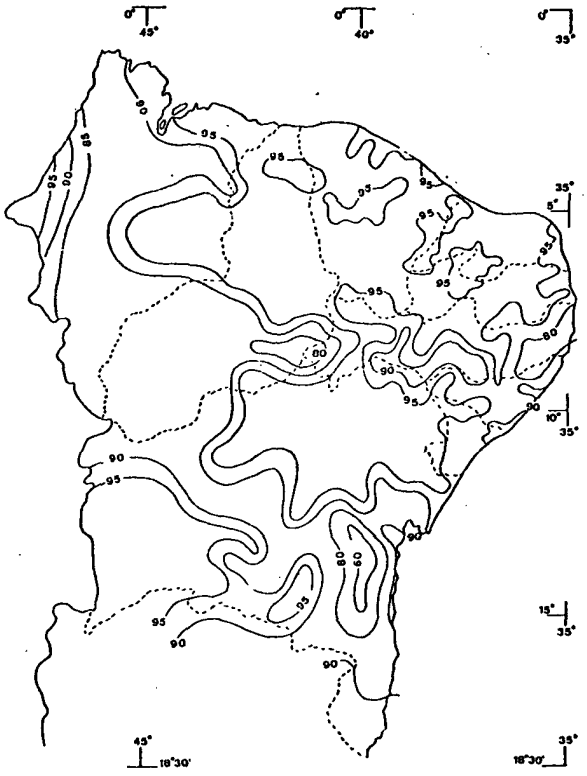


Fig. 7. Probability isolines of annual rainfall occurrence less than or equal to the recorded in 1924.

4. REFERENCES

Cramér. H., 1953, Métodos Matemáticos de Estatística, Aguilar, p. 443.

Geary, R.C., 1935, The ratio of the mean deviation to the standard deviation as a test of normality. Biometrika, Cambridge, 27 p. 310.

Linacre E. and Hobbs J., 1977, The Australian Climatic Environment, Australian, John Wiley et Sons, p. 90 - 96.

Magarino T. and Mortera A., 1948, Atlas Pluviométrico do Brasil, Brasil.

Observações Pluviométricas do Nordeste do Brasil, 1969, Departamento Nacional de Obras Contra as Secas, Brasil.

Petterssen S., 1958, Introduction to Meteorology, Second Edition. Mc Graw-Hill Book Campany. Inc.

Serra A., 1974, Climatologia do Brasil, Brasil.

Wilson C.V, 1971, The Climate of Québec, Climatic Atlas, part one, Canadian Meteorological Service, Otawa, Canadá.

AN APPROACH TO ANALYSIS AND MODELLING MULTIVARIATE, SPATIAL DATA IN METEOROLOGY AND CLIMATOLOGY*

K. Ruben Gabriel

Department of Statistics and Division of Biostatistics
University of Rochester
Rochester, NY 14627

1. STATISTICS RELATED TO GEOGRAPHICAL LOCATION

(i) Dependence of meteorological observations on location

Meteorological and climatological variables are usually related to the locations at which they are observed. Unlike some experimental variables which could equally have been recorded at one laboratory or another, most observational data depend critically on where they are observed. The average level and the variability usually vary with the topography, and the observations at different locations tend to be correlated in a manner that often reflects geographical proximity and the direction of prevalent winds.

Statistical evaluations of meteorological variables often incorporate the locations of the observations by such time-honored representations as isohyets, isobars, etc.. These displays present only the average levels at the different locations, though they may be elaborated to allow for the representation of variability as well. The representation and modelling of the correlations between the variables observed at different locations is more difficult, because it requires joint representation for all locations rather than several separate representations for the different locations.

(ii) Usefulness of study of spatial variation

Study of the spatial variation and covariation of meteorological variables is interesting both in itself, for the models it can suggest, and as a tool for the study of phenomena which might otherwise be obscured by prevailing topographical patterns. Description and summarization of spatial variation of weather occurences over a region can lead to the formulation of models and parameters which may have interesting physical interpretations. This may help to understand the phenomena and relate them to other relevant parameters. Moreover, such models may provide the basis for studies of other weather related events, such as inadvertent weather modification. (Changnon, 1968; Changnon and Huff, 1968; Changnon, Huff, and Semonin, 1971; Changnon, Semonin, and Huff, 1976.) In studies of this kind one tries to disentangle the local effects of some special factor from the existing background variation in natural weather which is known to have a strong geographical component and can be described by a spatial model. In fact, the methodology of studying inadvertent weather modification seems to consist mostly of attempts to adjust, compensate for, partial out, or otherwise eliminate, the background variation which is mostly a function of the topography. Similar methods can be used in the analysis of cloud seeding experiments and commercial operations in which topographical features may either simulate or obscure a possible seeding effect (Huff, 1968; Scott, 1980). Models of spatial variation may then be usefully employed to isolate the topographical compenent and thus allow more precise evaluation of the effect that may be ascribed to seeding.

(iii) Methods of study of spatial variation

Methods of studying spatial variation mostly model the relation of average levels to location by fitting trend surfaces, isoline contours, etc. There is a very considerable literature on such statistical methods and its applications, ranging from geology (e.g. Matheron, 1963, 1965) and biology (e.g. Matern, 1960; Sokal and Oden, 1978; Ripley, 1981) to the social sciences (e.g. Bachi, 1968) and including meteorology and climatology (e.g., Gandin and Kagan, 1974; Huff, 1968).

There have also been attempts to study correlation over a region, including those of meteorological and climatological variables. An early example is Fisher and Mackenzie's (1923) work on the correlation of precipitation at stations in terms of the distances between them. The phenomenon of decay of correlation with increasing distance has been studied by several workers (Watson, 1927; Huff and Shipp, 1969; Longley, 1974; and others). This work has, however, suffered from the difficulty of finding mathe-

*Supported by ONR Contract N00014-80-C-0387 on "Biplot Multivariate Graphics," K. R. Gabriel, Principal Investigator

matical functions that can adequately, model this correlation-versus-distance relationship.

It is difficult to model the decay of correlation with distance, i.e., to find a monotone decreasing function r(d) that expresses the correlation r between the observations at two stations d miles apart. One would have thought it was pretty straightforward to take the data for, say, 50 stations, calculate the (50 x 49)/2 correlations and distances between all pairs, and fit a suitable function to this scatter (Julian and Thiebaux, 1975). One soon finds out, however, that a seemingly well-fitting curve may give inconsistent results. Thus, if one considers a new, i.e., unobserved, set of stations with given distances between them and uses the fitted r(d) curve to interpolate the correlations for these stations, one may well obtain a non-positive-definite correlation matrix. In other words, some of the variances of combinations of these stations' observations turn out to be negative--surely not a sensible result, or one that one could apply in practice. To avoid such inconsistencies one must do the curve fitting within a very restrictive class of functions (Thiebaux, 1976, 1977). It is not a simple matter to identify this class or to find statistical methods that identify the particular function from this class which fits the given data most closely.

It would seem that two things are required for the development of a good methodology of spatial statistics for meteorological and climatological data. One of them is a class of models which are adequate mathematically to deal with such phenomena, and which avoid the risk of absurdities such as negative variances. The other is a set of statistical techniques that will allow the exploration of spatial data and the diagnosis of models that may fit them. It would be best if the statistical techniques were tailored directly to the diagnosis of models suitable for the description of spatial data.

We propose a geometric-geographical model for spatial data and show that biplot graphical display of data is particularly suited to diagnosing this type of model.

2. THE EMPIRICAL BASES FOR THE PROPOSED MODELS

(i) The first observed phenomenon: Correlational Decay

The model we propose is based on two phenomena which have been repeatedly observed on meteorological data. One is the decay of correlations with distance and its modification in certain directions; the other is the geometric pattern of correlations between localities which appears to be a systematic distortion of their locations on the globe. These two phenomena will be explained in some detail.

It is an intuitively plausible, and empirically well established, phenomenon that meteorological variables observed at nearby locations are more highly correlated than the same variables observed at locations farther apart. This is readily confirmed by studies of spatial patterns. However, these studies also show that correlation is not a function of distance only, but of direction as well, i.e., it is not isotropic. For example, Sharon (1974, 1978) plotted contours of correlations of precipitation at pairs of stations as functions of both distance between the stations and the direction of one with respect to the other. He noted that, at a given distance, the correlation was highest if the direction between them was that of the prevailing winds, and lowest if the direction was normal to that. Other workers have also noted such an anisotropy, and some have modelled rainfall patterns by elliptic isohyets, with the major axis in the direction of the prevailing winds (Silverman, Rogers and Dahl, 1981; Gabriel, 1981a). Clearly, a distance-decay model of precipitation correlation must incorporate something like a scaling parameter to allow for the evident anisotropy.

(ii) The second empirically observed phenomenon: Concordance of geographic and correlational geometries

A further phenomenon has been observed when spatial data were analyzed by statistical methods which are essentially geometric, or display data graphically in a biplot of principal components or other axes. These methods have shown that precipitation data for a number of stations within a given region have a large coordinate on the first axis (or principal component). That in itself is not surprising, since there obviously is a strong "average" rainfall element affecting an entire region during a particular storm, day, or month. However, this first component has usually been found to be accompanied by two further coordinates which correspond fairly well with the geographical coordinates of the locations. The second coordinate is often found to correspond with a coordinate in the prevailing wind direction and the third coordinate with the direction normal to that. Thus, if one plots the stations by their loadings on the second and third principal components, one finds a fairly close reproduction of the map of the locations, rotated to correspond to the wind direction and somewhat expanded in the direction normal to it. (This contraction reflects the relatively higher correlations per given distance in the prevailing wind direction.)

(iii) Examples of meteorological phenomena

One example of such a pattern comes from data on rainfall in convective bands passing over the Santa Barbara area of Southern California. Those data were recorded during the weather modification experiment carried out there by North American Weather Consultants, and resulted from an analysis by Elton Scott (1980). The loadings of all stations on the first principal component were found to be uniformly large--suggesting a strong overall rainfall component affecting the entire area pretty much equally. A biplot of the control area stations' loadings on the second and third components - Figure 1 - is found to have much similarity with the map of these stations - Figure 2 (Gabriel, 1980a, also contains some other illustrations of this phenomenon.)

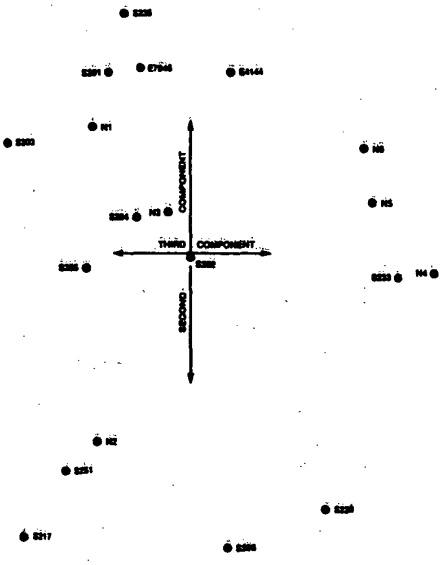


Fig. 1. Santa Barbara Rainfall in Control Area. Biplot loadings on 2nd and 3rd axes.

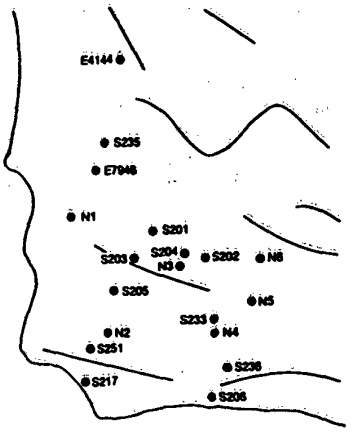


Fig. 2. Santa Barbara Rainfall: Map of Control area Stations

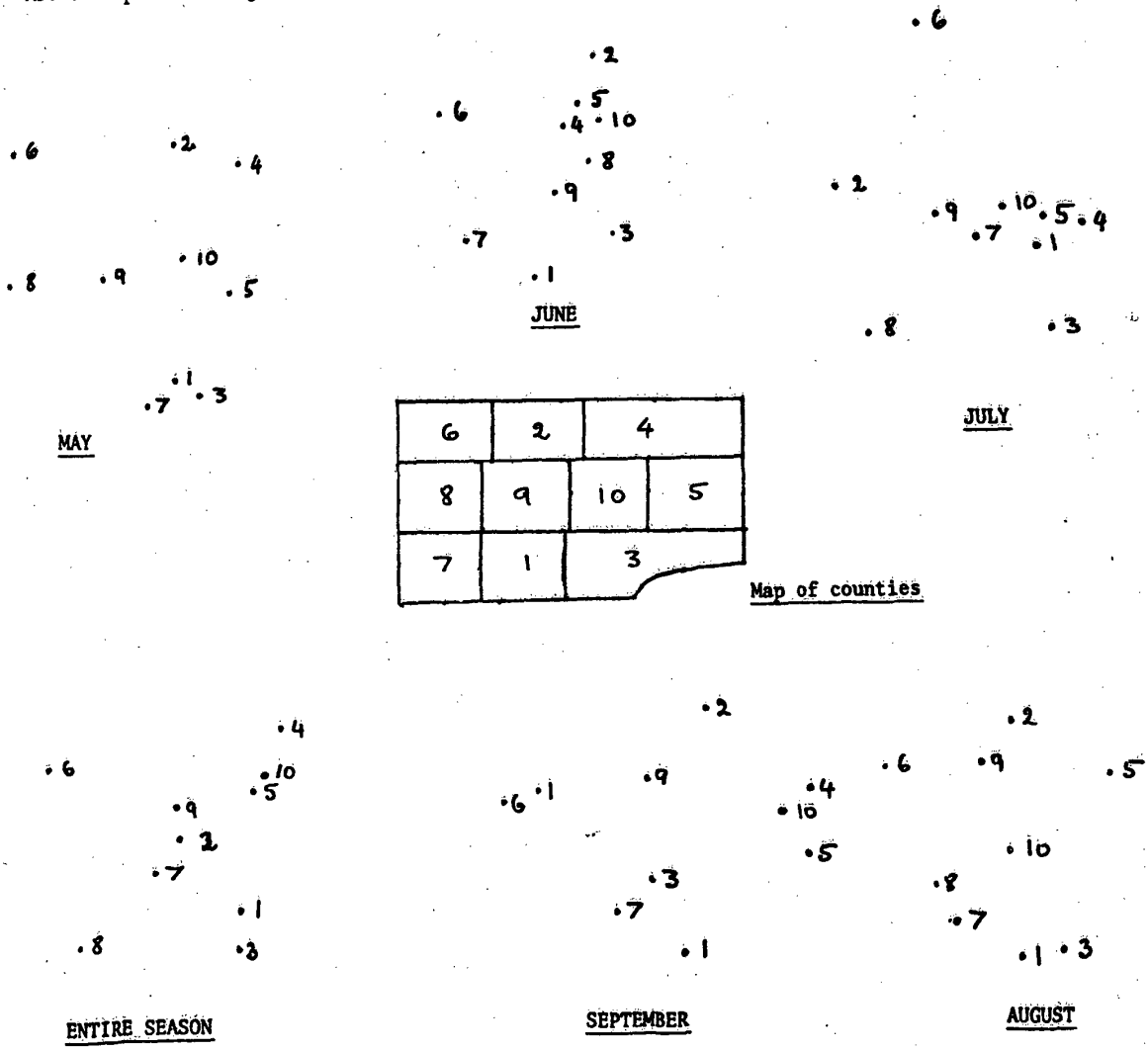


Fig. 3. Ten Kansas Counties Map and Biplots of Monthly Rainfalls (Components 2 & 3)

A similar example comes from monthly precipitation records of ten counties in Kansas. The correlations between the counties were obtained separately for each month--May through September--as well as for all months together (Hsu, 1983). Figure 3 shows the map of the ten counties' locations and the second and third components' biplots of the correlations for each month's rainfall. (The latter having been rotated so as to allow easier comparison with the map.) The pattern of the loadings on the biplots for most of the months is seen to correspond quite well with the geographical configuration of the counties.

The usefulness of these patterns in analyzing weather modification experiments has been illustrated by Corsten and Gabriel (1976). They biplotted data of the first Israeli randomized cloud seeding experiment and found that both inter-area correlation and variability of rainfall was larger during seeding than otherwise. They were able to disentangle these effects from the covariance matrices largely because the overall pattern of variability and covariability corresponded closely to the geographic configuration of the sub-areas of the experiment. They could therefore ascribe deviations from that pattern to the experimental treatment, i.e., to cloud seeding.

(iv) Formulation of a model relating correlations to locations

Recurring observations of these phenomena suggests that there must be some systematic pattern of spatial variation. It presents the challenge of formulating a general model, or family of models, to account for these patterns. And, of course, one would wish for the chosen models also to account for the phenomenon of decay of correlation with distance and the anisotrophy observed on it. An obvious direction to search for a suitable method of modelling is the well-known geometric representation of correlational and variance-covariance configurations. One could then seek to relate the geometry of the geographical locations to the geometry of the correlations.

The geographic geometry can be visualized as m vectors $\underline{\ell}_1, \underline{\ell}_2, \ldots, \underline{\ell}_m$ from the globe's center to its circumference. Clearly, if the localities are in a region smaller than, say, a continent, these m vectors will form a fairly tightly knit sheaf. For a suitable choice of reference axes--the origin being at the globe's center and the first axis through, or near to, the localities' centroid--the coordinates with respect to the first axis must be relatively large and pretty much constant (especially if the region considered is small). Moreover, the projection onto the plane of the second and third axes forms a planar projection of the map of the localities.

The correlational geometry models variables as vectors in Euclidean space, and correlations as cosines of the angles between these vectors. Thus, for a set of m variables $Y_1, Y_2, \ldots, Y_m$, say, one postulates m vectors $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_m$ such that

$$\text{Corr}\ (Y_j, Y_g) = \text{Cos}\ (\underline{v}_j, \underline{v}_g)$$

$$j, g = 1, 2, \ldots, m.$$

If the vectors $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_m$ are further postulated to be of unit lengths, then

$$\text{Corr}\ (Y_j, Y_g) = \underline{v}_j' \underline{v}_g,$$

i.e., the correlation is the inner product of the corresponding vectors. In that case, if one writes R for the $(m \times m)$ matrix of correlations and V for the matrix whose j-th row is $\underline{v}_j'$, it is clear that

$$VV' = R.$$

This shows that the non-negative definite property of a correlation matrix is always satisfied by such a vector model. Also, any correlation matrix, being by definition non-negative, can be factored as above, so that a vector model is always available.

For data on variables Y_j, $j=1,\ldots,m$, at m localities, these thus are two geometries. First, the geographic 3D geometry of the locations $\underline{\ell}_1, \ldots, \underline{\ell}_m$, on the globe, and secondly, the higher-dimensional correlational geometry of m vectors $\underline{v}_1, \ldots, \underline{v}_m$. If the vectors of the correlational geometry are largely concentrated in a three dimensional subspace, they may be closely related to the geographic geometry.

The phenomenon that has been repeatedly observed is that the 3D approximation to the correlational vector geometry is quite similar to the geographical geometry. First, the coordinates on the first principal axis of the vectors were found to be large and pretty much the same for all m vectors. Secondly, plots of the coordinates on the second and third principal axes were found to reproduce the map of the localities more or less closely. Suitable rotation of the plane of the second and third axes and differential scaling along the two axes, yielded quite good concordance of the two geometries.

The relation between the geographic geometry of the locations and the correlational geometry can be modelled in a variety of ways. The simplest approach assumes them to be equal, i.e., the angle subtended at the center of the globe by the radiuses of any two locations equals the arc cosine of the correlation between them. This may be called the simple basic model. An elaboration--the scaled basic model--assumes that the two geometries could be rotated in such a manner that they differed only by different scalings along the three axes. Other elaborations fix certain directions a priori, e.g., the radial direction to the centroid of the points and perhaps the direction of prevalent winds. Thus only limited rotations and scalings are allowed.

The mathematical definitions of some models and methods of fitting them to data are reviewed next. This work is by no means complete at this time, nor is the study of the relation between these models and the anisotropic functions of correlation decay.

3. MODELS OF THE RELATION BETWEEN THE GEOGRAPHIC GEOMETRY OF LOCATION AND THE CORRELATIONAL GEOMETRY OF VARIABLES OBSERVED AT THESE LOCATIONS

(i) The basic models

Consider a set of variables $Y_j (j=1,\ldots,m)$ associated with localities $\underset{\sim}{\ell}_j=(\ell_{j,1},\ell_{j,2},\ell_{j,3})'$ $(j=1,\ldots,m)$, where $\ell_{j,i}$ is the coordinate of locality j with respect to the ith of some 3 given axes. Since the locations are on the globe, one would obtain, if the origin were set at the globe's center.

$$\| \underset{\sim}{\ell}_j \|^2 = \sum_{e=1}^{3} \ell_{j,e}^2 = \gamma^2, \tag{1}$$

where γ is the radius of the globe. (It is unlikely that the small non-sphericity of the earth and the altitudes of the stations need be taken into account for the present modelling purposes.)

To model observed correlations $r_{j,g}$ between variables $Y_j, Y_g (j,g=1,\ldots,m)$ one may transform the location coordinates $\underset{\sim}{\ell}_j$ to new, so called, model vector coordinates $\underset{\sim}{\lambda}_j$. This is done by means of a rotation--using an orthogonal matrix $C_{(3x3)}$--and scalings along the new (rotated) axes--which use a diagonal matrix $S=\mathrm{Diag}(s_1,s_2,s_3)$. The resulting transformation is

$$\underset{\sim}{\lambda}'_j = \underset{\sim}{\ell}'_j C'S \qquad j=1,\ldots,m \tag{2}$$

or, equivalently

$$\Lambda = LC'S, \tag{3}$$

where $L_{(mx3)}$ has rows $\underset{\sim}{\ell}'_1,\ldots,\underset{\sim}{\ell}'_m$ and $\Lambda_{(mx3)}$ has rows $\underset{\sim}{\lambda}'_1,\ldots,\underset{\sim}{\lambda}'_m$.

The model vectors $\underset{\sim}{\lambda}_j$ are seen to be located on an ellipsoid whose semi-axes are of lengths $\gamma s_1, \gamma s_2$, and γs_3, respectively, along the direction of the rotated axes. This is evident from

$$\begin{aligned} \sum_{i=1}^{3} (\lambda_{j,i}/s_i)^2 &= \underset{\sim}{\lambda}'_j S^{-2} \underset{\sim}{\lambda}_j \\ &= \underset{\sim}{\ell}'_j C'SS^{-2}SC\underset{\sim}{\ell}_j \\ &= \underset{\sim}{\ell}'_j \underset{\sim}{\ell}_j \\ &= \gamma^2. \end{aligned} \tag{4}$$

Thus, the transformation from locations to model vectors is from points on a sphere to points on an ellipsoid.

The proposed models express the correlations as cosines of the model vectors, that is,

$$r_{j,g} \underset{m}{=} \cos(\underset{\sim}{\lambda}_j, \underset{\sim}{\lambda}_g), \tag{5}$$

where $\underset{m}{=}$ indicates that the right hand expression models, i.e., approximates, the left hand side. This can be written

$$r_{j,g} \underset{m}{=} \underset{\sim}{\lambda}'_j \underset{\sim}{\lambda}_g / (\| \underset{\sim}{\lambda}_j \| \quad \| \underset{\sim}{\lambda}_g \|) \tag{6}$$

or

$$R \underset{m}{=} D_\Lambda^{-1} \Lambda\Lambda' D_\Lambda^{-1}, \tag{7}$$

where

$$D_\Lambda = \mathrm{Diag}(\| \underset{\sim}{\lambda}_i \|) \tag{8}$$

is the (mxm) matrix such that $D_\Lambda^2 = \mathrm{Diag}(\Lambda\Lambda')$.

Since $\Lambda\Lambda'$, and hence also $D_\Lambda^{-1}\Lambda\Lambda' D_\Lambda^{-1}$, cannot have rank greater than 3, it may be more reasonable to consider the approximate model

$$R_{[3]} \underset{m}{=} D_\Lambda^{-1} \Lambda\Lambda' D_\Lambda^{-1}, \tag{9}$$

where $R_{[3]}$ is the least squares rank three approximation to R.

This approximation can be obtained from the singular value decomposition of R,

$$R = \sum_{a=1}^{m} \rho_a^2 \underset{\sim}{q}_a \underset{\sim}{q}_a', \tag{10}$$

with $\underset{\sim}{q}'_a \underset{\sim}{q}_b = 1$ or 0 according to whether $a = b$ or $a \neq b$,

and $\rho_1^2 \geq \rho_2^2 \geq \cdots \geq \rho_m^2 \geq 0$.

The approximation--in the least squares sense--is

$$R_{[3]} = \sum_{a=1}^{3} \rho_a^2 \underset{\sim}{q}_a \underset{\sim}{q}_a', \tag{11}$$

according to the Householder-Young (1938) Theorem. If one writes

$$H_{(mx3)} = (\rho_1 \underset{\sim}{q}_1, \rho_2 \underset{\sim}{q}_2, \rho_3 \underset{\sim}{q}_3), \tag{12}$$

a loadings matrix with rows $\underset{\sim}{h}'_1,\ldots,\underset{\sim}{h}'_m$, then (9) becomes

$$HH' \underset{m}{=} D_\Lambda^{-1} \Lambda\Lambda' D_\Lambda^{-1} \tag{13}$$

and it follows that

$$H \underset{m}{=} D_\Lambda^{-1} \Lambda T \tag{14}$$

for some orthonormal (3x3) matrix T.

This shows the normalized vectors

$$\underset{\sim}{v}'_j = \underset{\sim}{\lambda}'_j / \| \underset{\sim}{\lambda}_j \| \tag{15}$$

i.e., the columns of $D_\Lambda^{-1}\Lambda$, to be rotated into the loadings $\underset{\sim}{h}'_j$ by

$$\underset{\sim}{h}'_j \underset{m}{=} \underset{\sim}{v}'_j T. \tag{16}$$

Note that rotation T is different from rotation C--though both are (3x3) orthonormal matrices--which rotates the locations $\underset{\sim}{\ell}'_j$ to the axes of the variable vectors $\underset{\sim}{v}'_j$. In general, these two rotations will not be the same.

The above model relationship may be rewritten in terms of the original location coordinates $\underline{\ell}_j$ as:

$$H \underset{m}{=} D^{-1}_{LC'ST} LC'ST, \tag{17}$$

since $D_{LC'ST}=D_{LC'S}$, by definition of D in (8). But it is well known that the set of all products "rotation-by-scaling-by-rotation," i.e., in matrix terms "orthogonal-by-diagonal-by-orthogonal," is equal to the set of all matrices, and so one may rewrite (17), without loss of generality, as

$$H \underset{m}{=} D^{-1}_{LM} LM \tag{18}$$

for some matrix $M_{(3x3)}$, which is given by

$$M = C'ST. \tag{19}$$

(In fact this factorization yields its singular value decomposition.)

(ii) Correlational Decay

To develop the relation between correlation and distance between locations, one may use approximation

$$r_{j,g} \underset{m}{=} 1-\tfrac{1}{2} \| \underline{v}_j-\underline{v}_g \|^2 \tag{20}$$

which follows from the equality

$$2\underline{v}_j\underline{v}_g = \| \underline{v}_j \|^2+\| \underline{v}_g \|^2-\| \underline{v}_j-\underline{v}_g \|^2 \tag{21}$$

in view of the normalization that makes $\| \underline{v}_j \| =\| \underline{v}_g \| =1$. Thus the correlation between Y_j and Y_g is a monotene decreasing function of the distance between $\underline{v}_j$ and $\underline{v}_g$ which are on a unit sphere but subtend the same angle at the origin--center of the sphere--as the model vectors $\underline{\lambda}_j$ and $\underline{\lambda}_g$. In fact, the $\underline{v}_j$'s may be visualized as truncated versions of the $\underline{\lambda}_j$'s, truncation occuring where the model vector intersects with the unit sphere.

In terms of the original locations, the above relation becomes

$$r_{j,g}=1-\tfrac{1}{2}(\tilde{\underline{\ell}}_j-\tilde{\underline{\ell}}_g)'C'S^2C(\tilde{\underline{\ell}}_j-\tilde{\underline{\ell}}_g), \tag{22}$$

where

$$\tilde{\underline{\ell}}_j=\underline{\ell}_j / \sqrt{\underline{\ell}_j'C'S^2C\underline{\ell}_j} \tag{23}$$

is a normalized form of the original localities. Since the matrix $C'S^2C$ rotates the location vectors and scales them to fit on the ellipse discussed above, the above expression is close to proportional to the distance between the model vectors $\underline{\lambda}_j,\underline{\lambda}_g$ on the ellipsoid.

(iii) Least squares methods of Fitting

At times, the model may be assumed even though the matrices C,S and T are not known. In that case, they may be "estimated" in a manner that will produce closest fit of the model to the data.

If neither C nor S nor T are known in any way we may estimate them by finding $\hat{M}$ which solves

$$\min_M \| H-D^{-1}_{LM}LM \|^2 = \| H-D^{-1}_{L\hat{M}} L\hat{M} \|^2. \tag{24}$$

The singular value decomposition then yields a factorization.

$$\hat{M} = \hat{C}'\hat{S}\hat{T}, \tag{25}$$

which provides estimated rotation and scaling matrices.

If the rotation matrices C and T are known--the latter usually as $T=I_3$--then S would be estimated by $\hat{S}$ which solves

$$\min_{S,\text{diagonal}} \| HT'-D^{-1}_{LC'S}LC'S \|^2=\| HT'-D^{-1}_{LC'\hat{S}}LC'\hat{S} \|^2. \tag{26}$$

Note that the three non-zero elements of $\hat{S}$ can be multiplied by a constant without changing the solution. One may therefore set $\hat{s}_{1,1}$ arbitrarily at any chosen positive value and solve for $\hat{s}_{2,2}$ and $\hat{s}_{3,3}$.

(iv) Use of Biplots

Least squares fitting of the basic model was seeen to use the singular value decomposition of a fitted coefficient matrix. Evidently, graphical display of the singular (i.e., principal) components will allow inspection of the fit and its residuals. The biplot method of display-especially in its 3D version which has been developed recently (Tsianco, 1980)-is especially suitable for this purpose. Indeed, it was the inspection of biplots of precipitation data that first suggested the type of model discussed here. Computer programs for 3D-biplot display are available in portable form (based on the NCAR or DISPLAY graphics package).

(v) Extensions and Applications

It is unlikely that the basic models proposed above will adequately fit all meteorological and climatological data. Data analyses and biplot displays are likely to suggest further modelling. It is premature to speculate about that now.

To the extent that the proposed models do fit data like precipitation, it will be of interest to see how the scale parameters are related to different types of observations. For example, the correlation of two stations at a given distance is likely to be larger for weekly rain totals than for daily, or hourly, precipitation. The locational geometry is, however, the same. Hence the scaling parameters

should vary with the length of the period of observation. It should be interesting to see whether such a trend will be found and to what extent it can be further modelled.

The range of applications of models of the kind proposed may be considerable. Obvious instances are in studies of weather modification, advertent or inadvertent, and in interpolation of meteorological observations between gauge locations. Again, it is too early to say more.

4. FINAL COMMENTS

The study of the proposed geographic-correlational model is in its first stages. A basic model has been formulated and methods of analysis proposed. Some small examples lend credence to the model and method, but a thorough study has yet to be carried out.

5. REFERENCES

Bachi, R. (1968). Graphical Rational Patterns: A New Approach to Graphical Representation of Statistics. Jerusalem: Israel Universities Press.

Bradu, D. and Gabriel, K.R. (1978). "The Biplot as a diagnostic tool for models of two-way tables." Technometrics, 20, 489-498.

Changnon, S. A. Jr. (1968). "The La Porte weather anomaly--Fact or Fiction?" Bull. Amer. Meteor. Soc., 49, 4-11.

Changnon, S. A. Jr. and Huff, F. A. (1977). "La Porte again: A new anomaly?" Bull. Amer. Meteor. Soc., 58, 1069-1072.

Changnon, S. A. Jr., Huff, F.A. and Semonin, R.G. (1971). "METROMEX: An investigation of inadvertent weather modification." Bull. Amer. Meteor. Soc., 52, 958-967.

Changnon, S. A. Jr., Semonin, R.G. and Huff, F.A. (1976). "A hypothesis for urban rainfall anomalies." J. Appl. Meteor., 15, 544-560.

Corsten, L.C.A. and Gabriel, K.R. (1976). "Graphical exploration in comparing variance matrices." Biometrics, 32, 851-863.

Fisher, R.A. and Mackenzie, W.A. (1922). "The correlation of weekly rainfall." Quart. J. Roy. Meteor. Soc., 48, 234-245.

Gabriel, K.R. (1971). "The biplot--graphic display of matrices with application to principal component analysis." Biometrika, 58, 453-467.

Gabriel, K. R. (1972). "Analysis of meteorological data by means of canonical decomposition and biplots." J. Appl. Meteor., 11, 1071-1077.

Gabriel, K.R. (1978). "Least squares approximation of matrices by additive and multiplicative models." J. Roy. Statist. Soc., B, 40, 186-196.

Gabriel, K. R. (1980). "Comments on the reanalysis of the Santa Barbara II cloud seeding experiments", in Statistical Analysis of Weather Modification Experiments. (E. J. Wegman and D. J. DePriest, eds.). New York: Dekker. pp. 113-129.

Gabriel, K. R. (1980). "Biplot". Encyclopedia of Statistical Sciences. (S. Kotz and N. L. Johnson, eds.) New York: Wiley.

Gabriel, K. R. (1981a). "Gage density and variability of rainfall estimates: A complement to Silverman, Roger and Dahl's study." J. Appl. Meteor., 20, 1537-1542.

Gabriel, K. R. (1981b). "Biplot display of multivariate matrices for inspection of data and diagnosis." Chapter 8 of Interpreting Multivariate Data. (V. Barnett, ed.) London: Wiley.

Gabriel, K. R. and Odoroff, C. L. (1983). "Resistant lower rank approximation of matrices." University of Rochester, Statistics Technical Report 83/02.

Gabriel, K. R. and Zamir, S. (1979). "Lower rank approximations of matrices by least squares with any choice of weights", Technometrics, 21, 489-498.

Gandin, L. S. and Kagan, L. R. (1974). "Construction of a system for objective analysis of heterogeneous data based on the method of optimum interpolation and optimum agreement." Meteor. Hydrol., 5, 1-10.

Gower, J. C. (1975). "Generalized Procrustes Analysis." Psychometrika, 40, 33-51.

Gower, J. C. and Digby, P.G.N. (1981). "Expressing complex relationships in two dimensions." Chapter 6 of Interpreting Multivariate Data (V. Barnett, ed.) London: Wiley.

Householder, A.S. and Young, G. (1938). "Matrix approximation and latent roots." Amer. Math. Monthly, 45, 165-171.

Hsu, C. F. (1983). Report of OSET Research

Huff, F. A. (1968). "Area-Depth Curves--A useful tool in weather modification experiments." J. Appl. Meteor., 7, 940-943.

Huff, F. A. and Shipp, W. L. (1969). "Spatial correlation of storm, monthly and seasonal precipitation." J. Appl. Meteor., 8, 542-550.

Julian, P. R. and Thiebaux, H. J. (1975). "On some properties of correlation functions used in optimum interpolation schemes." Monthly Weather Rev., 102, 605-616.

Longley, R. W. (1974). "Spatial variation of precipitation over the Canadian prairies." Monthly Weather Rev., 102, 307-312.

Matern, B. (1960). Spatial Variation. Meddelanden fran Statens Skogsforskningsinstitut, 49, 5, 1-144.

Matheron, G. (1963). "Principles of geostatistics," Economic Geology, 58, 1246-1266.

Matheron, G. (1965). Les Variables Regionalisees et leur Estimation. Paris: Masson.

Pearson, K. (1901). "On lines and planes of closest fit to a system of points." Phil. Mag., Ser. 6, 2, 559-572.

Ripley, B. D. (1981). Spatial Statistics. New York: Wiley.

Scott, E. (1980). "A multivariate methodology for the analysis of weather modification experiments," in Statistical Analysis of Weather Modification Experiments. (E.J. Wegman and D.J. DePriest, eds.). New York: Dekker, pp. 61-77.

Sharon, D. (1974). "On the modelling of correlation functions for rainfall studies." J. Hydrol., 22, 219-224.

Sharon, D. (1978). "Rainfall fields in Israel and Jordan and the effect of cloud seeding on them." J. Appl. Meteor., 17, 40-48.

Silverman, B. A.; Rogers, C.K. and Dahl, D. (1981). "On the sampling variance of raingage networks." J. Appl. Meteor., 20, 1468-1478.

Sokal, R. R. and Oden, N. L. (1978). "Spatial autocorrelation in biology. I. Methodology. II. Some biological applications of evolutionary and ecological interest." Biol. J. of Linnean Soc., 10, 199-228, 229-249.

Thiebaux, H. J. (1976). "Anisotropic correlation functions for objective anlaysis." Monthly Weather Rev., 104, 994-1002.

Thiebaux, H. J. (1977). "Extending estimation accuracy with anisotropic interpolation." Monthly Weather Rev., 105, 691-699.

Tsianco, M. C. (1980). Use of Biplots and 3D Bimodels in Diagnosing Models for Two-way Tables. University of Rochester, Ph.D. Thesis.

Tsianco, M. C. and Gabriel, K. R. (1983). "Modelling Temperature Data: An illustration of the use of Biplots and Bimodels in non-linear modelling." Submitted to J. Appl. Meteor.

Tsianco, M. C.; Odoroff, C. L.; Plumb, S. and Gabriel, K. R. (1981). BGRAPH--Program for Biplot Multivariate Graphics: User's Guide. University of Rochester, Statistics Technical Report 81/20.

Watson, R. A. (1927). "Correlation of daily rainfall on the east and west of Scotland." Quart. J. Roy. Meteor. Soc., 53, 183-185.

Whittle, P. (1962). "Topographic correlation, power-law covariance functions and diffusion." Biometrika, 49, 305 - 314.

RECONSTRUCTING SURFACE PRESSURE PATTERNS USING PRINCIPAL COMPONENTS REGRESSION ON TEMPERATURE AND PRECIPITATION DATA

P.D. Jones, T.M.L. Wigley and K.R. Briffa

Climatic Research Unit,
University of East Anglia,
Norwich NR4 7TJ, U.K.

1. INTRODUCTION

In order to properly understand the causes of climatic change it is essential to have a long and detailed record of past climates. Although there are isolated single-site or limited-area records extending back many centuries (e.g. Manley, 1974; a record of Central England temperatures back to 1659), these are of limited value since their representativeness in terms of larger-scale climatic change is not only poorly understood but also liable to vary with time (Jones and Kelly, 1983). Spatially detailed records of past climates which give information on the patterns of climatic change over large areas extend back only to the mid-nineteenth century, considerably less than the longest single-site records. The earliest circulation maps based on instrumental data are those of Lamb and Johnson (1966) back to 1750, maps of monthly-mean pressure for Januarys and Julys. Complementing these charts are time series of digitized pressure data extending from the present-day back into the late nineteenth century. For these series the coverage includes much of the Northern Hemisphere back to the beginning of this century. The area covered reduces as one goes back to the beginning of the continuous digitized data series in 1873. For the first few decades the mapped data cover only the eastern-Atlantic-European sector. Most of these data are available in grid point form on daily and monthly time scales. There are, however, some reservations regarding the quality of the data. Williams and van Loon (1976) and Trenberth and Paolino (1980) have noted errors and inhomogeneities in data for certain regions and it is generally thought that high northern latitude pressure values are too high in the early years when meteorologists believed that there was a semi-permanent Arctic high (Namias, 1958). 1873 marks the beginning of the digitized data. Before this, undigitized maps for the British Isles region exist covering the period 1861-1872. The quality of these maps is probably poorer than the more recent data. Prior to 1861 the only circulation pattern data available are the monthly mean maps of Lamb and Johnson (1966) and the daily maps constructed by Kington (1980) for 1781-86.

The purpose of this paper is to add to the existing circulation data using the longest available homogenized station records of pressure, temperature and precipitation. To construct circulation maps from these data by hand would be a formidable task, a task begun by Lamb and Johnson (1966) but never completed. The procedure used by Lamb and Johnson (see later) was similar to that employed by early meteorologists in constructing synoptic charts in areas of sparse data coverage. Instead we have used an objective multivariate regression technique. We have constructed monthly mean maps back to 1780 spanning the region 35-70°N, 30°W-40°E. The method is similar to that used in multivariate climate reconstruction work (reviewed by Webb and Clarke, 1977). In particular, the procedure parallels the pressure pattern reconstructions of Fritts and co-workers (e.g. Fritts et al, 1979; Fritts et al, 1981) based on tree-ring data. The method we employ is most similar to that used by Kutzbach and Guetter (1980) in their analysis of the potential limitations of such tree-ring-climate reconstructions.

2. METHOD

The basic technique is to develop a series of regression equations relating grid-point surface pressure patterns over the whole study region as predictands to the available station values of pressure, temperature and precipitation as predictors. Because the relationships between temperature (or precipitation) and the circulation vary seasonally (for example easterly winds in western Europe may be associated with cold weather in winter and warm weather in summer), separate equations must be developed for each month. Furthermore, since the network of available data varies with time new equations must be developed whenever the network changes significantly. We have simplified this latter problem by using only three different predictor networks.

The matrix of grid-point pressure data (i.e. the pressure network), Y_{ij} ($i = 1,...I$ grid points; $j = 1,...J$ years), is first transformed into principal components (PCs) using the correlation matrix:

$$\beta_{kj} = \sum_{i=1}^{I} v_{ki}\, y_{ij} \qquad (k = 1,...I;\ j = 1,...J)$$

(in this case I = 60 and J = 75; the models were calibrated using data for the period 1900-1974 which we refer to later as the calibration period). Similarly, the chosen station data of pressure, temperature and precipitation (i.e. the climate network), X_{nj} ($n = 1,...N$ sites; $j = 1,...J$ years), are also transformed into principal components:

$$\alpha_{mj} = \sum_{n=1}^{N} u_{mn} x_{nj} \qquad (m = 1,...N;\ j = 1,...J)$$

Here, lower case x and y denote normalized variables, α and β are the principal component amplitudes and u and v are the loadings.

In principal components analysis (PCA) the original intercorrelated data series are rearranged into linear combinations (the principal components) each representing decreasing amounts of the total variance of the original data. Generally, a high proportion of the variance can be represented by a relatively small number of PCs and the higher-order components can be considered as unwanted noise which may be eliminated by omitting these components from the analysis. There are many methods available for deciding how many PCs to retain (see, e.g. Preisendorfer et al, 1981). We have used the PVP method devised by Guiot (1981), a method which we have found to be satisfactory in other work (Briffa et al, 1983). In this method only those components for which the cumulative eigenvalue product exceeds 1 are retained. Usually between one and two-thirds of the PCs are retained by this criterion. The amount of retained variance is generally over 95%. In this particular application we begin with I = 60 grid-point pressure PCs and retain I^* of these where I^* varies from 14 to 20 depending on the month (fewer in the winter months). On the predictor side we begin with N components (N = 25 or 32 depending on the network) and retain N^*, where N^* varies between 10 and 18.

Each predictand component is then regressed against the N^* retained climate PCs which form a candidate set of predictors. This analysis is simplified because the candidate predictors are orthogonal and techniques like stepwise regression are unnecessary. We then eliminate some of the candidate predictors on the basis of their significance in the regression analysis to obtain a final set of N^{**} predictors where N^{**} may vary from predictand to predictand. There is no universally accepted criterion for eliminating candidate predictors at this stage. For the present analysis we choose to reject candidate predictors with t values less than 1 (see Briffa et al, 1983, for a more detailed discussion).

For each month and each of the three climate station networks we produce a set of I^* regression equations. These are then recombined to obtain the original grid-point pressure values. The overall regression equation may be written as ($i = 1,...I;\ j = 1,...J$):

$$\hat{y}_{ij} = \sum_{k=1}^{I^*} \sum_{m=1}^{N^*} \sum_{n=1}^{N^*} v_{ik} b^*_{km} u_{mn} x_{nj}$$

where b^*_{km} are the regression coefficients relating β_{kj} and α_{mj} (with the * used to indicate that N^*-N^{**} of these have been set equal to zero).

These equations (one for each of the I grid points) are then used for reconstructing grid-point pressure data back in time prior to the calibration period (1900-1974). As with any multivariate regression technique, the reliability of the regression equations cannot be determined solely from the calibration performance. In this present case the models are all highly statistically significant, but verification over an independent period is still essential. Performance can be assessed in three ways: the principal components can be tested using the individual equations for each component; results can be examined at each grid point separately; or real and reconstructed data can be compared over the whole grid month by month by examining spatial correlations. Results and model performance are discussed below. We first describe the data used in our analysis.

3. DATA

Over Europe there are many sites where records of air temperature and precipitation extend back on a monthly time base to the first half of the eighteenth century. A significant fraction of these data has lain dormant in meteorological and other archives for many years and is not in readily available sources (such as World Weather Records). Recently, much of the precipitation data for western Europe has been critically assessed and long homogeneous series produced for at least 200 sites: see Tabony (1980), Jones (1981, 1983) and Craddock and Wales-Smith (1977). Similar amounts of air temperature data are also available and those used here have all been critically examined and homogenized.

For the study area 11 long precipitation records and 12 long air temperature records were selected, the selection being based on length of record and/or location. At certain sites in Europe pressure data have also been recorded for long and relatively continuous periods. For the present study 32 sites of continuous surface pressure records were collected. All 32 have records back to at least 1865. The locations of the pressure, temperature and precipitation sites are shown in Fig. 1.

As with the homogenization of air temperature and precipitation data, corrections must be made to pressure records for changes in station location and/or instruments. However, with pressure, homogenization is further complicated by the necessary corrections of pressure to $0^{\circ}C$ and to gravity at $45^{\circ}N$, and the reduction of data to mean sea level. In certain cases because individual daily or twice-daily readings were not corrected for temperature at the time, the correction of monthly data for temperature has been made using the long term mean temperature for the station. This introduces a possible maximum error of ±0.5 mb in some of the early station data, with the largest possible errors occurring at those stations which are at highest elevations.

Reconstruction of monthly surface pressure charts from air temperature, precipitation and pressure station data should logically be better when more station pressure data are used in the analysis. To make best use of the available station pressure data, a new regression equation

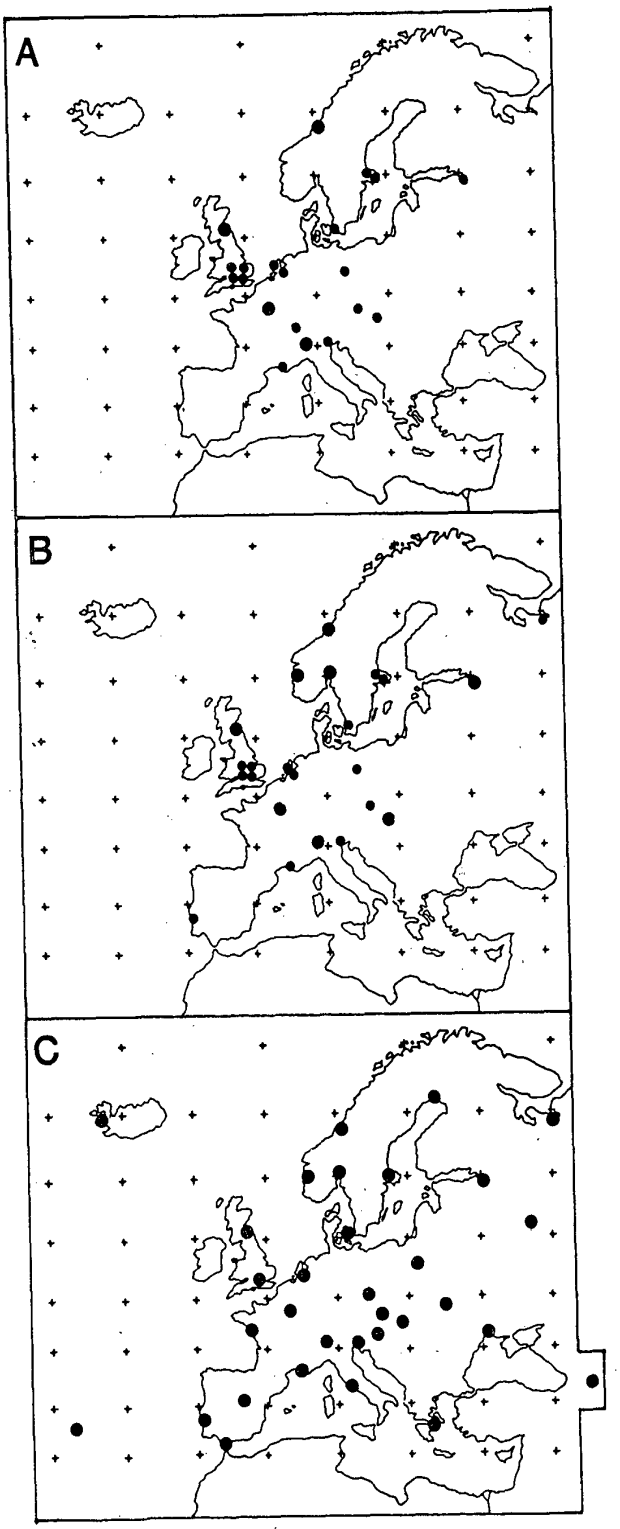


Fig. 1. Station networks used in the regression analyses: A, 1780-1836; B, 1837-1864; C, 1865-1872. Large dots denote sites with pressure data, small dots are sites with precipitation and/or temperature data. For network C only pressure data were used. In networks A and B most of the pressure sites were also temperature and/or precipitation data sites.

should be derived for every change in the pressure station network. However, minor network changes have no significant effect on the results and we judged that only three different networks were necessary at this stage. These were: climate network A back to 1780, 11 temperature stations, 10 precipitation stations and 4 pressure sites; climate network B back to 1837, 12 temperature stations, 11 precipitation stations and 9 pressure sites; climate network C back to 1865, 32 pressure sites (see Fig. 1).

The grid point predictand pressure data were taken from the U.K. Meteorological Office (UKMO) data set. This data set has been compared with a similar version produced by NCAR by Williams and van Loon (1976). An analysis of the NCAR data set has been undertaken by Trenberth and Paolino (1980) in which they compared gridded data with neighbouring monthly station data. Both sets of workers found both random and systematic errors in both data sets particularly for the Asian highlands, the Rockies and the Arctic Ocean. However, for the study area (35°N-70°N, 30°W-40°E) the UKMO data set is of high quality and these data were used in our analysis.

The data are gridded onto a 5° latitude by 10° longitude grid and our predictand network used 60 grid points (20° longitude spacing used at 70°N). Gridded data in the UKMO data set extend back to 1873 for this region giving 27 years (1873-1899) for verification. It was not possible to use all the data up to the present (1975-1981) as some of the recent station data are not readily available.

4. RESULTS

We have produced a set of 1116 pressure maps extending the monthly-mean coverage back from 1873 to 1780. Some examples are shown later. It is impossible to describe all of the results here. Even the preliminary principal components analyses are of interest in their own right. We will restrict ourselves to a discussion of regression model performance and the reconstructions.

Overall model performance can be assessed by correlating data at all grid points simultaneously. This allows us to calculate the overall explained variances month by month for each of the three predictor networks over both the calibration and verification periods. The results are shown in Fig. 2. As expected, there is some loss of variance between calibration and verification: although it should be noted that the verification data are themselves of lower quality than the post-1900 data and some of the variance loss could be attributed to this. Model performance is generally poorest in the summer months, but even in the worst case (June, network A) the regression accounts for 28% of the variance over the whole study region in an independent period. In the centre of the region (i.e. within the area in which the predictor network is located) performance is much better.

We have already noted that model performance can be assessed overall or at individual grid points. Following a similar study by Fritts et al, (1976) we have examined the spatial character of model performance by producing maps of the correl-

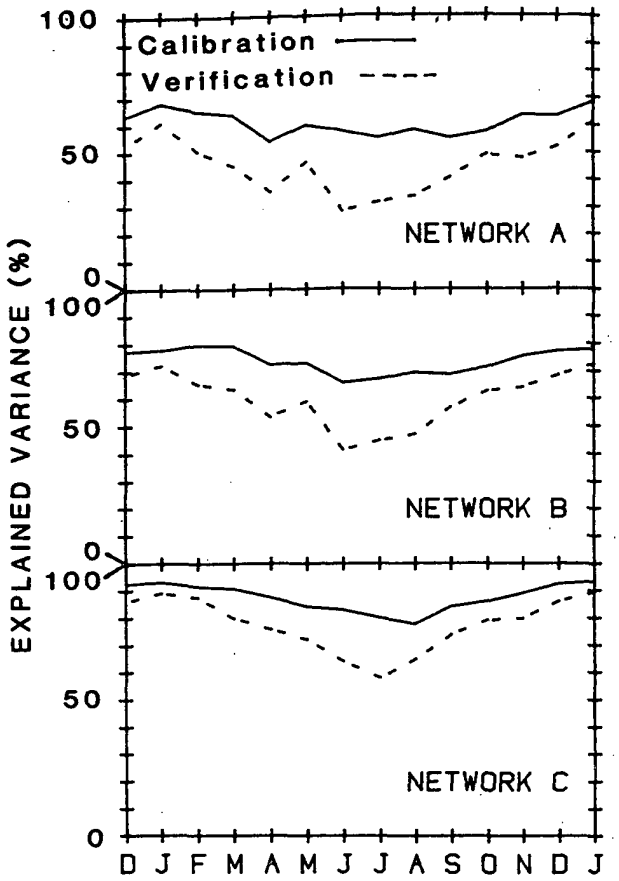


Fig. 2. Variations in explained variance (%) over calibration (full lines) and verification periods (dashed lines).

ations between real and reconstructed data over both the calibration and verification periods (one correlation map for each month and each climate network). We show here, in Fig. 3, only the best (Januarys, network C) and worst (Junes, network A) of these.

Even in the worst case the reconstructions are excellent in the centre of the study region and more than 50% of the variance is explained over a substantial part of the study region in the independent verification period. The poorer performance in the summer months can be seen to be due to effects around the edge of the study region, and it is apparent that the models' ability to extrapolate beyond the data area is considerably better in the winter months.

Poor model performance, when it occurs, is related to a failure to predict the behaviour of the higher-order pressure principal components. In Fig. 4 we show calibration and verification correlation coefficients for each PC down to PC10 for the best and worst cases (Januarys, network C; Junes, network A) and for two other cases (Junes, C; Januarys, A). For network C, the PC equations show no significant loss of explained variance in verification down to PC7 or 8. For network A, calibration performance is generally poorer for all PCs and the number of verifiable components is less (only 5 in the June example shown here). The network A cases generally show a loss of variance

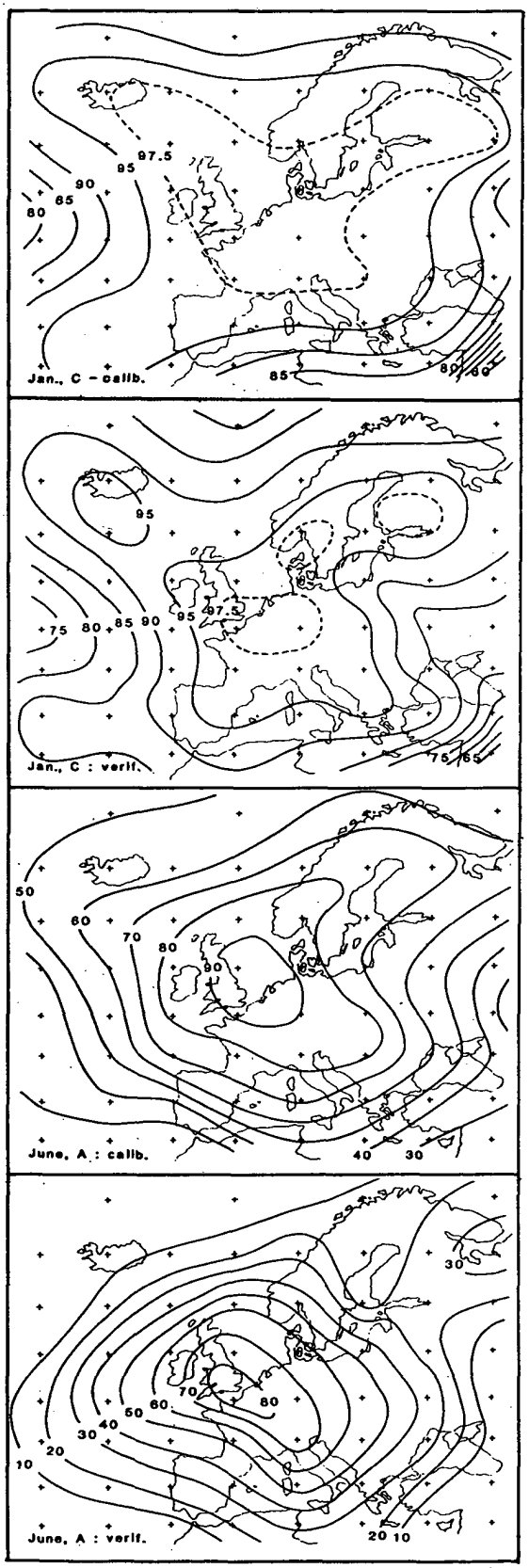


Fig. 3. Spatial patterns of explained variance (%) for calibration (1900-74) and verification (1873-1899) for the best (Januarys, network C) and worst (Junes, network A) of the 36 models.

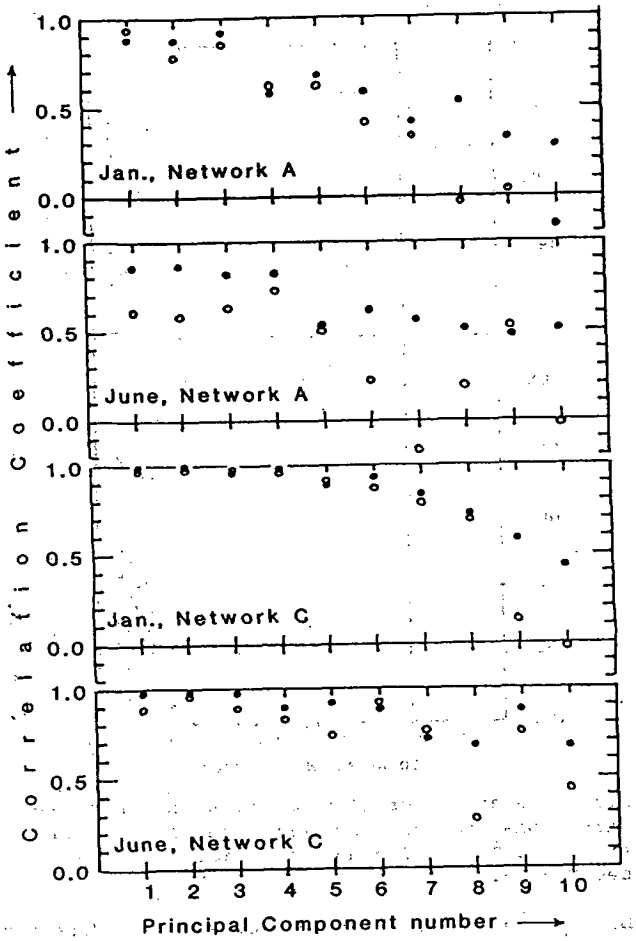


Fig. 4. Calibration (full circle) and verification (open circle) correlation coefficients between real and reconstructed data for individual grid-point-pressure principal components.

in verification even for the low order components. There is some indication in these results that we have retained too many PCs at the preliminary screening stage and that the PVP procedure that we used is not stringent enough in this particular application.

5. COMPARISON WITH OTHER RESULTS

Although there is some loss of variance in verification all models perform well and reproduce enough of the variance of the real data during verification to be judged successful. For model C (back to 1865, the year in which pressure data first became available in the Azores), given the known limitations of the late-19th-century grid-point pressure data set, our reconstructed values are probably as good as the real values. Prior to 1865 the reconstructions make use of almost all of the available early pressure data and employ temperature and precipitation data in an objective and statistically rigorous way. Given that all the data employed have been carefully examined for errors and homogeneity we judge the reconstructions to be close to the optimum that can be obtained using the available data.

Some of our reconstructions can be compared with the independently-produced circulation maps for the European region of Lamb and Johnson (1966) and Kington (1980). Note that these analyses, and our own, are not really independent in terms of data, since they have many data sources in common, but they are all independent in terms of the method of construction. Each method has its limitations and it is not possible to say that any one set of results is better than the others.

Kington (1980) has produced daily charts for the period 1781-86 using the available instrumental data for this period. His charts, which are quite detailed, cover only part of our study region since he (wisely) did not attempt to extrapolate far outside the area of data coverage. Kington's charts generally cover the region 45-65°N, 10°W-30°E and have between 20 and 35 grid points in common with our analyses. To facilitate the comparison we digitized his daily charts for the period 1781-85 and averaged them to produce monthly grid-point data. We then calculated the spatial correlations with our network A reconstructions month by month (see Fig. 5). The overall agreement is high (amounting to 64.4% explained variance). There is no significant difference in performance between seasons (contrary to the early verification results, but supporting our earlier suggestion that the poorer performance in the summer months was the result of larger errors around the periphery of the study region). There are, however, occasional months (see Fig. 5) when the correlation between Kington's maps and our reconstructions is extremely poor. These instances require further study. They most probably highlight particular months when the

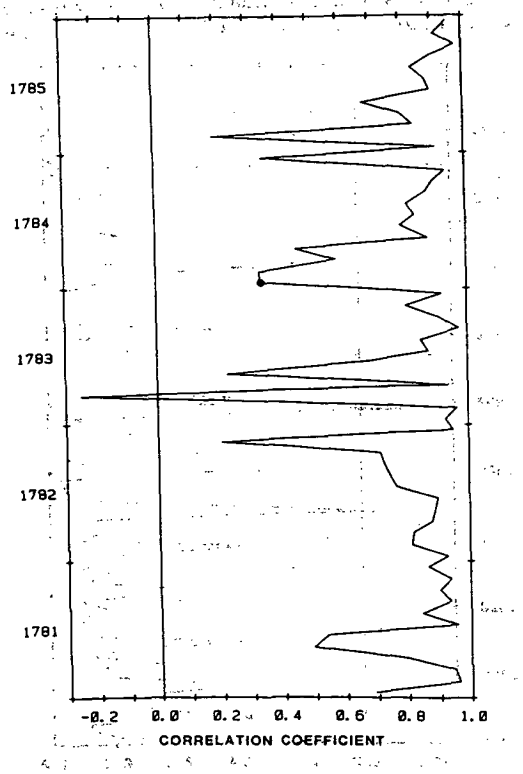


Fig. 5. Spatial correlations between our reconstructions (using network A) and monthly mean pressure patterns derived from Kington's daily charts, Jan. 1781 to Dec. 1785. The dot identifies Jan. 1784; see Fig. 8.

temperature and precipitation patterns show anomalous links with the prevailing circulation. This is a perennial problem in long-range forecasting and it is well known that subjective long-range predictions of temperature and precipitation based on a forecast surface pressure map can occasionally (amd embarrassingly) be grossly in error even when the pressure forecast is correct.

Lamb and Johnson (1966) have constructed monthly surface pressure charts for Januarys and Julys back to 1750 by plotting and subjectively contouring monthly station sea level pressure data. Other relevant information such as wind data, temperature and precipitation records and descriptive material from diaries were also employed in their analyses for the eighteenth and early nineteenth centuries. Their charts extend to near-global coverage during the present century, but for the earliest period the coverage is confined to the eastern-Atlantic-European region. For the period 1780 onwards Lamb and Johnson's charts cover all of our study region. The charts have been digitized by the U.K. Meteorological Office from the original working charts held at UKMO. Working charts were also prepared for April and October but never published (H.H. Lamb, personal communication).

As with Kington's data the spatial correlations between the Lamb and Johnson charts and our own reconstructions have been plotted as time series for Januarys (Fig. 6) and Julys (Fig. 7). The agreement is generally excellent, although there are occasional Januarys in the early years for which the match is poor. Agreement tends to

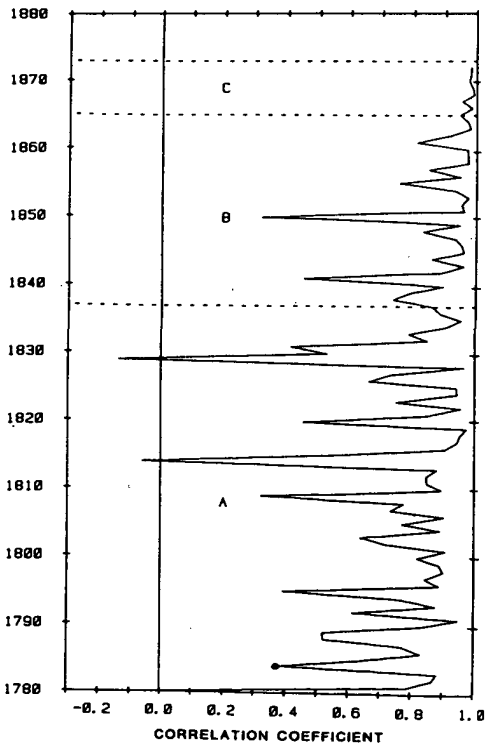


Fig. 6. Spatial correlations between our reconstructions and Lamb and Johnson's maps for Januarys from 1780 to 1872. The dot identifies Jan. 1784; see Fig. 8.

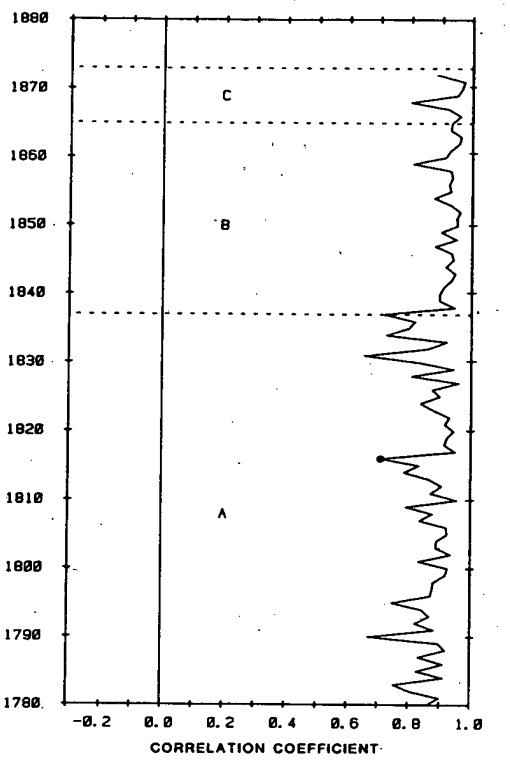


Fig. 7. Spatial correlations between our reconstructions and Lamb and Johnson's maps for Julys from 1780-1872. The dot identifies July 1816; see Fig. 8.

be less good in the earlier years as the data coverage deteriorates. This is largely due to increasingly disparate results in the data-poor area of extrapolation around the edge of the study region, although even here agreement is often very good.

The average squared correlations between our reconstructions and Lamb and Johnson's charts are given in Table 1 using all available years for each of the three climate networks. To put these values in context, Table 1 also shows the strength of the relationships between our reconstructions and the real data over the calibration and verification periods.

	network A		network B		network C	
	Jan.	July	Jan.	July	Jan.	July
Calibration period 1900-1974	0.67	0.50	0.75	0.63	0.90	0.75
Verification period 1873-1899	0.55	0.23	0.64	0.32	0.84	0.45
Lamb and Johnson 1780-1872 (A) 1837-1872 (B) 1865-1872 (C)	0.64	0.77	0.79	0.84	0.95	0.86

Table 1. Overall common variance (R^2) between the reconstructions and real data and between the reconstructions and Lamb and Johnson's pressure maps.

The overall agreement between our reconstructions and Lamb and Johnson's maps is excellent. In the early years the agreement is similar

to that between our results and Kington's maps (viz. 64.4% variance in common over the whole year). The correlations between our reconstructions and the Lamb and Johnson charts are higher than the reconstruction-real data correlations over both verification and calibration periods.

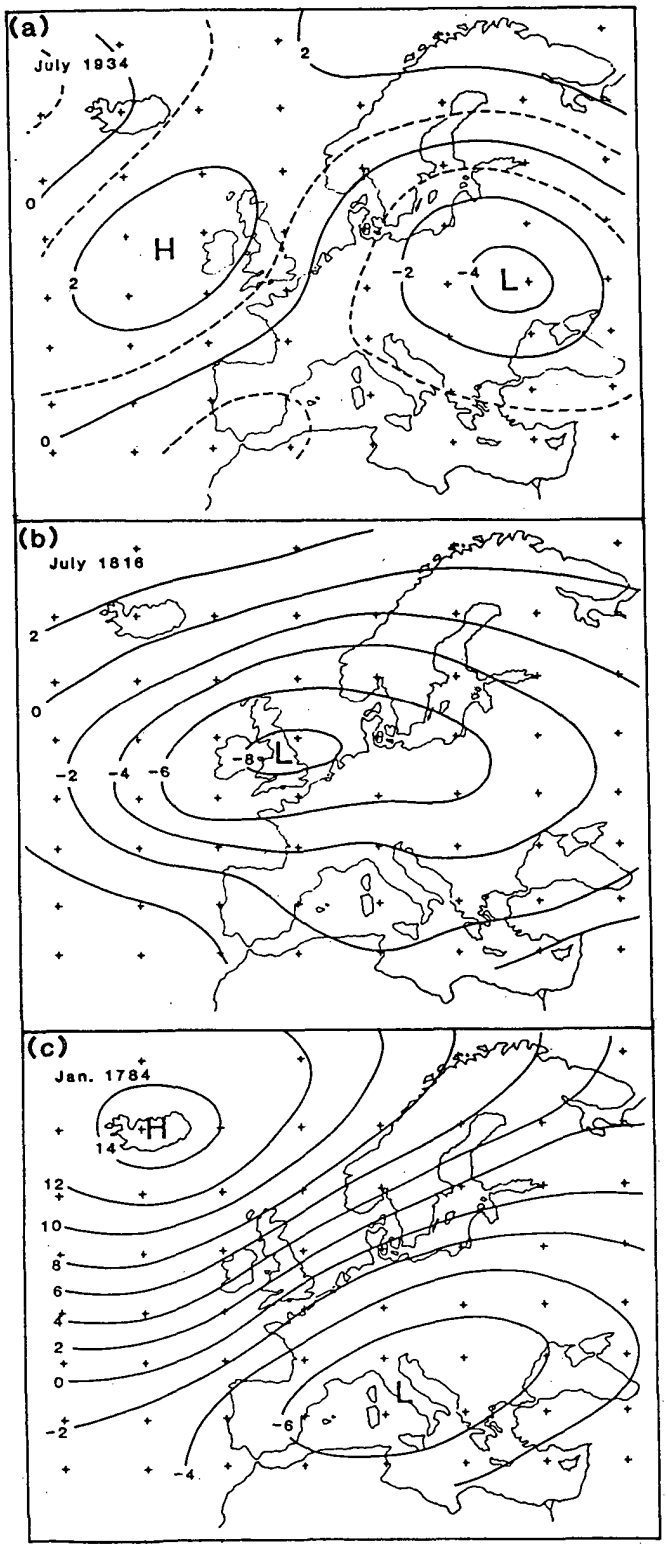


Fig. 8. Three reconstruction examples, July 1934, July 1816 and Jan. 1784. Pressures are anomalies in mb from the 1900-74 mean.

Furthermore, unlike the calibration and verification results there is no reduction in the strength of the link in the summer months. Since Lamb and Johnson's charts and our own reconstructions are based on a similar network of station data it is apparent that each achieves the same ends and makes the same sort of errors in extrapolation. More importantly, it appears that the objective statistical relationships derived by our regression analysis must be similar to the subjective relationships, based on years of experience, used by Lamb and Johnson.

6. CONCLUSIONS

Having described the general character of our reconstructions and tested and demonstrated their validity we now present three examples, pressure anomaly maps for July 1934, July 1816 and January 1784. These three were chosen because they illustrate particularly anomalous circulation patterns. Summer 1934 (Fig. 8a) was one of the most severe agricultural droughts in the past three centuries in England (Wigley and Atkinson, 1977) with July rainfall considerably below average. The prevailing circulation was an anomalous blocking situation. The correlation coefficient between our reconstruction (using network C) and the real grid-point data is 0.89. July 1816 was a central month in the well-known "year without a summer" (Landsberg and Albert, 1974). July precipitation was well above average, one of the ten wettest Julys in England and Wales in the past 200 years (Wigley et al, 1983). The pressure distribution for this month is given in Lamb and Johnson's work and shows a strong cyclonic anomaly over the centre of our study region. Our reconstruction (Fig. 8b) gives the same result and is well correlated with Lamb and Johnson's map ($r = 0.71$). From Fig. 7 one can see that 1816 was, in fact, one of the worst Julys in terms of the correlation between our reconstructions and Lamb and Johnson's maps. January 1784 (Fig. 8c) was a particularly cold month in a famous cold winter, a winter in which the River Thames froze over sufficiently solidly for a frost fair to be held on the ice. From Kington's maps and those of Lamb and Johnson we know that the month was one of predominant easterly winds over western Europe and a characteristic winter blocking pattern. Our reconstruction shows this well even though it is not highly correlated with Kington's data ($r = 0.34$) nor with Lamb and Johnson's data ($r = 0.37$). From Figs. 5 and 6 it can be seen that January 1784 is one of the poorest months in terms of the correlation with these other circulation maps.

Minor improvements to our reconstructions are possible and we plan to revise the results before publishing the complete map series. For some periods short data series can be used to supplement the existing continuous network, and a substantial improvement may be possible back to about 1850 since we have not made use of all available pressure data for the period 1850-1864. One simple method of improvement is simply to average the monthly maps into seasonal maps. On a month-to-month basis the data are not autocorrelated so the standard errors should be reduced by a factor of about $\sqrt{3}$ in this way. Combining monthly data into decades should also reduce the uncertainties and provide an extremely reliable indicator of

circulation changes on this longer time scale.

The next stage, and part of a long term plan, is to extend our analysis to other parts of the world. Considerable extensions of the available digitized grid point pressure data will be possible, both spatially and backwards in time. The method can also be used to help to fill in data-poor areas in the more recent records and to produce climatologically consistent and objective corrections for the errors identified in existing data sets by Williams and van Loon (1976) and Trenberth and Paolino (1980).

Two other areas deserve further study. Firstly, we have already noted that, in isolated months, our regression models perform rather poorly. The models hinge on the physically-based and statistically strong relationships between pressure patterns and temperature and precipitation on a monthly time scale, and an analysis of instances where these relationships break down may give useful insight into the physical processes involved. Secondly, we have found that our statistical extrapolation procedure works better in the winter months implying, in the European region at least, that teleconnections are strongest at this time of the year. Our technique provides a basis for examining teleconnections in a multivariate framework over much larger regions.

Finally, our results have some important implications for multivariate climate reconstruction work based on proxy data such as tree rings. Following Fritts et al (1976) and Kutzbach and Guetter (1980), pressure pattern reconstructions using real temperature and precipitation data should provide an upper limit to the accuracy with which one can hope to reconstruct pressure patterns from tree rings, assuming a similar predictor network. It is unfortunate that our reconstructions are worst in the summer months, since this is the time of the year when trees in Europe respond best to climate. Nevertheless, even at this time of the year our reconstructions are excellent in the immediate vicinity of the climate data network.

Acknowledgements This research was supported in part by the U.S. Department of Energy, Carbon Dioxide Research Division under Contracts DE-AC02-79EV10098 and DE-AC02-81EV10739 and by the U.K. Natural Environment Research Council Grant GR4/79/AAPS/12.

7. REFERENCES

Briffa, K.R., P.D. Jones, T.M.L. Wigley, J.R. Pilcher and M.G.L. Baillie, 1983: Climate reconstruction from tree rings: Pt. 1, Basic methodology and preliminary results for England. J. Climatology, in press.

Craddock, J.M., and B.G. Wales-Smith, 1977: Monthly rainfall totals representing the East Midlands for the years 1726 to 1975. Met. Mag., 106, 97-111.

Fritts, H.C., T.J. Blasing and G.R. Lofgren, 1976: Climatic variations reconstructed from tree rings. Paper presented at the Centennial Academic Assembly entitled "Man and Nature - Makers of Climatic Variations", Sept. 30-Oct. 1, Texas A&M Univ., College Station, Texas.

Fritts, H.C., G.R. Lofgren and G.A. Gordon, 1979: Variations in climate since 1602 as reconstructed from tree rings. Quat. Res., 12, 18-46.

Fritts, H.C., G.R. Lofgren and G.A. Gordon, 1981: Reconstructing seasonal to century time scale variations in climate from tree ring evidence. In, Climate and History (T.M.L. Wigley, M.J. Ingram and G. Farmer, eds.) Cambridge University Press, 139-161.

Guiot, J., 1981: Analyse mathematique de donneés geophysique. Applications à la dendroclimatologie. Unpublished Ph.D. dissertation, Université Catholique de Louvain, Faculte des Sciences, Louvain-la-Neuve, Belgium.

Jones, P.D., 1981: A survey of rainfall recording in two regions of the Northern Pennines. Met. Mag., 110, 239-252.

Jones, P.D., 1983: Further composite rainfall records for the United Kingdom. Met. Mag., 112, 19-27.

Jones, P.D., and P.M. Kelly, 1983: The spatial and temporal characteristics of Northern Hemisphere surface air temperature variations. J. Climatology, in press.

Kington, J.A., 1980: Daily weather mapping from 1781. Climatic Change, 3, 7-36.

Kutzbach, J.E., and P.J. Guetter, 1980: The design of palaeoenvironmental data networks for estimating large-scale patterns of climate. Quat. Res., 14, 169-187.

Lamb, H.H., and A.I. Johnson, 1966: Secular variations of the atmospheric circulation since 1750. Geophys. Mem., 110. London (HMSO., for Met. Office).

Landsberg, H.E., and J.M. Albert, 1974: The summer of 1816 and volcanism. Weatherwise, 27, 63-66.

Manley, G., 1974: Central England temperatures: monthly means 1659-1973. Quart. J. Roy. Meteor. Soc., 100, 389-405.

Namias, J., 1958: The general circulation of the lower troposphere over Arctic regions and its relation to the circulation elsewhere. In, Polar Atmosphere Symposium, Pt. 1, Meteorology, Pergamon Press, 45-61.

Preisendorfer, R.W., F.W. Zwiers and T.P. Barnett, 1981: Foundations of principal component selection rules, SIO Ref. Series 81-4, Scripps Institution of Oceanography, La Jolla, Calif.

Tabony, R., 1980: The homogenization and analysis of European rainfall records. U.K. Met. Office, M.O. 13, Branch Memo. 76, Bracknell, U.K.

Trenberth, K.A., and D.A. Paolino, 1980: The Northern Hemisphere sea-level pressure data set: trends, errors and discontinuities. Mon. Wea. Rev., 108, 855-872.

Webb, T., and D.R. Clarke, 1977: Calibrating micropaleontological data in climatic terms: a critical review. Annals N.Y. Acad. Sci., 288, 93-118.

Wigley, T.M.L., and T.C. Atkinson, 1977: Dry years in south-east England since 1698. Nature, 265, 431-434.

Wigley, T.M.L., J.M. Lough and P.D. Jones, 1983: Spatial patterns of precipitation in England and Wales and a revised, homogeneous England and Wales precipitation series. J. Climatology, in press.

Williams, J., and H. van Loon, 1976: An examination of the Northern Hemisphere sea-level pressure data set. Mon. Wea. Rev., 104, 1354-1361.

A MULTIPLE CLIMATIC REDUCTION TECHNIQUE APPLIED TO THE EXTENSION OF PRECIPITATION SERIES.

Vicente R. Barros and Irma G.Schmidt

Centro Nacional Patagónico.
Consejo Nacional de Investigaciones Cietíficas
Puerto Madryn, 9120, Chubut, Argentina.

1. INTRODUCCION

It is fairly usual that when there is an specific interest for climatic information in certain sites, meteorological records are not long enough to estimate the required statistical parameters whith a desired level of confidence. This is particularly true in those regions where there is not a long technological tradition. Anyway, the urgency for taking a decision, usually press the application of techniques wich make use of only the available information.

A methodology which successfully extends in time some meterological series is a valuable tool for both the recovering of missing data and for the estimation of long term statistics. Such methods may allow sensible reductions in the required measuring time (in a specific site) to periods of time compatible with the nature of the request, being this either scientific, social o economic in its origin.

In this paper, the use of a Multiple Climatic Reduction Technique (MCRT) for extending in time, short series of precipitation in a region of South Argentina is explored. The MCRT is based on the series representation through the eigenvectors of the covarience matrix of the series of the region. The multipliers of the series to be extended are calculated by minimization of the RMS error between data of the short series and its representation. Once the multipliers are calculated, the series may be virtually extended using the respective eigenvectors.

This technique, which is described in section 2.1 does not guarantee the extended series will be exactly the true one or differ from it by a small error. The parameters and their spatial and temporal scales to which this technique is applicable as well as the degree of accuracy to be obtained are points which require further studies. In connection with it, this paper shows the use of MCRT for extending in time precipitation series in a mountainous region where this parameter is quite variable,both in space and time.

2. MCRT USE FOR THE EXTENSION OF SHORT PRECIPITATION SERIES.

2.1 The method

This technique was described by Barros and Rodriguez Seró (1981), and used to estimate annual wind variation, mean wind speed and meteorlogical wind power, based in only a few months of data (Barros and Estevan, 1983). It will be briefly reviewed here as follows. Let V be a matrix whose columns are m data series of a parameter in a given region. It is always possible to write.

$$V=AW \qquad (1)$$

Where A is the orthonormal matrix of the eigenvectors of VV^t and $W=A^tV$ is formed by the projections of data series over each eigenvector. Therefore a datum at location j and time i is given by

$$v_{ij}= \sum_{k=1}^{r} a_{ik}\, w_{kj} \quad ,i=1...r, \; j=1...m \qquad (2)$$

r is the number of items in the series and k refers to the eigenvector number.

If the series, as it may be true for many meteorological parameters of a given region are not all statiscally independent, a reasonably accurate representation may be achieved by

$$v_{ij} = \sum_{k=1}^{n} a_{ik}\, w_{kj} \quad , \; i=1...r \;\; j=1...m \qquad (3)$$

Where $n < r$

When for a particular series $j=q$ in the region, there is only a short record O_{iq}, whit $i = L...p$ and $L > 1$ and $p < r$, O_{iq} may also be expressed by (3). Though O_{iq} has not been used for the calculation of the eigenvector matrix A, the values of w_{kq} may be found by minimizing

$$F_q = \sum_{i=L}^{p} (O_{iq} - \sum_{k=1}^{n} a_{ik} w_{kq})^2 \quad (4)$$

This leads to a linear system of N-equations with N-unknowns: w_{kq}, $k=1...n$. This system may be easily solved by standard methods, (see Forsythe and Moler, 1967). Therefore using model (3) the series are virtually extended to the range $i=1...r$

2.2 Extension of precipitation series using MCRT

Fig. 1 shows the location of the only 16 observing stations where complete precipitation series of the period 1972 - 1981 were available to the authors. The region is in the southwest of Argentina between 37°and 42°south and from 68° west up to the argentine-chilean border. Annual mean precipitation on this set of stations goes from 180 mm in the east up to 2000 mm in the west (Table 1), but precipitation exceeding 3000 mm may be expected in same locations in the Andes area where strong horizontal gradients are observed due to complex an steep terrains. Except in the arid zone, precipitation is considerably more important in fall and winter than in summer and spring.

In order to test the MCRT, each of the 16 series was assumed unknown in part of the ten year period and reconstructed according to what has been explained in section 2.1

The elements of the series were the precipitation of each season of the ten year period of 1972-81, so the matrix V was of 40 X 15 since it should be mentioned that for each reconstruction of a series, this one was not considered in the calculation of A.

Only 5 eigenvectors were used since they accounted for 94% of the variance.

In each case the assumed known period of the series was chosen in 4 to 8 years at the begining and end of the ten year period. For simplicity, results of the 4 years case are only presente here.

Tables 2 and 3 show correlation coefficients, percentage reduction of variance, and mean relative error resulting from comparison between the true observed series and the simulated through MCRT. It may be seen that correlations

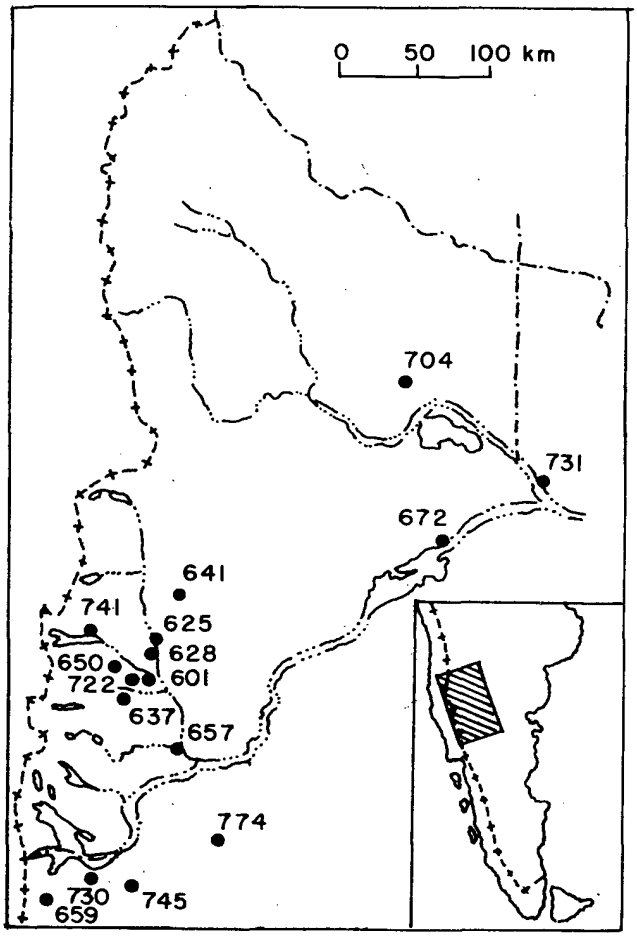


Fig. 1. Location of the stations used in this paper

Table 1. Mean precipitation (mm). Station numbers according to Centro Nacional Patagonico numeration.

Station	Annual	Summer	Fall	Winter	Spring
601	597	55	267	209	66
625	522	47	225	197	59
628	497	53	204	188	56
637	672	60	272	269	71
641	530	48	242	184	56
657	270	55	108	74	33
659	2006	232	776	613	384
672	201	74	57	31	39
680	912	75	386	355	96
704	181	35	66	16	34
722	648	60	285	234	69
730	1355	128	585	511	131
731	255	68	65	45	56
741	1235	127	528	432	148
745	729	69	318	262	79
774	250	52	99	61	37

coefficients are near 0.9 on the average.

Fig. 2 shows as an example the actual and reconstructed series of stations 680 and 657. This last one is representative of the average situation as may be seen from inspection of table 2. It is worth -

Table 2. Statistics of reconstructed series with MCRT based on the first 4 years of the 1972-81 period as compared to actual data

$$Z_{pp}^2 = (1 - 1/N \frac{\sum_{i=1}^{N} (PP_i - P_{ei})^2}{\sigma_{pi}^2})$$

with PP_i being the precipitation, P_{ei} the respective estimate, and σ_{pi}^2 the variance of PP_i

Station	Correlation Coefficient	$1 - Z_{pp}^2$	Mean relative error				
			Annual	Summer	Fall	Winter	Spring
601	.95	.09	.02	.02	.05	.04	.05
625	.97	.07	.08	.09	.09	.07	.05
628	.98	.04	.04	.06	.08	.01	.06
637	.97	.07	.02	.06	.09	.04	.01
641	.92	.15	.01	.23	.06	.00	.04
657	.87	.24	.03	.16	.09	.00	.29
659	.68	.68	.02	.09	.01	.07	.28
672	.73	.75	.04	.17	.04	.39	.16
680	.99	.02	.00	.03	.02	.05	.05
704	.69	.63	.11	.05	.21	.33	.32
722	.97	.05	.01	.02	.03	.01	.01
730	.93	.16	.00	.06	.01	.04	.11
731	.75	.60	.08	.12	.08	.21	.10
741	.96	.09	.00	.07	.00	.01	.02
745	.97	.05	.05	.08	.06	.01	.12
774	.80	.36	.02	.18	.00	.19	.06
Average	.88	.25	.03	.09	.06	.09	.11
Maximum	.99	.75	.11	.23	.21	.39	.29
Minimum	.68	.02	.00	.02	.00	.00	.01

Table 3 Same as Table 2 but for MCRT based on the last 4 years of the 1972-81 period.

Station	Correlation Coefficient	$1 - Z_{pp}^2$	Mean relative error				
			Annual	Summer	Fall	Winter	Spring
601	.96	.07	.06	.22	.02	.05	.24
625	.97	.06	.06	.22	.04	.03	.05
628	.98	.04	.00	.09	.05	.00	.05
637	.97	.06	.00	.01	.07	.06	.06
641	.93	.13	.02	.22	.08	.11	.00
657	.83	.40	.10	.23	.00	.04	.36
659	.76	.61	.03	.22	.07	.07	.35
672	.63	.67	.16	.46	.00	.10	.07
680	.99	.02	.01	.00	.03	.04	.08
704	.71	.49	.03	.02	.08	.10	.11
721	.97	.05	.03	.05	.01	.06	.04
730	.93	.13	.03	.07	.04	.15	.00
731	.79	.72	.03	.20	.18	.29	.36
741	.97	.06	.06	.23	.08	.02	.10
745	.98	.05	.02	.12	.02	.02	.17
774	.83	.41	.06	.19	.05	.04	.21
Average	.89	.25	.04	.16	.05	.07	.14
Maximum	.99	.72	.16	.46	.18	.29	.36
Minimum	.63	.02	.00	.00	.00	.00	.00

while to mention that the only value sustantially wrong in the reconstructed series is that of 1978 summer. This was due to an exceptional convective rain only registred at this station.Other exceptional precipitation was the very low registered in 1979 fall, and since this phenomenum was observed at other stations, the reconstruction fits the datum quite well. Station 680 is probably the best simulated of the set.

Series of the arid zone are poorly

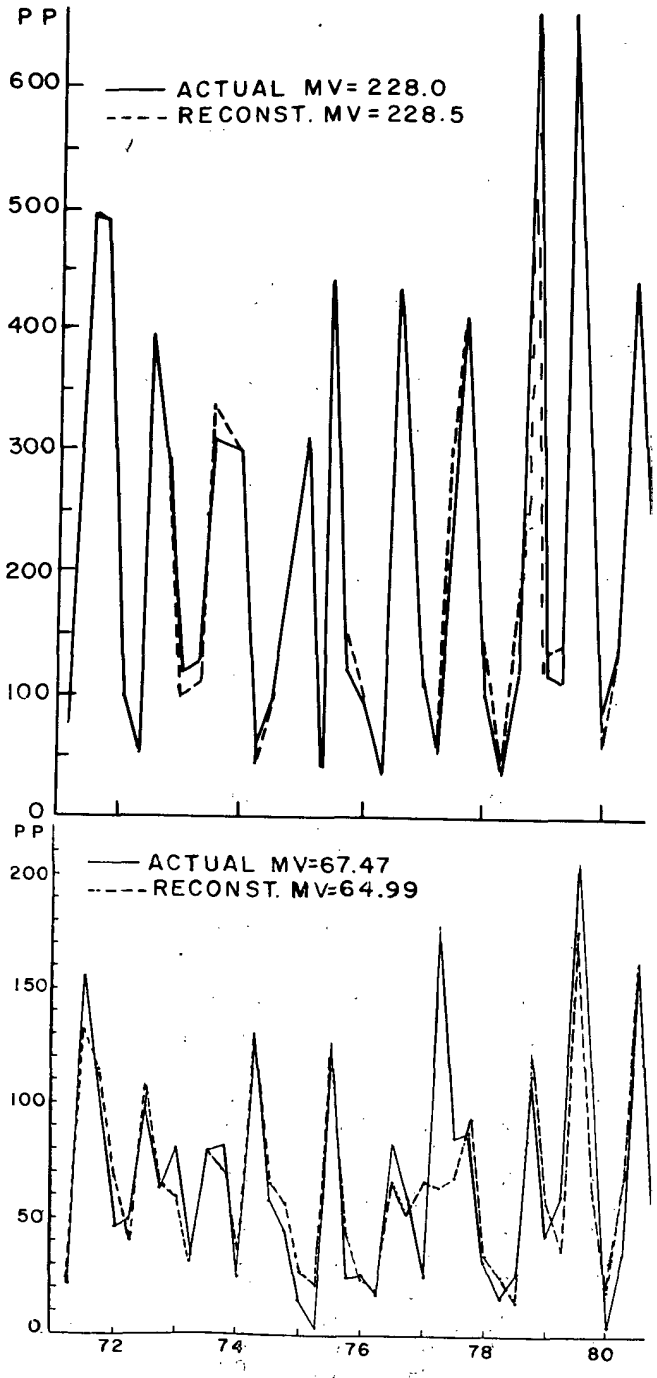


Fig. 2 Comparison of the actual and reconstructed series using MCRT and the first 4 years(station 657,botton) and the last 4 years (station 680, top).

simulated because no care has been taken in normalizing the series before the calculation of the eigenvectors. So, these are biased toward the higher values of the west region which has a precipitation regime different fron the arid zone. Anyway, the simulation of these series has a considerable improvement when more than four years are assumed known. In any case, The mean annual and seasonal values over the ten year period are calculated whith small errors. This is particularly true for the annual case where the average relative errors are .03 and .04, table 2 and 3.

Though some relative errors in the mean seasonal estimation are higher than .2, this happens only in those cases where the precipitation means are usually well below 100 mm, either because are spring or summer values or because belong to stations of the arid zone and so the absolute error is very small. There is one exception, station 659, which in all sense is not well reconstructed because is the only one situated to the west of the highest mountains and probably having a different precipitation regime.

3. COMPARISON WITH SIMPLER METHODS

Other techniques were used with the same data set in order to compare them with MCRT.They are the ratios method (Conrand an Pollack, 1962), the Corotis method (Corotis, 1979) and simple linear regression. In all this cases mean annual and seasonal values were estimated assuming known the last 4 years of the series and the 10 year average of the nearest station.

Table 4 shows average results for the whole set of 16 series. In the case of mean annual values, the MCRT does not show any improvement over the other simpler methods, but in the estimation of mean seasonal precipitation appears as the only method which gives acceptable results. This is so, conisidering that errors of the order of 15% in spring and summer mean very small absolute errors (table 1). In turn, the higher precipitations of fall and winter are estimated on the average with 5 and 7% error.

Table 4. Comparison of different methods for the estimation of mean precipitation. Average relative errors for the 16 stations.

	Annual	Summer	Fall	Winter	Spring
Ratios	.04	.18	.32	.40	.32
Corotis	.07	.15	.70	1.09	.24
Regression	.03	.20	1.03	.71	.37
MCRT	.05	.16	.05	.07	.14

4. RECOVERING OF THE W_k

It is fairly usual that multipliers w_k are correlated with some geographic features, like for instance station high. The drawing of w_k maps using these characteristics conduces to

more spatially detailed and precise precipitation maps. Therefore, it is interesting to check the degree of accuracy that MCRT may produce in the estimation of each multiplier w_k. If this estimation is done with small errors, it may allow the use of incomplete series in the drawing of w_k maps.

Fig. 3 shows the values of w_1 and w_2 calculated by MCRT as a function of the actual ones. The adjustment in the case of the first multiplier is indeed good, with small errors in relation to the range of variability of the parameter. For the second multiplier the fitting is still acceptable though errors are sensible higher and station 659 value is completely wrong.

The problem with station 659 remains for the other multipliers and as it was explained in 2.2. This station may be considered as pertaining to other precipitation regime region.

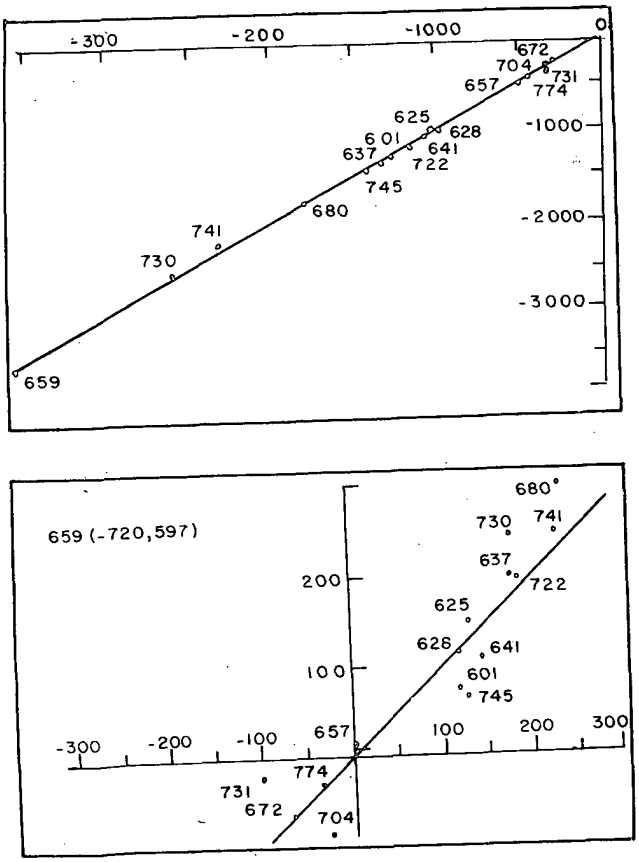


Fig. 3 Adjustment between actual values of M_k (in the ordinate) and evaluated through MCRT (abscissa). Figure on the top: first multiplier, figure on the bottom second multiplier

Table 5. shows a measure of the adjustment between the actual and re-constructed values of m_k when station 659 is excluded. Of course, results improve as the number of years assumed known increases.

Table 5. Percentage reduction Z^2_M for each multiplier M_{kq}, q=1..16

Multiplier number	4 Years	8 Years.
1	.99	.99
2	.84	.96
3	.40	.35
4	.45	.72
5	-1.87	- .27

It may be seen again that the fitting on the two first multipliers is good, and that in the case of the fifth, MCRT produces no information. Fortunately, in this case the contribution of the fifth eigenvector is negligible.

5. CONCLUSION

The MCRT has been shown to be able of extending in time precipitation series, at least in the studied region. Though mean annual values were not estimated better than with simpler techniques, mean seasonal estimates using MCRT were far superior than those of simpler methods.

Finnally, good estimations of the first multipliers were found. This may allow the use of incomplete series in the drawing of precipitation maps.

6. REFERENCES.

Barros, V.R, and E. Estevan, 1983: On the evaluation of wind power, from short wind records. To be published in the Journal of Applied Meteorology.

Barros, V. R. and J. A. Rodriguez Seró, 1981: Measurements strategies: Use of short observation records for estimating the annual wind variation. Proceedings of the International Colloquium on Wind Energy. BHRA Fluid Engineering. Brighton U.K., 23-28 pp (Available from BHRA Fluid Engineer

ing, Cranfield, Bedford, MK43 OAJ, U. K.)

Conrand, V. and L.W. Pollack, 1962 : Methods in Climatology. Harvard University Press 459 pp.

Corotis, R. B, 1979 : Statistical models for wind characteristics at poten - tial wind energy conversion sites. DOE/ ET/ 202831-1 Northwestern University, 86 pp.

Forsythe, G. E. and C.B. Moler, 1967 : Computer Solution of Linear Alge - braic Systems. Prentice Hall Inc., Englewood Cliffs, N. Y.

EOF SPECTRAL ESTIMATION IN CLIMATE ANALYSIS

Robert F. Cahalan

Climate and Radiation Branch
Goddard Laboratory for Atmospheric Sciences
NASA/Goddard Space Flight Center
Greenbelt, Maryland 20771

1. INTRODUCTION

EOF's and their associated variance spectra, the eigenfunctions and eigenvalues of the covariance matrix, respectively, are a property of the parent probability distribution of the climate, and thus are of great interest. Empirical estimates of them are based on a finite number, say n, independent realizations of the instantaneous state of a geophysical field during a limited time period, and the field is sampled at a finite number of points, say p, covering a limited spatial domain. Errors in these estimates depend on the true spectrum of the field, which determines whether p and n have been appropriately chosen.

The characteristic types of errors which occur in EOF estimation may be illustrated by considering two extreme cases for the true spectrum:

case A: All of the variances are equal ("measurement error")

case B: None of the variances are equal ("red noise")

In case A the covariance between any two points vanishes. Such "measurement error" can be simulated by generating Gaussian random variables with zero mean and unit variance independently at each point. One can show that the addition of such uncorrelated variations to any "true" climatic variations leaves the "true" EOF's and their spectrum undistorted in the limit $n \to \infty$. Anderson (1963) and others have derived large n corrections, but such results assume $n >> p$, and give incorrect results in geophysical applications (Preisendorfer and Barnett, 1977). I have derived a simple analytical formula for the distribution of sample eigenvalues in case A for an arbitrary ratio n/p (Cahalan, 1982). The mathematics required was developed more than 20 years ago in nuclear physics (Wigner, 1955 and 1957; Dyson, 1962). Although the problem is discussed in most multivariate statistics books, the analytical solution was not known. All of the results for case A follow from the probability density ρ as a function of the EOF variance (covariance eigenvalue) λ :

$$\rho(\lambda) = \mathrm{Re}\,[-\lambda^2 + 2(1+p/n)\,\lambda - (1-p/n)^2]^{1/2} / (2\pi\lambda\, p/n). \qquad (1)$$

Since F. J. Dyson first derived this in a different context, I refer to it as "Dyson's distribution". It is easy to show from (1) that the eigenvalues satisfy

$$\text{range} \qquad (1-\sqrt{p/n})^2 \leq \lambda \leq (1+\sqrt{p/n})^2, \qquad (2)$$

$$\text{mean} \qquad \overline{\lambda} = 1,$$

$$\text{variance} \qquad \overline{\lambda^2} - (\overline{\lambda})^2 = p/n.$$

The variance spectrum as a function of modenumber as well as the cumulative variance spectrum as a function of modenumber are both derived from Dyson's distribution. This distribution maximizes Wishart's joint probability distribution for large p and fixed p/n. The Gaussian assumption is unnecessary when p is large. Comparisons to Monte Carlo simulations show excellent agreement, and EOF spectra of monthly averaged precipitation are also consistent with the analytical error spectrum (see below)

In case B the probability density approaches a sum of independent Gaussian distributions centered on each of the true eigenvalues, and each has a standard deviation equal to $\sqrt{2/n}$ times the true eigenvalue. From linear perturbation theory it can be shown that when the spacing between a given eigenvalue λ, and a neighboring one is on the order of the standard deviation, $\sqrt{2/n}\,\lambda$, then the sampling error in the EOF is comparable to the "nearby" EOF. This phenomenon is referred to as "mixing" and the nearby eigenvalues are said to be "effectively degenerate". An instructive example is a homogeneous process on the sphere, in which the

covariance between any two points depends only on the great circle distance between them. Such an assumption is equivalent to assuming that the true EOF's are the spherical harmonics (Obukhov, 1947, North and Cahalan, 1981). The rule of thumb that "to avoid mixing, the spacing must exceed the standard deviation" was borne out in Monte Carlo simulations in which the spherical harmonics were given a spectrum appropriate for the 500 mb height field (North, Bell, Cahalan, and Moeng, 1982). In this study the spherically homogeneous data was restricted to a finite domain of the sphere, and the exact EOF's ("patch harmonics") were derived, and compared to estimates determined from finite samples. The sample patterns are reminiscent of many empirical studies, and clearly illustrate the dependence of EOF's on domain boundaries, as well as the tendency to mix. Actual height field data also exhibit mixing, as can be seen in the well-known teleconnection study of Wallace and Gutzler (1981). The above rule of thumb shows that their first two pairs of EOF's are effectively degenerate, and this is borne out in their analysis of an independent data set.

We concentrate below on case A. Section 2 provides some preliminary theory, and describes Dyson's brownian motion analogy, which aids in the derivation of Eq. (1), and gives some physical feeling for the behavior of random matrices. Section 3 compares the analytical spectrum with Monte Carlo simulations, and Section 4 compares it with precipitation data. Section 5 concludes with a summary and discussion of further questions.

2. PRELIMINARY THEORY

Let us assume that we have N independent samples of a field on a grid of p points. The time mean is subtracted, leaving n = N -1 independent samples, and the spatial cross-covariance is computed for each pair of points. The eigenfunctions of this p x p matrix are the EOF's. They are orthogonal over the given spatial domain and uncorrelated over the given temporal sample. These two properties of EOF's are certainly convenient, but also sample-dependent. Here we wish to focus not on the functions themselves, but rather on their associated variances, which are the eigenvalues of the p x p covariance matrix, referred to collectively to as the EOF variance spectrum:

$$\{ \lambda_k,\ k = 1, 2, \ldots \min(p,n) \} . \qquad (3)$$

We shall assume p > n unless otherwise noted.

In classical spectral analysis the total variance may be partitioned among the various Fourier modes in many conceivable ways. In EOF analysis each successive EOF mode is determined by requiring that it have the maximum allowable variance, under the constraint that it be statistically independent of previous modes. This strict ordering property implies a simple interpretation of EOF spectra in probabilistic terms, as follows. Imagine an infinite ensemble of planets identical to Earth. The same field is measured on each, with the same sampling procedure, and an ensemble of covariance matrices and associated EOF spectra are determined. The planets are labelled by a parameter ϵ , and the largest eigenvalue is plotted as a function of ϵ, along with the second-largest, etc... A portion of this plot is shown in Figure 1.

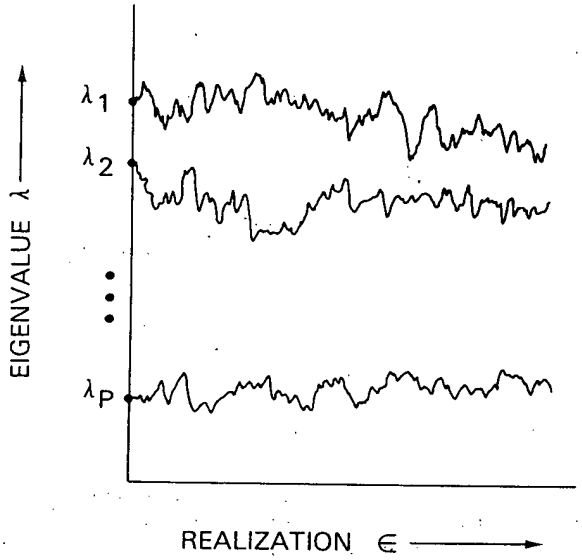


Figure 1

From this data we determine the cumulative distribution function, f(λ), namely the probability of finding an eigenfunction greater than or equal to some given value, say λ. It is unity at $\lambda = 0$, since the eigenvalues are all positive. It decreases monotonically for increasing λ since any eigenvalue which exceeds λ must exceed all lower values. Finally, it vanishes at some finite λ , since we assume that the total variance is finite. This function could be estimated from a single realization by simply counting the eigenvalues exceeding λ , and dividing by the total number of eigenvalues. This gives a "staircase" form, since f decreases by 1/p each time λ passes an eigenvalue. This can be written as follows:

$$f(\epsilon,\lambda) = \frac{1}{p} \sum_{k=1}^{p} \theta (\lambda_k(\epsilon) - \lambda)$$

where θ is the unit step function. The inverse of this function, namely λ (f), is a staircase with a level for each eigenvalue, and a step down at each f = k/p, for k = 1, 2,...p. Thus the traditional plot of eigenvalue versus mode-number (if lines are drawn down from and to the left of each eigenvalue to form steps and the modenumber is expressed as a fraction of p) may be interpreted as a realization of the cumulative distribution function turned on its side. This interpretation is especially helpful when p, and thus the number of steps, is large since then the distribution computed from a single realization typically approaches that of the full ensemble. This ergodic property will be demonstrated explicity for case A in the next section.

Let us now consider the special case in which N independent samples of a p-component vector (i.e. a field measured at p gridpoints) are drawn from a multivariate normal distribution having covariance Σ. The sample covariance matrix then has a Wishart distribution with mean Σ and $n = N - 1$ degrees of freedom. In case A the correlation between different components (i.e. gridpoints) vanishes, and all variances are equal, so Σ is the identity matrix. The joint density distribution of the sample eigenvalues is known for this case (Anderson, 1963). It may be written in the form

$$f(\lambda_1, \lambda_1 \cdots \lambda_p) =$$

$$C(p,n)\exp[-V(\lambda_1,\lambda_2,\cdots \lambda_p)] \qquad (5)$$

where $C(p,n)$ is independent of the λ_i, and where

$$V = \sum_{i=1}^{p} \left(\frac{1}{2}\lambda_i - \frac{1}{2}(n-p-1)\,\ell n\,\lambda_i - \sum_{j=i+1}^{p} \ell n(\lambda_i - \lambda_j)\right). \qquad (6)$$

Dyson's interpretation of this result is based on the following physical analogy. Let us imagine that the λ_i shown in Figure 1 represent the vertical coordinates of a set of p straight lines which are restricted to move in a plane perpendicular to the paper, and ε represents time. Each unit length of line has a uniform weight of 1/2 and a uniform charge of 1. There is a fixed charge of n-p-1 per unit length at $\lambda = 0$, and the system is immersed in a heat bath at temperature $k_BT = 1$. Statistical equilibrium is maintained by a balance between the electric and gravitational forces. The joint probability density for finding the p lines in the neighborhood of $\lambda_1, \lambda_2, \cdots \lambda_p$ at time ε satisfies the Fokker-Planck equation. As $\varepsilon \to \infty$ the solution approaches a stationary distribution independent of the initial state, and identical to (5) and (6), where V is the total potential energy. This exact analogy allows us to think of the last term in (6) as a Coulomb repulsion between the eigenvalues of a random matrix.

Following Wigner, we now consider the limit $p \to \infty$ with p/n fixed. The lines merge into a sheet having an equilibrium distribution of weight and charge denoted by $\rho(\lambda)$. The sums in (6) become integrals over λ weighted by ρ. The most probable ρ is determined by minimizing V under variations of ρ with a normalization constraint. This gives an integral equation for ρ which has the solution given by (1). This is the most probable density distribution for large p. The most probable cumulative distribution is then given by

$$f(\lambda) = \int_{\lambda}^{\infty} \rho(\lambda')d\lambda'. \qquad (7)$$

This may be evaluated in terms of elementary functions, and its inverse gives the eigenvalue as a function of fractional modenumber, $\lambda(f)$.

An additional quantity of interest is the cumulative fraction of the total variance contained in all modes up to some given modenumber f. To derive this we first transform variables. Let

$$y = (\lambda - 1 - p/n)/(2\sqrt{p/n}). \qquad (8)$$

Then

$$\rho(y) = \rho(\lambda)\frac{d\lambda}{dy} = \frac{2}{\pi}\,\frac{\sqrt{1-y^2}}{1 + p/n + 2\sqrt{p/n}\,y} \qquad (9)$$

and

$$-1 \le y \le +1. \qquad (10)$$

The cumulative variance is then

$$\int_0^{f(y)} \lambda(f')df' = \int_y^1 \lambda(y)\,\rho(y)dy = \int_y^1 \rho_o(y)\,dy = f_o(y) \qquad (11)$$

Here ρ_o and f_o are the density and cumulative distribution for p/n = 0. The density ρ_o is Wigner's semicircular distribution. To apply (11), one plots f_o versus f using y as a parameter. When $p \ll n$ then $f \to f_o$ and one obtains a straight line, so that the cumulative variance increases linearly from 0 to 1 as f goes from 0 to 1. This is due to the fact that the sample spectrum is approaching the true spectrum, a constant $\lambda = 1$ as can be seen from (2). For nonzero p/n we have $f_o > f$ for any given y. Thus the first f fraction of the modes contain $f_o > f$ fraction of the variance. These results will be compared with Monte Carlo experiments in the next section, and with precipitation data in Section 4.

3. MONTE CARLO SIMULATIONS

To simulate the results for case A a large number of random covariance matrices were computed from samples of p-component vectors, with each component taken independently from a normal distribution with unit variance. Two different double-precision random number generators were used with many different values of p and n. Figure 2 shows the effect of increasing p from 9 to 144. In each case we take $p \ll n$, and compare the staircase of ensemble mean eigenvalues with the smooth curve, which is $f_o(y)$. The top 9-level staircase represents an average of 100 spectra. Individual realizations tend to deviate from this mean by less than about one step. In terms of Dyson's analogy, this is due to the Coulomb repulsion between the levels. When p is increased to 36 as for the middle staircase, the step size is reduced by a factor of 4, and the Coulomb repulsion forces each realization to be

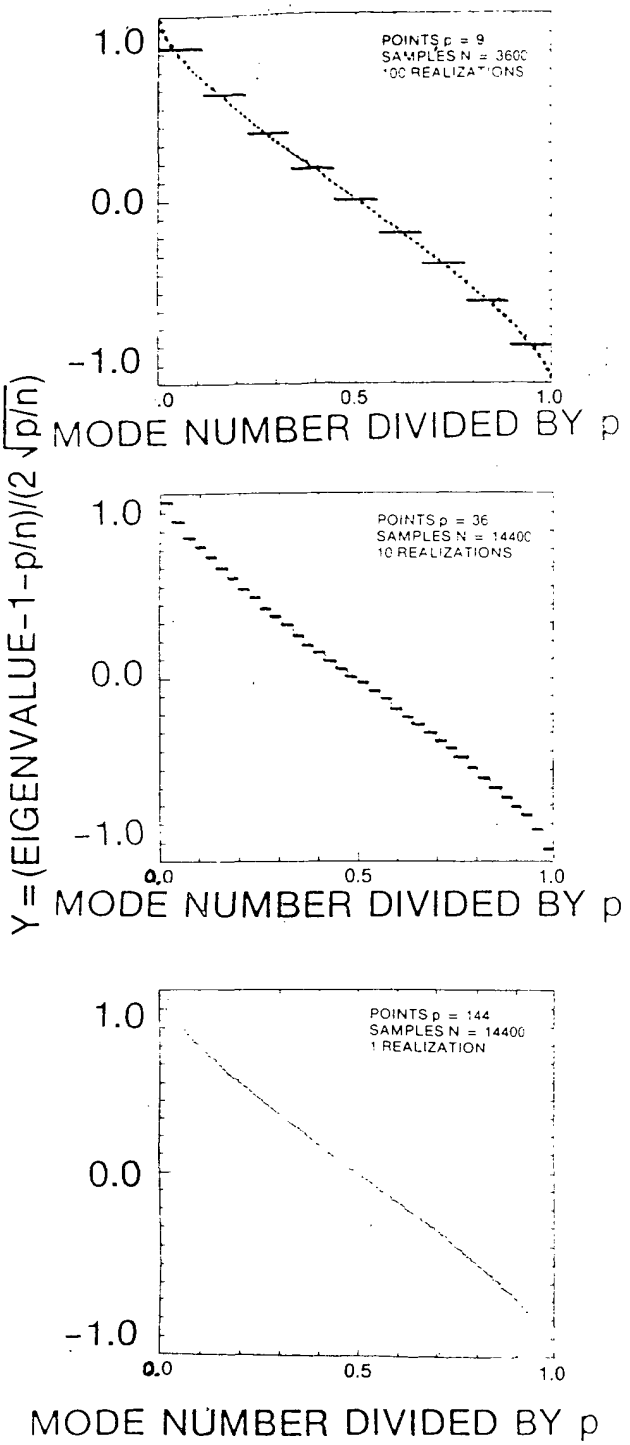


Figure 2

closer to the mean. In this case 10 spectra were averaged over. Finally, p is increased to 144 and a single realization is plotted in the bottom panel. Even the small deviations from $f_o(y)$ discernable near the top and bottom of the 144-step staircase would disappear if we had plotted f(y) with the correct value of p/n. We use the term "spectral ergodicity" to refer to this dramatic tendency of individual spectra to approach the most probable spectrum for large p.

To simulate the presence of a "signal" a vector with a fixed direction but variable magnitude was added to each random vector. The magnitude was normally distributed, and varied independently of the random components, with twice the variance. Since the "noise" is homogeneous the results do not depend on the direction (i.e. spatial pattern) of the signal.

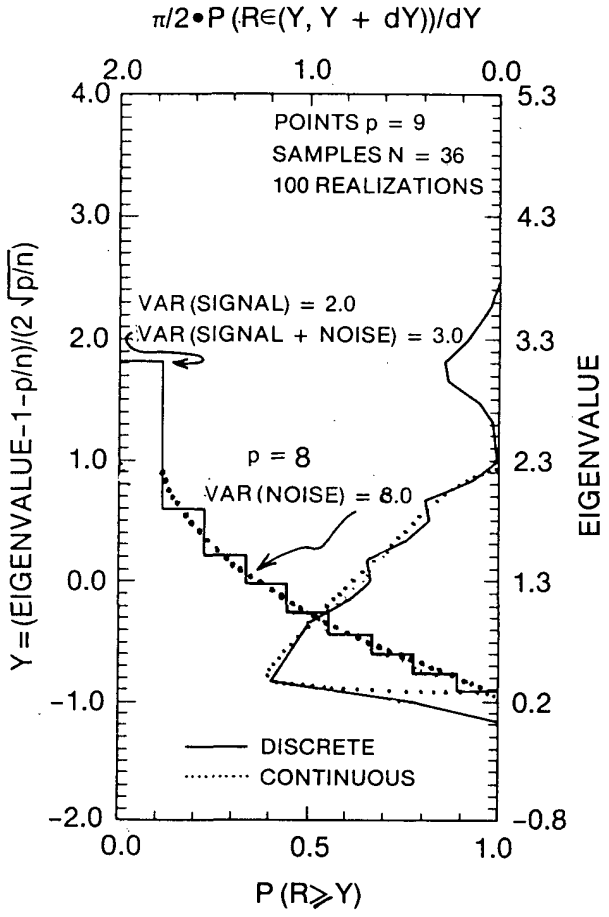


Figure 3

The staircase in Figure 3 is an average of 100 such spectra, each computed from 36 samples on a grid of 9 points. Superimposed on this is a second solid line plot representing a histogram of all 900 eigenvalues, plotted against the upper horizontal scale. The upper part of the histogram shows a normal distribution around a mean eigenvalue of 3, which is the sum of the signal and noise variances. The lower part of the histogram is a pure noise distribution for the remaining 8 degrees of freedom, namely ρ (y) for p = 8, plotted as a dotted curve. The staircase shows the same two features--a single high step containing the signal on top of the background, and 8 low steps which follow f(y) for p = 8, which is the dotted curve. (Note that f(y) the lower scale must be contracted by a factor of 8/9).

In the next section we apply the same ideas to see if there is any evidence of a "signal" in EOF spectra of precipitation.

4. PRECIPITATION

In contrast to the geopotential height field, the field of precipitation has relatively short length and time scales. Monthly mean precipitation at a given station is almost uncorrelated with that of stations a few hundred kilometers away, or with that measured at the same station one month later. This makes it an ideal candidate to compare with the analytic error spectrum given by (1) and (7). The variance of precipitation has a strong geographic dependence, but for the purpose of computing EOF's the anomalies are normalized to have unit variance at each point. The distribution of the anomalies is far from gaussian, but our analytic result is not sensitive to this assumption. When p is large, as Wigner has emphasized (and as we have verified with simulations not shown here), the one crucial assumption is the lack of any correlation between fluctuations at different points.

One of the pioneering applications of EOF's to geophysical fields was the fine study of Kutzbach (1967), who computed eigenfunctions and eigenvalues from measurements of monthly mean sea level pressure, surface temperature and precipitate at a number of stations in 23 regions of roughly equal area over the continental United States for 25 Januaries. The 8 largest eigenvalues of precipitation are expressed as a percent of the total variance and plotted versus the fractional modenumber in Figure 4.

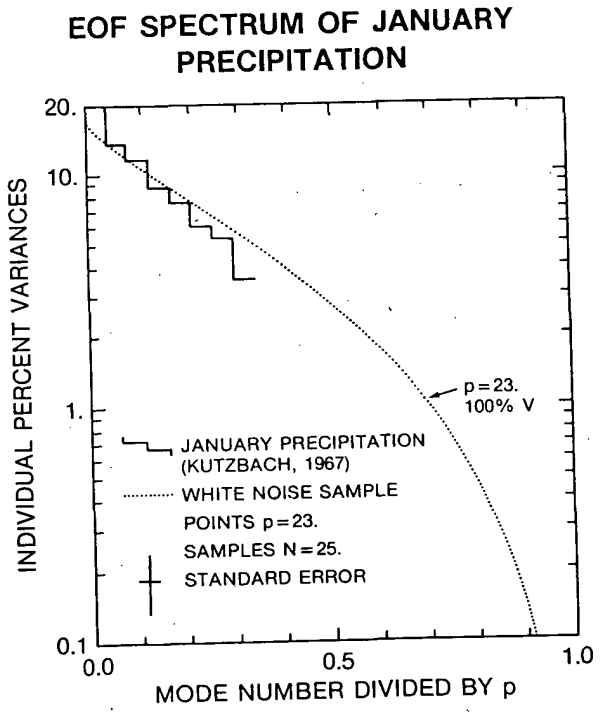


Figure 4

The dotted curve represents the most probable spectrum assuming that the 23 regions are uncorrelated and that the 25 Januaries are independent If the true eigenvalues are unequal then like the "signal" in Figure 3, the samples eigenvalues would be normally distributed about the true ones. The standard deviation is shown as an error bar in the lower left corner of Figure 4. None of the measured eigenvalues is further than one standard deviation from the noise spectrum. It is clear that the "signal" in the measured eigenvalues is consistent with zero, since a single realization could deviate from the most probable noise spectrum by the amount shown in this plot. Of course this does not imply that the record of January mean precipitation is nothing but measurement error, but rather that the length scales are so short that 25 years is insufficient to reject the hypothesis that the true eigenvalues are equal.

In a more recent study, Walsh and Mostek (1980) computed the EOF's and spectra of monthly mean precipitation from 78 years of data at 61 points over the United States. Only the first few eigenvalues for January and July were published, but the first author has kindly made the complete set available. The January spectrum is shown in Figure 5. The upper dotted curve is the analytic noise spectrum for p = 61, N = 78. The first four eigenvalues are more than one standard deviation above the noise spectrum. If they represent some "signal" then their variance should be removed before comparing the next eigenvalue with noise. If the next one is significantly above the noise, its variance is removed, and so on. If two eigenvalues are within one standard deviation of each other, they are effectively degenerate, and should therefore be treated as both signal or both noise. Eigenvalues 6-11 in Figure 5 are effectively degenerate. The first 11 eigenvalues contain 71.6% of the variance. For the sake of illustration we consider them as signal, and compare the remaining 50 with the lower dotted curve, which is the analytic noise spectrum for p = 50, N = 78, with the variance reduced and the horizontal axis rescaled. They are consistent with pure noise.

The July spectrum is shown in Fig. 6, along with the noise spectra for p = 61 and p = 50. In this case the first 11 eigenvalues contain only 54.1% of the variance, and only the first four can be distinguished from noise. The clear difference between the January and July cases is presumably due to the different length scales involved. January precipitation is associated with synoptic scale systems, while July precipitation is more convective, and the smaller scale makes it appear more like pure noise.

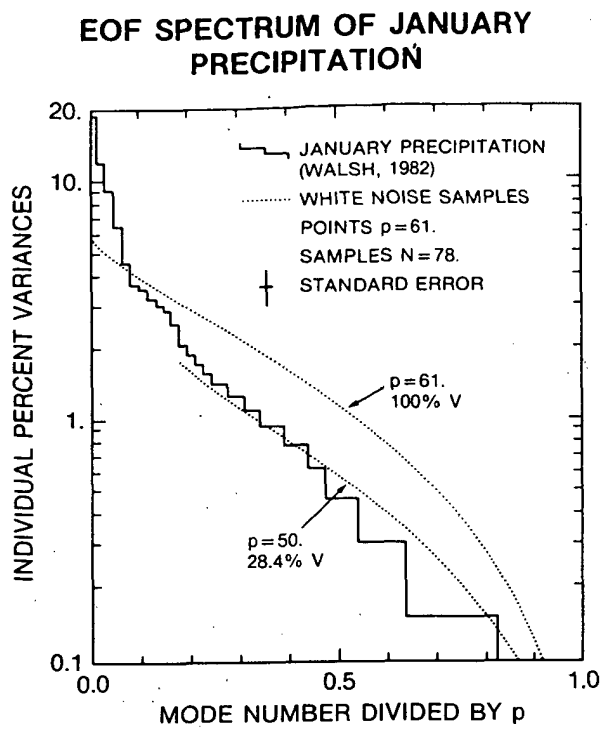


Figure 5

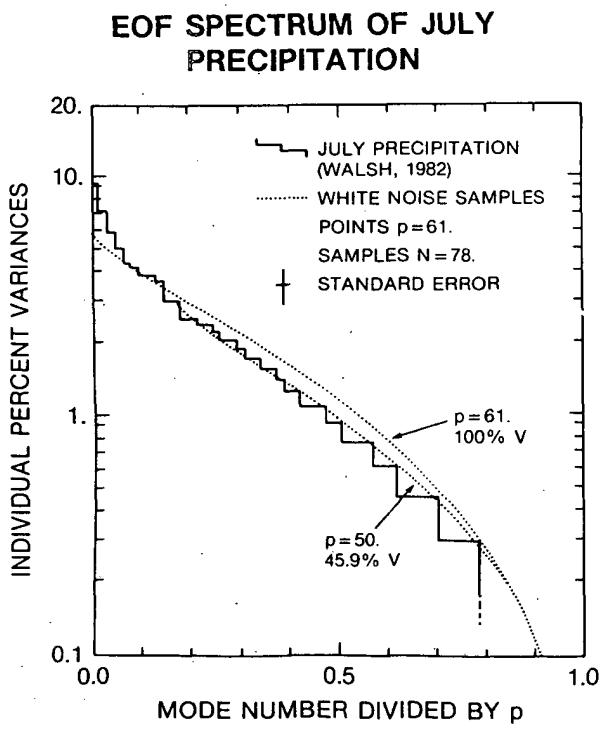


Figure 6

5. CONCLUSION

This paper can be summarized in terms of a recommended procedure for plotting EOF variance spectra: 1. Plot the logarithm of the eigenvalue versus the modenumber divided by p.

2. Since this is a cumulative distribution in λ, form a staircase by extending horizontal lines to the left and vertical lines below each eigenvalue.

3. Associate a standard error given by $\sigma_k = \sqrt{2/n}\ \lambda_k$ with each eigenvalue λ_k, and distinguish only those eigenvalues more than σ_k apart.

4. For comparison with spatially uncorrelated data, plot the analytic noise spectrum given by

$$\text{modenumber}/p = \frac{1}{2\pi p/n} \int_{\lambda p/V}^{(1+\sqrt{p/n})^2} \left[-y^2 + 2\,(1+p/n)y - (1-p/n)^2\right]^{1/2} dy/y.$$

where p and n are the number of points and degrees of freedom, V is the total variance and λ varies between

$$\lambda_{\min} = \left(\frac{1}{\sqrt{p}} - \frac{1}{\sqrt{n}}\right)^2 V$$

$$\lambda_{\max} = \left(\frac{1}{\sqrt{p}} + \frac{1}{\sqrt{n}}\right)^2 V\ .$$

5. If some of the eigenvalues are accepted as significant, the remaining ones should be compared with a noise spectrum having V and p reduced accordingly.

How does one estimate the degrees of freedom when the samples are not independent? Persistence in the data may greatly reduce n. One estimate of n is the total time divided by the point autocorrelation time. This is probably reasonable in comparing with homogeneous uncorrelated noise. However, fields with large spatial correlations, such as the height field, often have more persistence in the large scales, so that one has fewer degrees of freedom in estimating largescale EOF's.

Another approach to the time dependence is to first filter the data to separate low and high frequency components, and then compute EOF's separately in each case. Some of the differences seen may be due to the fewer degrees of freedom in the low frequencies.

ACKNOWLEDGEMENT

I would like to thank John Kutzbach and John Walsh for use of their results, and Thomas Bell, and Gerald North for stimulating conversations on this topic. I would also like to thank Valerie Robertson and Lora Thompson for the typing and preparation of the manuscript.

6. REFERENCES

Anderson, T. W., 1963: Asymptotic theory for principal component analysis, Ann. Math. Stat., 34, 122-148

Cahalan, R. F., 1982: EOF spectral estimation in climate analysis, unpublished.

Dyson, F. J., 1962: A brownian-motion model for the eigenvalues of a random matrix, J. Atmos. Sci., 38, 1191-1198.

Kutzbach, J. E., 1967: Empirical eigenvectors of sea-level pressure, surface temperature and precipitation complexes over North America, J. App. Meteor., 6, 791-802

North, G. R. and R. F. Cahalan, 1981: Predictability in a solvable stochastic climate model, J. Atmos. Sci., 38, 504-513.

_____, T. L. Bell, R. F. Cahalan, and F. J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions, Mon. Wea. Rev., 110, 699-706.

Preisendorfer, R. W. and T. P. Barnett, 1977: Significance tests for empirical orthogonal functions, Preprints 5th Conf. on Prob. and Stat. in Atmos. Sci., Las Vegas, Amer. Meteor. Soc., 169-172.

Wallace, J. M. and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the northern hemisphere winter, Mon. Wea. Rev., 109, 784-812.

Walsh, J. E. and A. Mostek, 1980: A quantitative analysis of meteorological anomaly patterns over the United States, Mon. Wea. Rev., 108, 615-630.

Wigner, E. P., 1955: Characteristic vectors of bordered matrices with infinite dimensions Ann. Math., 62, 548-564.

_____, 1957: Statistical properties of real symmetric matrices with many dimensions Can. Math. Congr. Proc., University of Toronto Press, Toronto, Canada, 174-184.

WEATHER MAP PATTERN CLASSIFICATION: REFINEMENTS OF A CORRELATION METHOD

Donald E. Petzold

Department of Geography
University of Maryland
College Park, Maryland 20742, U.S.A.

1. INTRODUCTION

Synoptic climatological analyses and applications often involve the classification of regional airflow, storm tracks or pressure patterns. These features may be grouped by employing subjective techniques which are based on inexact visual correlations. Objective classifications based on statistical procedures provide an alternative to the subjective methods. With these, the possibility of errant human judgement overshadowing the usefulness of the weather type determinations is reduced considerably.

A correlation technique, first presented by Lund (1963), has gained rather widespread acceptance and use. The objective of this method is to identify unique weather patterns which occur much more frequently than by chance. The identity of small scale pressure features is preserved by defining a point grid of representative pressure values for each map in the study period, such the distance between adjacent points defines an area suitable to the needs of the study. Pressure data from each map are correlated with the corresponding data from every other map in the study period. The degree of similarity between unique map patterns is determined by the selection of a threshold correlation coefficient (r_o). Correlations below this value are not considered.

One map will correlate with the greatest number of others at or above r_o. This map, and all others correlated with it, are removed from the data pool as type A. The remaining maps are scanned to produce type B and so on until the next type to be chosen is too small to be of practical use (usually between 1-2% of the total study period). The final step involves a reassignment of all maps to types with which they correlate most highly. Despite this, usually more than 10% of the study period remains unclassified. This may be the result of (1) the choice of r_o; (2) grid density; and/or (3) the predetermination of the lower limit of the type size (i.e., the "left-overs" are clumped into an unclassified group). It is most desirable to produce the smallest unclassified pool possible.

2. METHODOLOGY

The procedure was employed in a synoptic climatological study of lake evaporation over subarctic Quebec-Labrador. Daily surface weather maps of three evaporation seasons (459 days total) were classified into 32 unique types with 10% remaining unclassified. This study resulted in the revision of the classification procedure by adding and refining certain steps.

3. REFINEMENTS OF THE PROCEDURE

Perhaps the most important revision is the classification of all maps. This requires the continuation of the sorting process to define all types to a minimum of two members. (Previous applications stopped at 1-2% of the total study period.) The reason for this is that even the smallest types can gain significantly during reclassification. The second revision concerns the unclassified days. In previous studies, they were grouped together and disregarded. However, close scrutiny revealed at least two types of unclassified days: (1) those which are unique and do not correlate with any other day; and (2) those which become unclassified during the sorting procedure. Members of the second category may become type-representative days upon reclassification. The third revision is the definition of the smallest practical size of a weather type as the last step in the procedure. Table 1 lists the steps of the revised procedure.

Table 1. The revised procedure for the pressure pattern classification by the correlation method.

STEP	
1	INITIALIZATION
a.	Correlate pressure grid of each map with that of all other maps.
b.	Define the threshold value of the correlation coefficient (r_o)
2	THE FIRST COUNT
a.	Define Type 1 as the map whose pattern correlates with most other days at $r \geq r_o$.
b.	Extract all Type 1 maps from further analysis.
c.	Define Type 2 as the map whose pattern correlates with the greatest number of remaining days at $r \geq r_o$.
d.	Extract all Type 2 maps from further analysis.
*e.	Repeat Steps 2c and d until all maps are classified in a type which may have a minimum size of two members.
*f.	Separate unique unclassifiable cases from those unclassifiable cases eliminated in the first count.
3	RECLASSIFICATION
a.	Reassign those days which correlate at a higher r with another type-representative day.
b.	Identify any days from those eliminated in the first count (Step 2f) which may serve as a type-representative day of a new group. Include these groups into the type catalogue.
c.	Determine the minimum practical size of a type. Extract all types with fewer members and include them in an unclassified group.

*Steps added to the classification procedure.

4. APPLICATION OF THE REVISED PROCEDURE

The impact of these changes on the present study was noticeable:

(1) 11 types, initially each with $\leq 1\%$ of the total study period, grew significantly upon reclassification. Two of these ranked 8th and 10th in type size. Thus, 31 days were assigned type memberships which otherwise would have been left in the unclassified pool.

(2) Three unclassified days, eliminated during the initial sorting procedure, were found to be type-representative days of minor types which total 15 days. This represents the better placement of over 3% of the total data set.

From these new insights into the correlation analysis and procedure, it is evident that past applications of this technique are neither as complete nor as accurate as they purport to be. The 88 membrs of the present study period, that otherwise would have been unclassified or placed in less suitable types, serve as valid evidence that these refinements produce a most effective definition of a weather type catalog by the correlation method.

5. REFERENCES

Lund, I., 1963: Map-pattern classification by Statistical methods. J. Appl. Meteor., 2, 56-65.

SPECIFICATION OF COMPLEX MODES OF CIRCULATION WITH T-MODE FACTOR ANALYSIS

Michael B. Richman

Illinois State Water Survey
P.O. Box 5050, Station A
Champaign, Illinois 61820-9050, U.S.A.

1. INTRODUCTION

A difficult task a data compression technique should accomplish is the *realistic* depiction of a complex meteorological field over a large domain. An eigenvector procedure which achieves this goal more accurately than previously used techniques will be explored: T-mode obliquely rotated factor analysis. Over the past three years more than 60 applications of eigenvector techniques have been published in meteorological/climatological journals. This popularity arises from an advantage of using an eigentechnique, the *data reduction capabilities* of a truncated series of eigenvectors. Factor analysis (FA), in particular, has advantages over other eigentechniques (e.g., principal components (PC), empirical orthogonal functions (EOF)) since the factor model partitions measurement error out of the analysis, leaving only reliable variance (most of which is known as "common variance") to be analyzed.

The factor model can be specified in six possible operational modes depending on which parameters are chosen as variables, individuals and fixed entities. These six modes have been defined as O, P, Q, R, S, and T by Cattell (1952). Each of these six modes will result in a *unique* clustering of variables when *rotation* is applied to the factor model. A rotated analysis is desirable as Buell (1975; 1979) has shown that unrotated eigenvectors of *different* correlation functions on a geometrically shaped domain have similar patterns which do *not* reflect the covariation among the variables. Numerous published studies support Buell's findings as their unrotated eigenvectors of various unrelated fields at different geographical locations all show the characteristic pre-determined patterns. A limited inventory of recent unrotated EOF/PC studies (and a few pertinent older studies) has been compiled for those analyses with a square/rectangular domain. The shape of EOF/PC's are denoted by (+) corresponding to a positive anomaly over the entire domain, (+ -) corresponding to a positive anomaly in the west and a negative anomaly in the east, (- + -) corresponding to a negative anomaly in the west, positive anomaly in the center and negative anomaly in the east portion of the domain, (±) corresponding to a positive anomaly in the north and a negative anomaly in the south and finally ($\pm\ \mp$) which corresponds to positive anomalies in the northwest and southeast with negative anomalies in the southwest and northeast. It should be noted that all of the patterns are also valid for the inverse of what is shown (e.g., (+ -) and (- +) are the same pattern due to the reflectivity property of EOF/PC) and also the domain can be rotated 90° to be aligned with the longest axis of the rectangle (e.g., (- + -) and ($\mp\pm$) are also the same pattern). The results of the inventory are shown in Table 1 which indicates that the first EOF/PC is positive in all of the rectangular domains, regardless of meteorological parameter or geographical location. The same holds true for EOF/PC2. The third through fifth EOF/PC are typically mixed, some are (- + -) while others (±). This is exactly what Buell (1979, p. 191) found for another correlation function (Buell's Fig. 9) where the third pre-determined function was (±), the fourth ($\pm\ \mp$) and the fifth (- + -).

TABLE 1. Geometrical shape of various unrotated EOF/PC studies in comparison to Buell's (1975, p. 190, Fig. 5) *rectangular* domain.

Author, Year	Parameter	EOF/PC1	EOF/PC2	EOF/PC3	EOF/PC4	EOF/PC5
Buell (1975)	Correlated Noise	(+)	(+ -)	(- + -)	($\pm$)	($\pm\ \mp$)
Klaus (1978)	Precipitation	(+)	($\pm$)	(+ -)	($\mp$)	
Willmott (1978)	Precipitation	(+)	($\pm$)	(- +)	($\mp$)	
Dyer (1979)	Precipitation	(+)	($\mp$)	($\mp$)		
Bedi & Bindra (1980)	Precipitation	(+)	($\pm$)	(+ - +)		
Walsh & Mostek (1980)	Precipitation	(+)	(+ -)	(- + -)		
Diaz (1981)	Precipitation	(+)	(+ -)	($\pm$)	($\mp$)	
Diaz & Fulbright (1981)	Precipitation	(+)	(- +)	($\mp$)	(+ - +)	
Tabony (1981)	Precipitation	(+)	($\mp$)	($\mp$)		
Karl & Koscielny (1982)	Precipitation	(+)	(+ -)	($\pm$)	(+ - +)	($\mp\ \pm$)
McGuirk (1982)	Precipitation	(+)	(+ -)	(- + -)		
Lamb & Richman (1982)	Precipitation	(+)	($\pm$)	(- +)	($\mp$)	($\mp\ \pm$)
Lorenz (1956)	Sea-level pressure	(+)	(+ -)	($\pm$)	(+ - +)	
Veitch (1965)	Sea-level pressure	(+)	(+ -)	($\mp$)	(+ - +)	($\pm\ \mp$)
Davis (1976)	Sea-level pressure	(+)	($\mp$)	(- +)	(- + -)	($\pm\ \mp$)
Brier & Meltesen (1976)	Sea-level pressure	(+)	($\mp$)	($\mp$)		
Walsh & Mostek (1980)	Sea-level pressure	(+)	(- +)	(- + -)		
Brinkmann (1981)	Sea-level pressure	(+)	($\pm$)	(+ -)	($\mp\ \pm$)	
Walsh & Mostek (1980)	Temperature	(+)	(- +)	($\pm$)		
Walsh & Richman (1981)	Temperature	(+)	(- +)	($\pm$)	(- + -)	
Gray (1981)	Temperature	(+)	($\mp$)	(+ -)	($\pm\ \mp$)	($\mp$)
Karl et al. (1982)	Temperature	(+)	(+ -)	($\pm$)		

Therefore, it appears that Buell's functions 3 through 5 may emerge in any order, yet the important aspect is that they will all be present in some order. Overall, the results of Table 1 support Buell's contention that the unrotated EOF/PC approach may *not* yield realistic individual modes of variation.

Rotation of the factors has been suggested as a viable alternative to uncover realistic modes of variation by Walsh and Richman (1981, p. 780-781). In their experiment they decomposed a set of known sea-level pressure patterns into two unrotated PCs which agreed with the first two of Buells patterns. These two unrotated PCs were then rotated to the varimax criterion (Kaiser, 1958) and the two input modes were accurately reproduced. The experiment is graphically reproduced in Fig. 1 which illustrates the effect of *rotation* on the analysis.

As stated previously, applying rotation to the unrotated factors for each of the six modes of decomposition results in a unique clustering of the variables in an analysis. *Determining which of the six modes to apply to data under investigation requires prior knowledge about the nature of the field under investigation.* In the present analysis the data were 700-mb height values over the western half of the northern hemisphere. A study of the data indicated that 4 or 5 anomaly centers were present on the input maps in 75% of the cases. Therefore, the mode chosen will need to accurately reproduce the *inherent complexity of the 700-mb height field.* A method for determining which mode to choose will be shown in section 3.

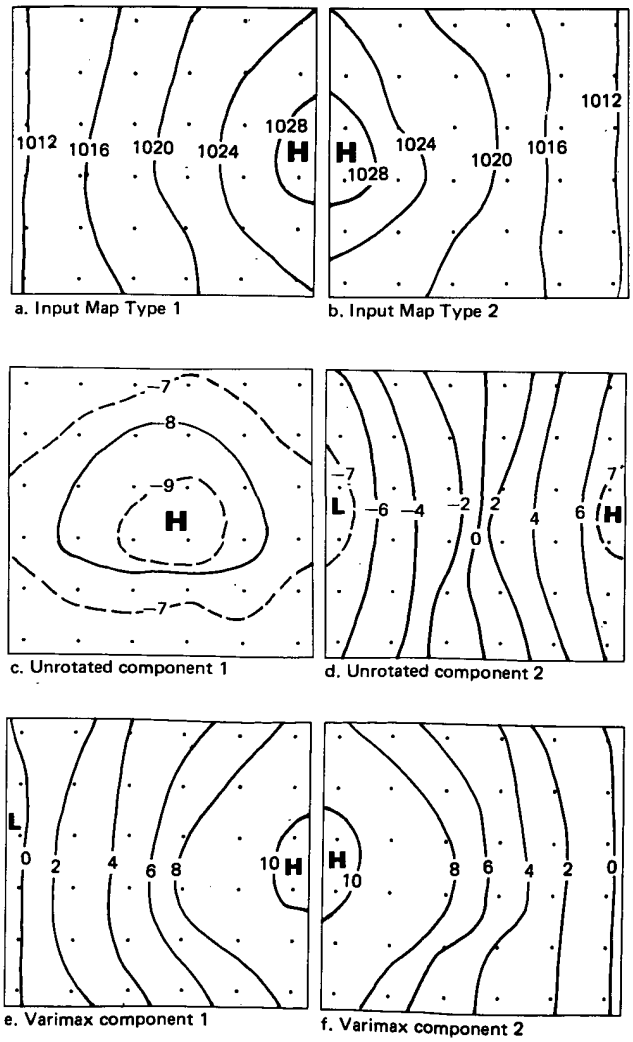


Fig. 1. Input maps of sea-level pressure, the first two unrotated PC of the covariance matrix and the first two PC rotated to satisfy the varimax criterion.

2. DATA AND PRELIMINARY HANDLING

The 700-mb data set consisting of 1664 fields of weekly observations for the period January 1947 through December 1978 were obtained from the National Center for Atmospheric Research data base. One hundred grid points were utilized which encompassed an area from 0° westward to 150°E; 20°N to 60°N as shown in Fig. 2. The domain was chosen in order to represent most of the Pacific and Atlantic Oceans as well as a large portion of North America. The data were standardized in order to remove long-term trends and seasonality by expressing the data as departures from the 32 yr. weekly means. Each year of data were then partitioned into separate yearly blocks and transposed such that the 52 weekly observations were treated as variables for each analysis. A question which immediately arises is: why not analyze all 1664 observations at once? This was done since the CDC CYBER 175 computer used in the analysis could not handle 1664 variables simultaneously. When the next generation of computers becomes available, it would be beneficial to analyze all 1664 variables at once rather than using yearly blocks of 52 variables. In the present study a subset of 10 yearly blocks (1947-1956) were separately analyzed to establish the basic circulation modes and combined (via higher-order factor analysis and procrustes target rotation) to achieve patterns and time series valid for the total period.

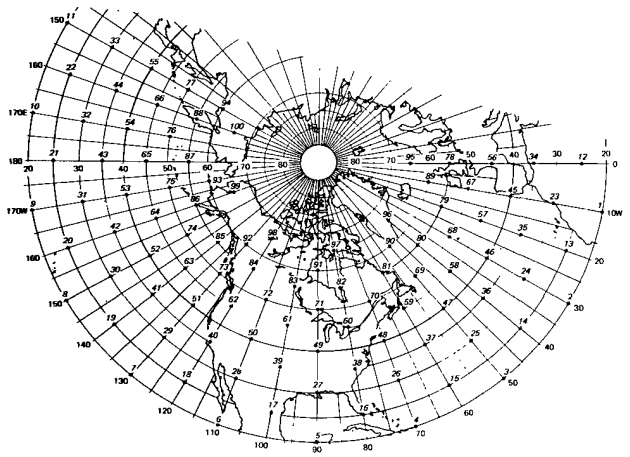


Fig. 2. Domain for the 700-mb T-mode analysis. Gridpoints selected are numbered 1-100.

3. ANALYSIS TECHNIQUES

a. Specification of the common factor model

The common factor model differs from other eigentechniques as the total data variance is separated into two categories. The first category, the common variance, contains only that variance which is common to two or more variables. Therefore, it is a measure of the overlap of variables in the analysis. *The interpretation of the analysis is based totally on this common variance.* The second category, the unique variance, is the remaining variance which is not common. This is typically considered to be largely comprised of error variance and, therefore, is treated as a residual.

Given: An N x n data matrix $\underline{Z}$

$$\underline{Z} = \{Z_{ij}: i=1, \ldots, N; j=1, \ldots, n\}$$

where i indexes individuals and j indexes variables. For every i,j, "Z_{ij}" denotes the value of some physical field or some derived field. The factor analysis model representation of Z_{ij} is

$$Z_{ij} = \sum_{r=1}^{m} f_{ir}\, a'_{jr} + e_{ij} + \Delta_{ij} = \tilde{Z}_{ij} + e_{ij} \quad (1)$$

$$+ \Delta_{ij}$$

$$i=1, \ldots, N; j=1, \ldots, n; r=1, \ldots, m; m \leq n$$

where e_{ij} = the *unique part* of Z_{ij}
$\tilde{Z}_{ij}$ = the *common part* of Z_{ij}
f_{ir} = the r^{th} *common factor score* for the i^{th} individual
a_{jr} = the *common factor loading* of the j^{th} variable
Δ_{ij} = the lack of fit between the factor model and the data. [This will be assumed to be zero in subsequent development].

In matrix form (1) becomes

$$\underline{Z} = \underline{F}\,\underline{A}' + \underline{E} \quad (2)$$

The factor scores (f_{ir}) and factor loadings (a_{jr}) can be developed with the following definitions (assumed to be in standardized form)

$$\underline{R} \equiv \underline{Z}'\underline{Z} \ (n \times n) \text{ (data correlation matrix)} \quad (3)$$

$$\underline{\Phi} \equiv \underline{F}'\underline{F} \ (n \times n) \text{ (factor score correlation matrix)} \quad (4)$$

$$\underline{S} \equiv \underline{Z}'\underline{F} \ (n \times n) \text{ (factor structure matrix)} \quad (5)$$

Since $\underline{Z}' = \underline{A}\,\underline{F}' + \underline{E}'$ can be substituted in (5)

$$\underline{Z}'\underline{F} = (\underline{A}\,\underline{F}' + \underline{E}')\,\underline{F}$$

$$\underline{Z}'\underline{F} = \underline{A}\,\underline{F}'\underline{F} + \underline{E}'\underline{F} \quad (6)$$

$\underline{E}'\,\underline{F} \equiv 0$ in the factor model by definition. Applying (4) yields

$$\underline{S} = \underline{A}\,\underline{\Phi} \quad (7)$$

These relationships can be applied to yield the correlation matrix representation of the factor model

$$\underline{R} = \underline{A}\,\underline{\Phi}\,\underline{A}' + \underline{U}^2 \quad (8)$$

where $\underline{U}$ are the uniquenesses.

In (8) the common portion of $\underline{A}\,\underline{\Phi}\,\underline{A}'$ can be seen as a combination of $\underline{A}\,\underline{\Phi}$ (the *factor structure matrix*) and $\underline{A}$ (the *factor pattern matrix*). In the initial unrotated solution (and under orthogonal rotation) $\underline{\Phi} = \underline{I}_n$, therefore, $\underline{A}\,\underline{\Phi}$ and $\underline{A}$ are identical.

In Section 1, the six modes of decomposition were introduced. In meteorological studies of physical fields in space and time there are three entities: a physical field, time and a location (or station). A factor analysis of a physical field can therefore be made by varying any two of these three entities and holding the third fixed. Examination of (1) and (2) reveals two basic matrices which are analyzed in a factor analysis. These are suggested by a_{jr} (the factor loadings) and f_{ir} (the factor scores). These are applied to the six modes in Table 2. For example, suppose $\underline{Z}$ has been rendered into an *S-mode* representation. From Table 2 note that j indexes stations, and i indexes times, of some fixed field. Then a plot of a_{jr}, for fixed r, as a function of j would be a map of the factor

TABLE 2. CATEGORIES OF DATA

FA mode	In Eq. (1) variable index j denotes:	In Eq. (1) individual index i denotes:	Fixed entity
O-mode	time	field	station
P-mode	field	time	station
Q-mode	station	field	time
R-mode	field	station	time
S-mode	station	time	field
T-mode	time	station	field
Graphical Device:	$\{a_{jr}, j=1, \ldots, n\}$ factor loading display	$\{f_{ir}, i=1, \ldots, N\}$ factor display or factor score display	

loadings as distributed over the spatial extent of the data of $\underline{Z}$. The plot of f_{ir}, for fixed r, as a function of i, would be a time series of the rth factor score. Again, if *T-mode* had been used to analyze $\underline{Z}$, the factor loadings, a_{jr}, would plot up as a time series for each r while f_{ir} would plot as a map of factor scores for each r. In a *P-mode* analysis, the factor loadings, a_{jr}, would give a series of n numbers, each associated with a member of a set of fields at a fixed station, while the factor scores, f_{ir}, would yield an associated time series. This display is not as visually interpretable as the more intuitive S- and T-mode displays, but it is just as logical as these two. Similar interpretations of the factor loadings and factor scores for the other three modes (O,Q,R) can be made as can be seen in Fig. 3.

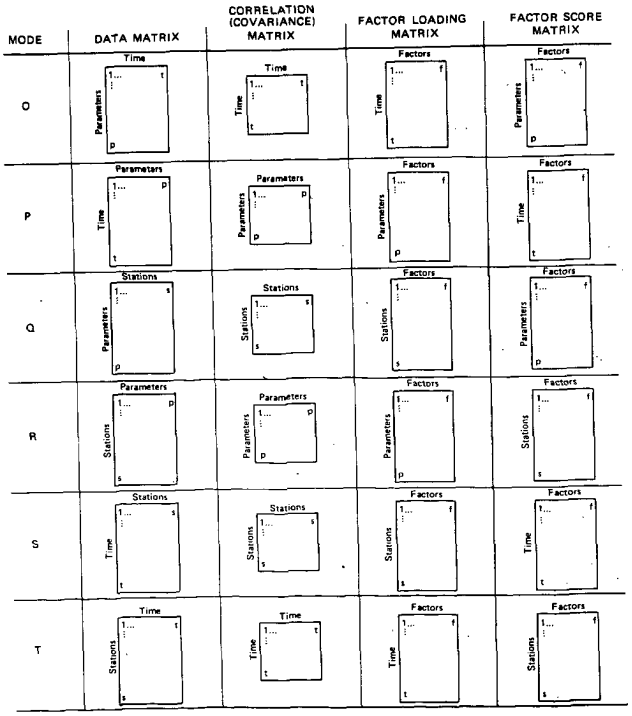


Fig. 3. Data matrix, dispersion matrix, factor loading matrix and factor score matrix under each of the six modes of decomposition.

b. Rotation and simple structure

Two types of rotation need to be examined: simple structure rotation and procrustes target rotation. From a meteorological standpoint, the validity of most rotations hinge on the significance of *simple structure* for the data under investigation. Thurstone (1947) proposed simple structure to resolve transformational indeterminacy in the factor model since an infinite number of alternative solutions ("linear transformations" or "rotations") satisfy the models equations (Richman, 1983).

Simple structure implies that variables in an analysis will cluster together when represented as vectors in factor space. Additionally, these vector clusters were assumed to have a greater density than surrounding areas. In terms of the factor loadings which are analyzed, this corresponds to high loadings *between* variables within a cluster and lower loadings for variables which do not fall within a cluster. *A factor which delineated one cluster would, therefore, exhibit high loadings on those variables within the cluster and near-zero loadings on those variables outside of the vector cluster.* Simple structure results in rotating the factors around the origin (via a linear transformation) either obliquely or orthogonally until each factor is aligned or colinear with a distinct cluster of variable vectors. Specifically, this is accomplished by rotating the factors such that the number of near-zero loadings (termed "variable loadings in the hyperplane") are maximized. *This is known as simple structure* and the higher the percentage of near-zero loadings, the stronger the simple structure. Numerous algebraic formulas have been developed to attempt to *approximate* Thurstone's simple structure mathematically (e.g., varimax (Kaiser, 1958)). In this study, a topological approach is used, known as DAPPFR (Tucker and Finkbeiner, 1982). This is an oblique rotation, therefore $\underline{\Phi} \neq \underline{I}_n$ necessarily. The advantage of using an oblique topological approach is twofold. First, oblique rotations tend to more accurately position the factors since no artificial orthogonality constraint is applied. Secondly, Thurstone's simple structure is a topological concept and application of a topological technique typically produces more near-zero factor loadings than application of a simple algebraic formula. Richman (1983) discusses this in more detail.

At this point, reviewing what has been developed, we find (1) the factor model introduced, (2) the six modes of decomposition defined and (3) simple structure introduced as a concept. In the introduction it was stated that obliquely rotated T-mode factor analysis yielded a more accurate representation of the 700-mb height field over a large domain. It was also stated that deciding which mode to apply required physical knowledge about the data under investigation. The importance of the application of simple structure rotation to the modes of decomposition results in a *unique clustering and simplification of the variables under each mode of decomposition.* All previous studies which have utilized rotation to isolate modes of height fields have been in the S-mode (e.g., Horel, 1981). In the purest sense of simple structure under an S-mode decomposition the rotated spatial maps would exhibit high loadings on a few gridpoints and as many near-zero loadings as possible. This approach is useful if identification of a subset of gridpoints which exhibit large fluctuations of variance is required. The disadvantages of the S-mode approach arise when the spatial loadings are used to represent modes of circulation for a complex field such as 700-mb. Cohen (1981, pp. 58-59; 162-166) presents an example of this disadvantage on 500-mb data when the varimax rotation is applied. In this case, each component isolates one half of a wave (i.e., one ridge or trough). The severity of this problem is partly a function of the size of the domain. For a large domain, such as presented herein, the effect is very undesirable. Further problems arise in S-mode with rotation as the factor scores do *not* have the simple structure applied (since simple structure is only applied to the variables in an analysis). Consequently, the factor scores are typically complex, which results in a series of factors being needed to properly explain one observation.

Conversely, the T-mode decomposition avoids many of the drawbacks of the S-mode when used to depict realistic 700-mb circulation modes over a large domain. Simple structure is still present, and in its purest form, would allow one or sometimes two factors to be needed to properly describe one observation. This occurs since the simple structure operates on the time series (which are treated as variables in T-mode) rather than the spatial patterns. Therefore, the factors are time intervals which have similar spatial characteristics and are clustered via the rotation. The most significant improvement occurs as there is *no* simple structure placed on the spatial patterns (factor scores). Accordingly, the complex wave patterns at 700-mb can be properly represented over the total domain. Recalling that the input data typically contain 4 or 5 anomalies, the T-mode spatial patterns are therefore superior to the S-mode spatial patterns for representing complex fields since as many waves as are needed to accurately depict the true mode will be present.

c. DAPPFR simple structure rotation

The obliquely rotated T-mode analysis was derived via the DAPPFR simple structure rotation (Tucker and Finkbeiner, 1982). In this topological approach, each variable (n) is assigned a specific probability of being located in the hyperplane (having a near-zero loading) which is an n-1 subspace. The hyperplane location is defined by a weighted least squares principle using the specific probabilities as weights. Richman (1983) outlines the mathematical aspects of the rotation which involve finding a transformation matrix based on the hyperplane location to fit the basic rotation equation:

$$\underline{B} = \underline{A}\,\underline{T}^{-1} \qquad (9)$$

where $\underline{B}$ = the matrix of rotated factor loadings (n x r)

$\underline{A}$ = the matrix of unrotated factor loadings (n x r)

$\underline{T}$ = the transformation matrix which fills to properties of simple structure (r x r)

d. Procrustes target rotation

The overwhelming majority of analytic rotations attempt to fulfill the principles of simple structure. These are most useful in trying to identify relationships and have limited use in trying to confirm existing theory. Alternatively, procrustes target rotation (Hurley and Cattell, 1962) is use-

ful in confirmatory applications where a hypothesized factor structure (in the loadings) is expected. The target rotation involves designating the hypothesis as a target and then rotating the unrotated factors to an optimized fit with these loadings. Instead of rotating factors to simple structure, they are transformed linearly to obtain factors having properties as similar as possible to those in the target matrix. The transformation involves a least squares solution between transformed factor weights (i.e., loadings) and the hypothesized factor weights. The goodness-of-fit between these two sets of weights is related to the validity of the hypothesis. Mathematically, the procrustes equation is

$$\underline{B} = \underline{A}\,\underline{T} + \underline{E} \tag{10}$$

where $\underline{B}$ = target matrix (n x r)
$\underline{A}$ = unrotated factor loading matrix (n x r)
$\underline{T}$ = transformation matrix (r x r)
$\underline{E}$ = discrepency of fit matrix or residuals (n x r)

We wish to find a transformation matrix such that:

$$\text{trace } (\underline{E}'\underline{E}) = \text{trace } (\underline{B}-\underline{A}\,\underline{T})' (\underline{B}-\underline{A}\,\underline{T}) \text{ is a minimum} \tag{11}$$

Mulaik (1972) illustrates such a solution by taking the partial derivative with respect to $\underline{T}$ and setting (11) to zero.

$$\frac{\partial}{\partial \underline{T}} (\text{tr } \underline{E}'\underline{E}) = -2\,\underline{A}'\underline{B} + 2\,\underline{A}'\underline{A}\,\underline{T} = 0 \tag{12}$$

This can be rewritten as

$$\underline{A}'\underline{A}\,\underline{T} = \underline{A}'\underline{B} \tag{13}$$

Solving for $\underline{T}$ yields

$$\underline{T} = (\underline{A}'\underline{A})^{-1}\,\underline{A}'\underline{B} \tag{14}$$

$\underline{T}$ is typically normalized so that diag $(\underline{T}'\underline{T}) = \underline{I}$. Analysis of the residuals ($\underline{E}$) is of utmost importance when using procrustes to gauge the validity of the hypothesis.

e. Higher-order factor analysis

Recalling from the development of simple structure rotation that $\underline{\Phi}$ does *not* necessarily equal $\underline{I}_n$ under *oblique rotation*, this typically provides us with the factor score intercorrelation matrix $\underline{\Phi}$ with off-diagonal values. The matrix $\underline{\Phi}$ can be thought of as a correlation matrix (or covariance matrix) between oblique factors. Factoring the matrix $\underline{\Phi}$ is known as a second-order factor analysis. Mathematically, this arises from (8), which can be subscripted as follows

$$\underline{R}_o = \underline{B}_1\,\underline{\Phi}_1\,\underline{B}_1' + \underline{U}_1^2 \tag{15}$$

where $\underline{R}_o$ = the *initial* data correlation matrix
$\underline{B}_1$ = the obliquely rotated factor loadings (pattern matrix)
$\underline{\Phi}_1$ = the interfactor correlation matrix
$\underline{U}_1^2$ = the matrix of uniquenesses

Since $\underline{\Phi}_1$ is a correlation matrix of factor scores, (15) can be rewritten as a *first-order solution*

$$\underline{R}_o = \underline{B}_1\,\underline{R}_1\,\underline{B}_1' + \underline{U}_1^2 \tag{16}$$

Since $\underline{R}_1$ is a correlation matrix, it can be replaced for $\underline{R}_o$ in (16) to yield

$$\underline{R}_1 = \underline{B}_2\,\underline{R}_2\,\underline{B}_2' + \underline{U}_2^2 \tag{17}$$

This is known as a *second-order factor solution*. The process can be repeated iteratively so that a *third-order solution* would be represented by

$$\underline{R}_2 = \underline{B}_3\,\underline{R}_3\,\underline{B}_3' + \underline{U}_3^2 \tag{18}$$

In general terms, anything from a second-order solution on up is known as *higher-order factoring*. The advantages of this approach for geophysical applications are twofold. First, a logical heirarchy of factors is developed by using oblique rotation to *cluster* first-order factors with related scores. Second, the approach yields further data reduction. In the 700-mb analysis, 134 first-order factors were described by only 19 second-order factors. These 19 second-order factors were used to describe the 1664 weekly observations. The procrustes target rotation was then applied to these 19 second-order factors to extend the time-series to the full 32 year data set from 1947-1978.

4. AN EXAMPLE OF THE T-MODE RESULTS

An example of some "typical" T-mode factors (Figs. 4 5,6) will now be shown, to illustrate the complex spatial nature of the technique. The time-series are also plotted in the upper right-hand corner for 4 year averages and a linear-trend test developed by van Loon and Williams (1977, p. 637) was performed. In the lower right hand corner is a reproduction of the Namias (1981) 700-mb catalog which represents the best spatial match between the highest absolute loading on the T-mode factor and the 1.0 correlation point on the Namias teleconnection map. Comparison of the geographical location of the dominant anomaly systems on the two maps provides further evidence of the realistic depiction of the modes of variation at 700-mb. Factor 1 (Fig. 4) is interesting both spatially and temporally. The spatial features are in close agreement with Horel and Wallace's (1981, p. 824) hypothesized Southern Oscillation teleconnection pattern as well as Wallace and Gutzlers (1981, p. 799; 806) Pacific/North American correlation pattern. Horel (1981, p. 2085) also found a similar pattern with his second rotated eigenvector. The Namias catalog match is from his page 14 (20°N; 180° for 1.0 correlation) with excellent seasonal matches in winter and spring. Considering the number of sources using different analysis techniques and finding this basic pattern, factor 1 appears to be a well-defined mode of atmospheric circulation. The time series of factor 1 indicates a well-defined linear trend from 1947-1978 which is significant >99% level. This linear trend indicates that the mean signs of the anomaly patterns have slowly reversed in a linear fashion.

Factor 6 (Fig. 5) is dominated by a large negative anomaly over central North America. There are two compensatory positive anomalies adjacent to this system. In the central Pacific a weaker trough is present, and in the central Atlantic an elongated ridge extends from Cuba northeastward toward Europe. This time series indicates no significant linear trend, although the mode was primarily the sign shown until 1950, then the inverse of what is shown from the 1950's through 1970. During the 1970's the mean pattern was the same as what is shown. The Namias catalog match is from page 152 (50°N; 90°W for 1.0 correlation) and has a noteworthy similarity to the T-mode factor as six of the anomalies have a good spatial (geographical) correspondence to factor 6.

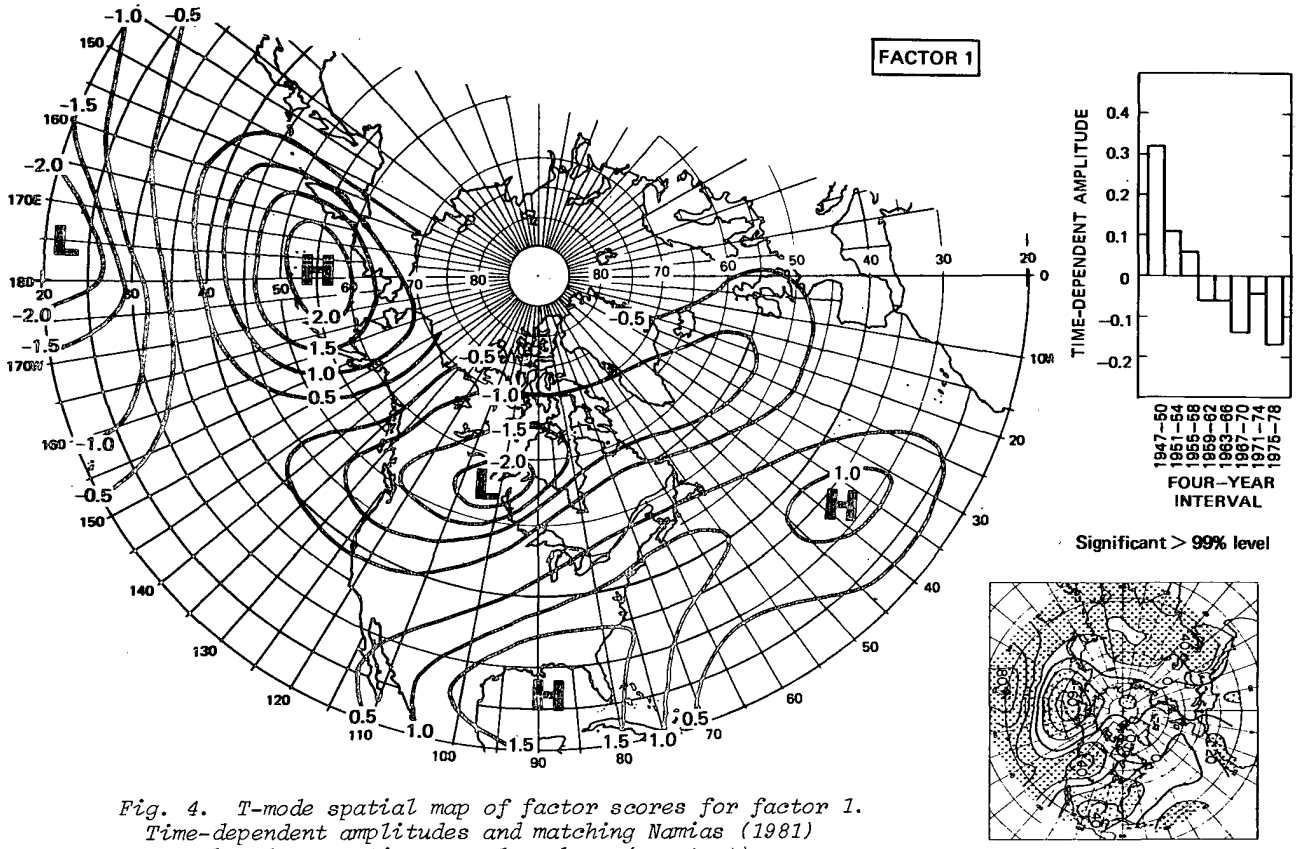


Fig. 4. T-mode spatial map of factor scores for factor 1. Time-dependent amplitudes and matching Namias (1981) 700-mb teleconnection map also shown (see text).

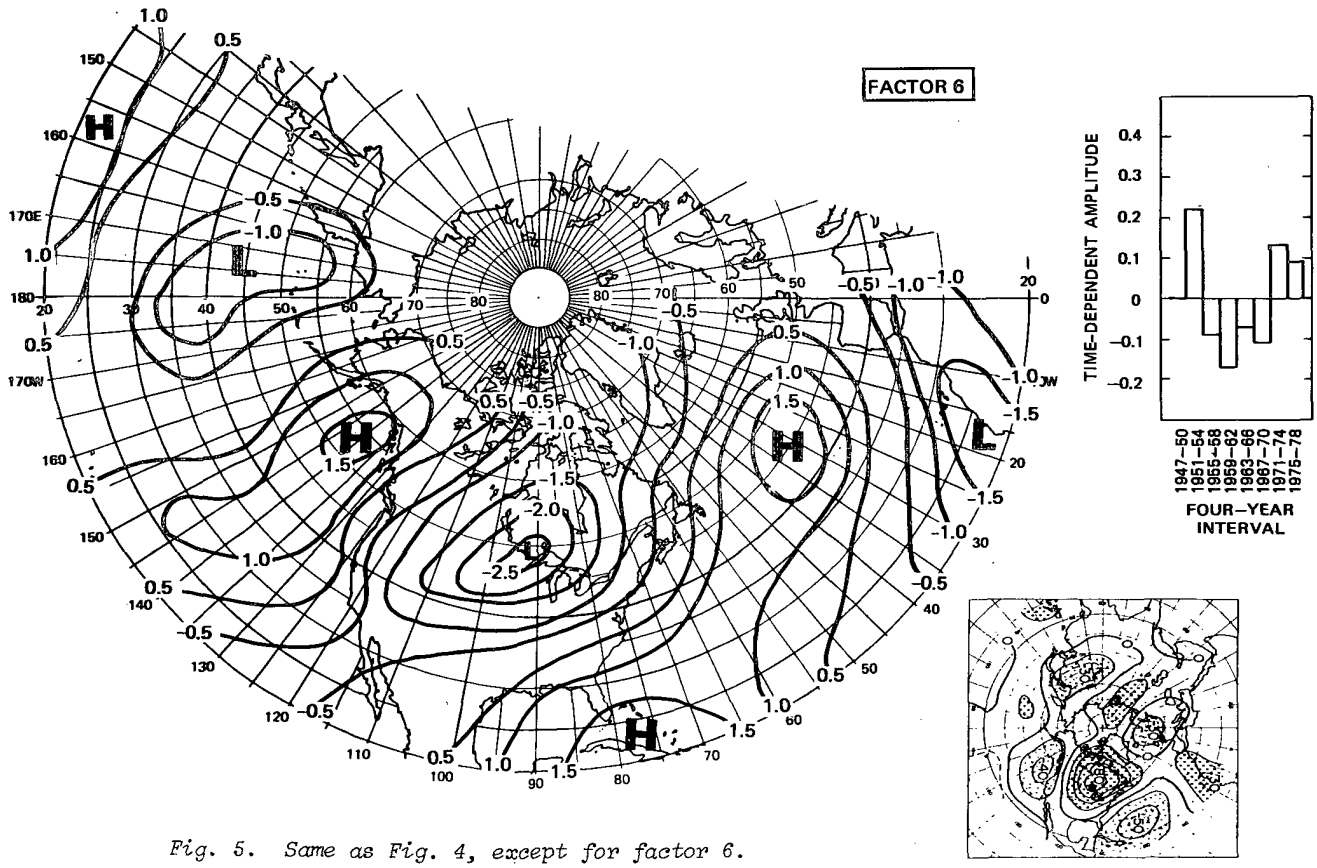


Fig. 5. Same as Fig. 4, except for factor 6.

Factor 13 (Fig. 6) has some similarities to Wallace and Gutzler's (1981, p. 805) western Pacific correlation map. The similarities include an elongated anomaly across the western Pacific and a strong ridge over North America. Differences between this T-mode factor and Wallace and Gutzlers map are that the factor has a much more developed North American anomaly and the Atlantic area is active (while fairly flat in Wallace and Gutzler's map). The Namias catalog match is from page 104 of the atlas (40°N; 100°W) for winter. The time series of the factor indicates a strong linear trend significat at >99% level which corresponds to the mean pattern being the sign shown from 1947-1954 and the inverse of what is shown from 1955-1978.

5. SUMMARY AND FUTURE WORK

Of the six basic modes of decomposition in eigenanalysis, only S-mode has been extensively utilized to filter geopotential height time series into unique modes of circulation. While S-mode with rotation has useful applications (e.g., finding selected isolated grids which frequently have large height fluctuations), its disadvantages when used to depict realistic circulation modes have been noted. The most serious shortcoming lies in the oversimplification of the spatial series which typically results in 1 to 3 anomaly centers per factor due to the simple structure criteria placed on the variables. On a large domain, such as presented herein, this is undesirable as examination of the input data indicate that 75% of the weekly 700-mb maps have 4 - 5 separate anomaly centers. Theoretically, the T-mode rotated approach has intrinsic advantages over S-mode when the domain is larger than one wavelength. The most notable advantage is the *lack of simple structure* in the spatial modes which corresponds to the complex input observations. Furthermore, the time series of the T-mode analysis will exhibit simple structure, facilitating easier interpretation of which observations are associated with specific factors.

The second-order factor analysis was used to combine 134 first order factors into 19 factors which were analyzed. This was accomplished as many of the 134 sets of factor scores had high intercorrelations and were clustered in 19 *more unique* groupings by the higher-order analysis. Therefore, if similar spatial patterns occurred on years 1, 2, 4, 6, 9, and 10 of the analysis, one second-order factor would be formed to describe the six individual factors. The application of procrustes target rotation has been employed to extend the time series of the T-mode factors through 1978. Procrustes offers numerous advantages to data reduction research in a confirmatory mode. Rather than rotating to simple structure, any hypothesis can be utilized and specified as a "target". One potential use might include weather modification research in which the expected modification area could be specified as the "target area" to note how well the stations within this region clustered. Another application might be to specify some useful synoptic pattern found in one analysis as the target and examine if this pattern was prevalent in other time periods or data bases. In T-mode, procrustes can be employed to isolate specific time periods in the target and used to study the effects of quasi-

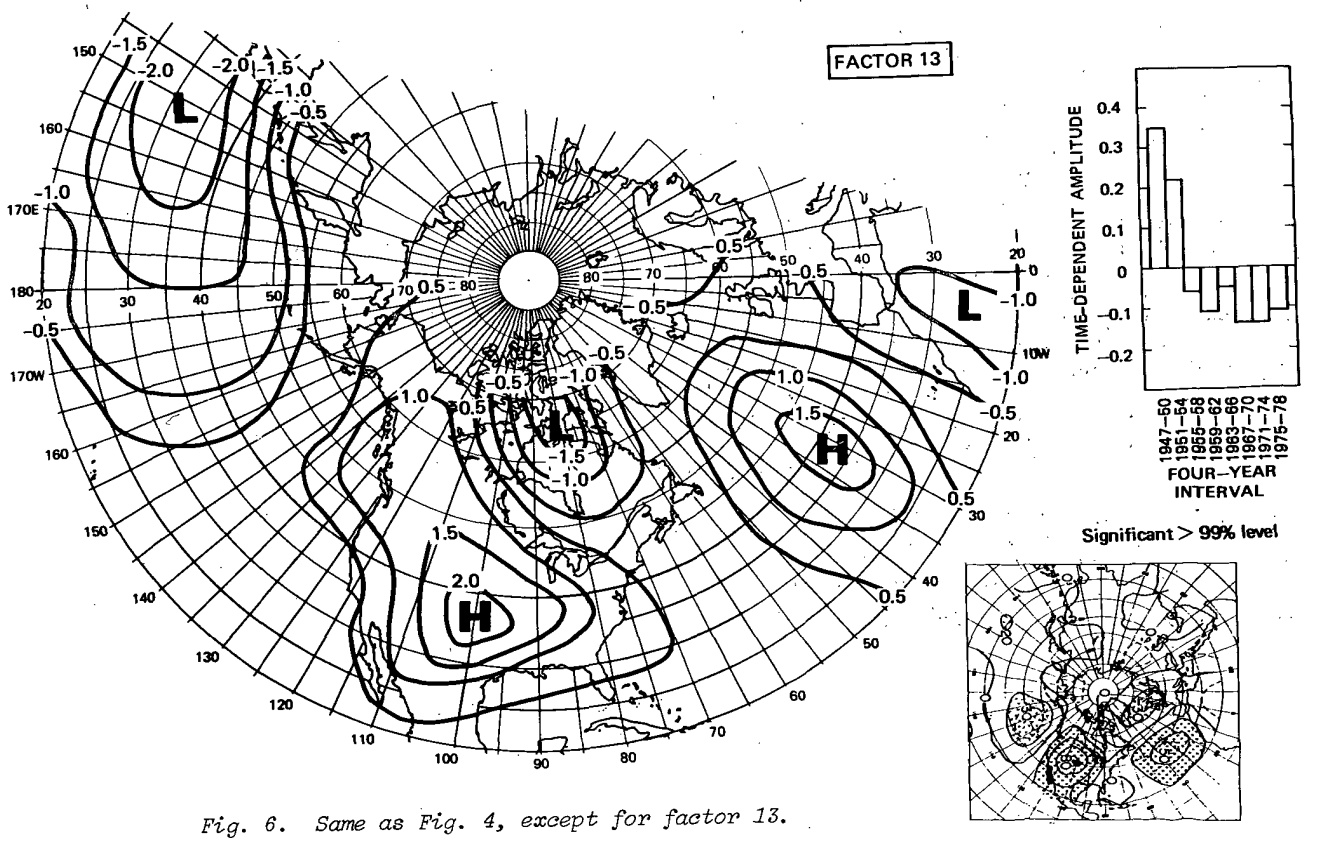


Fig. 6. Same as Fig. 4, except for factor 13.

periodic events (such as El Niño) on meteorological parameters (such as surface temperature). This approach has advantages over harmonics since *aperiodic intervals* can be specified.

6. REFERENCES

Bedi, H.S. and M.M.S. Bindra, 1980: Principal components of monsoon rainfall. Tellus, 32, 296-298.

Brier, G.W. and G.T. Meltesen, 1976: Eigenvector analysis for prediction of time series. J. Appl. Meteor., 15, 1307-1312.

Brinkmann, W.A.R., 1981: Sea-level pressure patterns over eastern North America. Mon. Wea. Rev., 109, 1305-1317.

Buell, C.E., 1975: The topography of empirical orthogonal functions. Preprints Fourth Conf. on Prob. and Stats, in Atmos. Sci., Tallahassee, FL., Amer. Meteor. Soc., 188-193.

Buell, C.E., 1979: On the physical interpretation of empirical orthogonal functions. Preprints Sixth Conf. on Prob. and Stats. in Atmos. Sci., Banff, Alta., Amer. Meteor. Soc., 112-117.

Cattell, R.B., 1952: Factor Analysis. Harper and Row, New York, NY, 462 pp.

Cohen, S.J., 1981: Climatic influences on residential energy consumption. Ph.D. thesis, Univ. of Illinois, 187 pp.

Davis, R.E., 1976: Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. J. Physical Ocean., 6, 249-266.

Diaz, H.F., 1981: Eigenvector analysis of seasonal temperature, precipitation and synoptic-scale system frequency over the contiguous United States. Part II: Spring, summer, fall and annual. Mon. Wea. Rev., 109, 1285-1304.

Diaz, H.F. and D.C. Fulbright, 1981: Eignevector analysis of seasonal temperature, precipitation and synoptic-scale system frequency over the contiguous United States. Part I: Winter. Mon. Wea. Rev., 109, 1267-1284.

Dyer, T.G.J., 1979: Rainfall along the east coast of southern Africa, the southern oscillation and the latitude of the subtropical high pressure belt. Quart. J. Royal Meteor. Soc., 105, 445-451.

Gray, B.M., 1981: On the stability of temperature eigenvector patterns. J. Climatology, 1, 273-281.

Horel, J.D., 1981: A rotated principal component analysis of the interannual variability of the Northern Hemisphere 500 mb height field. Mon. Wea. Rev., 109, 2080-2092.

Horel, J.D. and J.M. Wallace, 1981: Planetary-scale atmospheric phenomena associated with the southern oscillation. Mon. Wea. Rev., 109, 813-829.

Hurley, J.R. and R.B. Cattell, 1962: The Procrustes programs: producing direct rotation to test a hypothesized factor structure. Behavioral Science, 7, 258-262.

Kaiser, H.F., 1958: The varimax criterion for analytic rotation in factor analysis. Psychometrika, 23, 187-200.

Karl, T.R. and A.J. Koscielny, 1982: Drought in the United States. J. Climatology, 2, 313-329.

Karl, T.R., A.J. Koscielny and H.F. Diaz, 1982: Potential errors in the application of principal component (eigenvector) analysis to geophysical data. J. Appl. Meteor., 21, 1183-1186.

Klaus, D., 1978: Spatial distribution and periodicity of mean annual precipitation south of the Sahara. Arch. Meteor. Geoph. Biokl., Ser. B, 26, 17-27.

Lamb, P.J. and M.B. Richman, 1982: Regionalization of central United States for short-period summer rainfall. Proc. Seventh Climate Diagnostics Workshop. Boulder, CO, 180-188.

Lorenz, E.N., 1956: Empirical orthogonal functions and statistical weather prediction. Sci. Rep. 1, Statistical Forecasting Project, Dept. of Meteor., Mass. Institute of Technology, 48 pp.

McGuirk, J.P., 1982: A century of precipitation variability along the Pacific coast of North America and its impact. Climatic Change, 4, 41-56.

Mulaik, S.A., 1972: The Foundation of Factor Analysis. McGraw Hill, New York, NY, 453 pp.

Namias, J., 1981: Teleconnections of 700-mb Height Anomalies for the Northern Hemisphere. CalCOFI Atlas No. 29, Scripps Institution of Oceanography, 265 pp. [Available from Scripps Institute, La Jolla, CA 92093].

Richman, M.B., 1983: The common factor model: Specification of complex modes of circulation with T-mode factor analysis. Mon. Wea. Rev., 111, in press.

Tabony, R.C., 1981: A principal component and spectral analysis of European rainfall. J. Climatology, 1, 283-294.

Thurstone, L.L., 1947: Multiple Factor Analysis. The University of Chicago Press, Chicago, IL, 535 pp.

Tucker, L.R. and C.T. Finkbeiner, 1982: Transformation of factors by artificial personal probability functions. ETS research report 81-58, test and measurement no. TM 820429 [Available from ERIC clearinghouse, Educational Testing Service, Princeton, NJ 08541].

van Loon, H. and J. Williams, 1977: The connection between trends of mean temperature and circulation at the surface. Part IV. Mon. Wea. Rev., 105, 636-647.

Veitch, L.G., 1965: The description of Austrailian pressure fields by principal components. Quart. J. Royal Meteor. Soc., 91, 184-195.

Wallace, J.M. and D.S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. Mon. Wea. Rev., 109, 784-812.

Walsh, J.E. and A. Mostek, 1980: A quantitative analysis of meteorological anomaly patterns over the United States, 1900-1977. Mon. Wea. Rev., 108, 615-630.

Walsh, J.E. and M.B. Richman, 1981: Seasonality in the associations between surface temperature over the United States and the North Pacific Ocean. Mon. Wea. Rev., 109, 767-783.

Willmott, C.J., 1978: P-mode principal component analysis, grouping and precipitation regions in California. Archiv. Meteor. Geoph. Biokl., Ser. B., 26, 277-295.

Statistical Aspects of Empirical Orthogonal Functions Based on Small Sample Sizes

Hans v. Storch
Meteorologisches Institut der Universität Hamburg
Bundesstrasse 55; 2 Hamburg 13; FR Germany

Remark: A large portion (Sections 1-5) of this contribution is taken from a paper "Statistical aspects of estimated principal vectors (EOFs)" by the author and G. Hannoschöck. In the following, to this paper is referred by [SH]. Copies can be ordered from the author.

1. INTRODUCTION

Empirical orthogonal functions (EOFs) are widely used in atmospheric and oceanic research following Lorenz (1959) as an aid to describe climate (e.g. Kutzbach, 1967; Fechner, 1975; Barnett, 1977; Heddinghaus and Kung, 1980; Weare, 1982) or to compare simulations of general circulation models with climate (e.g. Savijärvi, 1978; Storch, 1982; Storch and Fischer, 1983; Storch and Roeckner, 1983a,b; Preisendorfer and Barnett, 1983). They are applied to construct response patterns (Hasselmann, 1979; Hannoschöck, 1983) or a spectral model (Rinne and Karhila, 1975) and as an aid for regression forecast techniques (Peagle and Haslam, 1982a,b).

The algebraic and geometric properties are reported in detail by ECMWF (1977) and their ability to represent meteorological data by Savijärvi (1978).

Preisendorfer and Barnett (1977) and Overland and Preisendorfer (1982) gave a procedure to test random eigenvalues against the hypothesis of an isotropic spectrum. Gray (1981) and North et al. (1982) showed that in case of small sample sizes the structure of the estimated eigenvectors become unstable. Kendall (1980) and Jackson (1981) report some asymptotic stochastic properties. James (1960) published a formula of the joint probability distribution of estimated eigenvalues. However, this formula seems not to be applicable in practice.

Concerning the quality of the eigenvalue estimator or the variances and covariances of the expansion-coefficients of EOFs no publication is known to me.

The concept of EOFs is identical to the concept of principal vectors (Kendall, 1980), which is used in social and economic sciences, especially. The difference in application is that in geophysical research we deal with a larger number of degrees of freedom, whereas the effective sample size is small.

In the following I use the expressions "principal vectors" (PVs) for the eigenvectors of the true 2nd moment matrix and "empirical orthogonal functions" (EOFs) for those obtained from an estimated 2nd moment matrix. Thus, EOFs are estimated PVs.

Estimated quantities are marked by a hat, e.g. a principal component is denoted by $\hat{y}$ and the corresponding EOF by $\hat{\gamma}$.

For the sake of shortness, the proofs of the following statements (2.2-4), (3.3) and (4.3-5) are not given here. They are included in [SH].

2. PRINCIPAL VECTORS

Let $x \in \mathbb{R}^n$ be a random vector. An ordered orthonormal basis of vectors $y_1,\ldots,y_n$ is called a system of principal vectors (PVs), if for every k with $1 \leq k \leq n$

$$E(||\, x - \sum_{i=1}^{k} (x,y_i)y_i \,||^2) = \min \tag{2.1}$$

is valid. E denotes expectation, $||.||$ the least square norm.
The linear form $a_j(x) := (x,y_i)$ is called the ith generalized Fourier coefficient or principal component (of x on y_i).

Let be $X := E(xx')$ the matrix of the second moments of x. Assume the order of the eigenvalues r_i of X as follows:

$$r_1 \geq r_2 \geq \ldots \geq r_n$$

(Note that X is symmetric and positive semidefinite.) Then the statements (2.2-4) are valid:

(2.2) A system of normalized eigenvectors y_i of X ordered like their corresponding eigenvalues forms a PV-system of the random variable x.

$$E(||x-\sum_{i=1}^{k} a_i(x)y_i||^2) = E(||x||^2) - \sum_{i=1}^{k} r_i = \sum_{i=k+1}^{n} r_i \tag{2.3}$$

(2.4) Provided $E(x) = 0$, the generalized Fourier coefficients have zero mean and are uncorrelated, i.e.:

$$E(a_i(x)) = 0 \text{ and } Cov(a_i(x),a_j(x)) = r_i \delta_{ij}$$

These properties are well-known (cf. Kendall, 1980; Jackson, 1981).

Assume $E(x)=0$. Then statement (2.4) means because of $E(||x||^2) = Var(x)$ that

$$Var\ x = \sum_{i=1}^{n} Var(a_i(x)) = \sum_{i=1}^{n} r_i \tag{2.5}$$

As the eigenvalues are non-negative, the Fourier coefficient of the ith PV describes on an average $100*r_i/Var(x)\%$ of the total variance. (2.4) states that the expectation (1st moment) of the Fourier coefficents vanishes and that their covariance matrix (2nd moment) is a diagonal one with the eigenvalues as diagonal elements. For this reason the 1st moments often are estimated before and then subtracted from the data.

3. ESTIMATION

In practice, the distribution of x is unknown. Therefore, one has to estimate the PVs and their eigenvalues by the eigenvectors and -values of an estimated 2nd moment matrix. These estimated PVs are called empirical orthogonal functions (EOFs). For that a set of m samples $x_1,\ldots,x_m$ of x is needed. The resulting matrix is

$$\hat{X} := \frac{1}{m} \left(\sum_{i=1}^{m} x_i x_i' \right) \tag{3.1}$$

If A is the matrix with x_i as i-th column, we have

$$\hat{X} = \frac{1}{m} A\, A' \tag{3.2}$$

If the number n of components of the random vector x is large, it may be difficult to compute the eigenvalues and eigenvectors of the n*n matrix X. But if the number m of samples is less than n, the problem may be reduced by a simple trick, which is based on the statement:

(3.3) If $r \neq 0$ is an eigenvalue of A'A with multiplicity s and eigenvectors $y_1,\ldots y_s$, then r is also an eigenvalue of AA' with the same multiplicity and eigenvectors $Ay_1,\ldots Ay_s$.

AA' is a n*n matrix and A'A a m*m one. A'A has m eigenvalues and AA' n. Each nonzero eigenvalue of A'A is one of AA' according to (3.3), too. It follows that maximally min(n,m) useful EOFs can be estimated.

Therefore, the algorithm is as follows: If $n \leq m$, solve directly

$$\hat{r}_i \hat{y}_i = \hat{X}\, \hat{y}_i \tag{3.4.1}$$

If $n>m$, solve

$$\hat{r}_i\, z = \frac{1}{m} A'A\, z \tag{3.4.2}$$

and define as EOF:

$$\hat{y}_i := \frac{Az}{||Az||} \tag{3.4.3}$$

Often, the data are "centered", i.e. the sample mean is subtracted. In this case only min(n-1,m) useful EOFs can be obtained.

4. THE ESTIMATOR $\hat{r}$

One application of PVs is to obtain characteristic patterns of the underlying ensemble and to attribute to these characteristic patterns their relative importance, i.e. their variance (e.g. Barnett, 1977; Lemke, 1980): If the ensemble mean vanishes and e.g. the first three eigenvectors are attributed with eigenvalues corresponding to, say, 60%, 30% and 5% of the trace of the covariance matrix, this is interpreted by means of (2.4) that the physical system under concern is dominated by two processes describing about 60% and 30% of the total variance.

Table 1

Comparison of the approximate formulae (4.4) and (4.5) with a series of 1000 Monte Carlo trials. Column A: Monte Carlo estimate of bias($\hat{r}$); column B: bias-approximation (4.4); column C: Monte Carlo estimate of the standard deviation of $\hat{r}$; column D: standard deviation according to (4.5); column E: A star denotes rejection of the nullhypothesis of normality at the 95% level.

sample size m	estimator r_1 A	B	C	D	E	estimator r_2 A	B	C	D	E
4	0.49	0.50	1.39	1.41	*	-0.46	-0.50	0.43	0.71	*
6	0.35	0.33	1.15	1.05	*	-0.31	-0.33	0.42	0.58	*
8	0.27	0.25	1.00	0.98	*	-0.27	-0.25	0.36	0.50	*
12	0.19	0.17	0.82	0.79	*	-0.18	-0.17	0.33	0.41	*
50	0.03	0.04	0.40	0.38	*	-0.04	-0.04	0.19	0.20	*
100	0.03	0.02	0.28	0.28		-0.02	-0.02	0.14	0.14	*

The question arises, whether the eigenvalues $\hat{r}$ of the estimated covariance matrix are satisfying estimators of the true eigenvalues r. Asymptotically, $\sqrt{m}(\hat{r} - r)$ is normally distributed with zero mean and variance $2 r^2$ (cf. Jackson, 1981). But, in case of smaller samples sizes, it turns out that, generally, $\hat{r}$ is a considerably biased estimator of r.

$$\begin{aligned} &E(\hat{r}_i) > r_i \text{ for large eigenvalues} \\ &E(\hat{r}_i) < r_i \text{ for small eigenvalues.} \end{aligned} \tag{4.1}$$

This is shown in the following analytically for the case n>m (a) and n=2 (b) and demonstrated by means of an example with n=7 (c).

(a) Case n>m

If m<n, at least n-m eigenvalues of $\hat{X}$ are zero according to (3.3), i.e. for i = m+1, m+2,...n is valid:

$$E(\hat{r}_i) = 0 \tag{4.2}$$

but generally we have $r_i \neq 0$.

(b) Case n=2

Theorem: Let x be a two-dimensional, normally distributed random variable with eigenvalues r_1 and r_2 of its 2nd moment matrix and $r_1 > r_2$. Then

$$E(\hat{r}_1) > r_1 \quad \text{and} \quad E(\hat{r}_2) < r_2 \tag{4.3}$$

Approximate formulae for biases and variances are:

$$\text{bias}(\hat{r}_1) = -\text{bias}(\hat{r}_2) \sim (r_1 r_2)/m(r_1 - r_2) \tag{4.4}$$

$$\text{Var}(\hat{r}_i) \sim \frac{2}{m} r_i^2 \quad \text{(cf. Jackson,1981).} \tag{4.5}$$

In order to check the quality of (4.4) and (4.5), a series of Monte Carlo simulations was done with

$$E(x) = 0 \text{ and } r_1 = 2 , r_2 = 1$$

with varying sample sizes m. Additionally, a Lillifors test (e.g. Conover, 1971) was performed to see whether the eigenvalue-estimators are normally distributed. In Table 1 the bias estimations and estimated standard deviations based on 1000 independent samples each are listed together with the approximations (4.4) and (4.5) and the result of the test of normality at the 95% level of significance. Seemingly, formula (4.4) is a good approximation even for small sample sizes, whereas (4.5) is partly not adequate. If the number of independent samples is 6 times (25 times) the number of degrees of freedom (=2), the larger eigenvalue will be overestimated by about 9% (1.5%), and the smaller one will be underestimated by about 18% (4%), on an average in this example. For sample sizes up to m = 50, the distribution of both eigenvalues is significantly not normal but skew.

(c) Example n=7

In order to see whether the above (b) found properties apply to random vectors with more components, a 7-dimensional example was considered, (a 2nd example is given by [SH]). Without loss of generality it is assumed that the expectation vector vanishes and that the covariance matrix is a diagonal one:

$$E(x) = 0 \quad \text{and} \quad X = \text{diag}(r_1,\ldots,r_7) \tag{4.6}$$

with

$$r_1 = 3.0,\ r_2 = 2.0,\ r_3 = 1.4,\ r_4 = 1.2$$
$$r_5 = 1.0,\ r_6 = 0.8,\ r_7 = 0.6$$

In this example, the relative magnitudes are comparable to what is occurring in practice.

In Table 2 the expectations of the eigenvalue estimators are listed. They were obtained by averaging 200 independent Monte Carlo trials. For n<7, they are given partly by (4.2). To see whether the estimators are significantly biased, the Wilcoxon signed ranks test (e.g. Conover, 1971) was applied to test the hypotheses

$$E(\hat{r}_i) \leqq r_i \quad \text{against} \quad E(\hat{r}_i) > r_i$$

or

$$E(\hat{r}_i) \geqq r_i \quad \text{against} \quad E(\hat{r}_i) < r_i$$

Which of both pairs of (one sided) hypotheses is the appropriate one, was decided by an additional series of independent Monte Carlo trials. A bias significant at the 95% level is denoted by a star in Table 2. As in the preceding example, the distribution of the eigenvalue estimator is tested for normality. An underlined value in Table 2 indicates that the nullhypothesis that the estimator is normally distributed is to be rejected at the 95% level of significance. (Note that

a lacking star or underlining does not mean that there is no bias or that the distribution is in fact normal. It it only the indication that the probability to decide erroneously "accept the alternative" is greater than 5%.) Up to m=15, most of the distributions of the estimators are significantly not normal.

The estimator with index 1 is overestimated and the one with highest index underestimated. This coincides with the result of the 2-dimensional problem. The bias of the first eigenvalue is significant even for m=100, whereas the negative bias seems to become negligible for m=100.

Table 2

Expectation of eigenvalue estimators of example (c) obtained by a series of 200 Monte Carlo trials. If a value is underlined, the respective estimator is significantly not normal, if it is denoted by a star, it is significantly biased (95% level). m is the sample size.

m	eigenvalue to be estimated						
	3.0	2.0	1.4	1.2	1.0	0.8	0.6
4	5.61*	2.60*	1.21	.40*	.0*	.0*	.0*
7	4.73*	2.53*	1.42*	.76*	.35*	.12*	.02*
10	4.36*	2.46*	1.50*	.94*	.52*	.26*	.09*
15	3.95*	2.34*	1.52*	.98*	.64*	.40*	.19*
50	3.31*	2.10*	1.48*	1.13	.86*	.64*	.44*
100	3.16*	2.06*	1.47*	1.17	.93	.72*	.52*

In Table 3, estimated standard deviations of $\hat{r}$ are listed for the same example. Especially for small sample sizes m, the standard deviations are quite large for the low indexed eigenvalues, pointing at the strong variability of these estimators.

Table 3

Sample standard deviation of 200 independent Monte Carlo trials of the eigenvalue estimator $\hat{r}$ applied to example (c).

m	eigenvalue to be estimated						
	3.0	2.0	1.4	1.2	1.0	0.8	0.6
4	3.25	1.18	0.12	-	-	-	-
7	2.32	0.96	0.43	0.16	0.07	0.03	0.01
10	1.98	0.83	0.37	0.17	0.07	0.04	0.02
15	1.83	0.70	0.37	0.16	0.07	0.04	0.02
50	0.96	0.49	0.24	0.12	0.05	0.03	0.02
100	0.70	0.32	0.19	0.08	0.04	0.03	0.01

Unfortunately, I have no idea how to overcome the problem outlined above.

That $\hat{r}>r$ holds for the largest eigenvalue is reasonable: r is the largest amount of the total 2nd moment of the continuous probability space {x} expressible by just one vector and $\hat{r}$ the respective number of the finite $\{x_1,\dots x_m\}$. Since the latter set is smaller than the former set, a bigger portion of its total 2nd moment may be represented by just one vector, i.e. $\hat{r}>r$.

The physical significance of the bias of the eigenvalue estimators is that the importance of the dominating (generally: large-scale) structures is overestimated and that of the less weighty (generally: small-scale) ones underestimated.

5. PROPERTIES OF ESTIMATED PRINCIPAL COMPONENTS

One problem occuring in atmospheric and oceanic research is the comparison of different climates, e.g. of a modelled climate with an observed one or two modelled ones (say: differing by their forcing boundary conditions). This is treated by a multivariate t-test (Storch, 1982; Storch and Fischer, 1983; Storch and Roeckner, 1983a) or by a nonparametric two-sample procedure (Preisendorfer and Barnett, 1983; Storch and Roeckner, 1983b). In order to get a tractable low-dimensional problem, the original data consisting of sometimes hundreds of components are a-priori expanded into a short EOF-series utilizing (2.1). The applied EOFs are obtained from one of the two samples to be compared. Then, instead of the original data set, the statistical test procedure is applied to the low-dimensional vector of generalized Fourier-coefficients ($\hat{a}_i$). For this, reliable estimates of the variances and covariances of the generalized Fourier coefficients $\hat{a}_i$ of a once fixed set of EOFs $\hat{y}_1,\dots\hat{y}_m$ are needed. The question is how to estimate these 2nd moments.

Provided E(x)=0, an obvious and asymptotic reasonable (Jackson, 1981) attempt is to use the properties (2.4) valid for the principal vectors with $\hat{r}$ instead of r, i.e. to estimate the covariances and variances by

$$\hat{\mathrm{Cov}}(\hat{a}_i(x),\hat{a}_j(x)) := \hat{r}_i \delta_{ij} \tag{5.1}$$

(a) Case n=2

In case of 2 dimensions

$$\mathrm{Var}(\hat{a}_1) < E(\hat{r}_1) \text{ and } \mathrm{Var}(\hat{a}_2) > E(\hat{r}_2)$$

is valid and one has to be aware of sometimes large covariances of $\hat{a}$ and $\hat{a}$: As in (b) of Section 4, we may assume without loss of generality that $X = \mathrm{diag}(r_1,r_2)$ with $r_1 > r_2$. Application of (2.1), (2.5) and the Theorem of Section 4 yields

$$\begin{aligned}\mathrm{Var}(\hat{a}_1) &\leq r_1 < E(\hat{r}_1)\\ \mathrm{Var}(\hat{a}_2) &\geq r_2 > E(\hat{r}_2)\end{aligned} \tag{5.2}$$

In order to get an idea of the magnitude of $\mathrm{Var}(\hat{a}_i)-\hat{r}_i$, example (b) of Section 4 was examined again ($r_1=2$, $r_2=1$). For the same numbers m as above, a Monte Carlo series of EOF sets was determined. For each of these EOF sets, the covariance matrix (5.1), $\hat{r}_i$ and $\mathrm{Var}(\hat{a}_i)-\hat{r}_i$ (i=1,2) was computed. Finally, the sample means of $\mathrm{Var}(\hat{a}_i)-\hat{r}_i$ of the 1000 Monte Carlo trials was evaluated. The resulting numbers are given in Table 4: the overestimation of $\mathrm{Var}(\hat{a}_1)$ by $\hat{r}_1$ is large. For e.g. m=8, 8 independent data are available to estimate the 2-dimensional covariance matrix. On an average, we have (cf. Table 1): $\hat{r}_1 = 2.26$ and $\mathrm{Var}(\hat{a}_1) = 2.26-0.48 = 1.78$. If $\hat{r}_1$ is utilized as estimator of $\mathrm{Var}(\hat{a}_1)$, one has to expect an overestimation of 27%. On the other hand, $\mathrm{Var}(\hat{a}_2)$ is underestimated by $\hat{r}_2$ on an average by 41%. A comparison of different samples basing on these variance estimators were conservative for EOF1 and radical for EOF2.

In Table 4 the frequency (of 1000 samples) of

$\mathrm{cov}(\hat{a}_1, \hat{a}_2) \geq 0.25$ is listed additionally. It turns out that the covariance of the Fourier coefficients based on EOFs is not negligible -especially for small samples sizes m.

Table 4
Sample means of $\overline{\mathrm{Var}(\hat{a}_i) - \hat{f}_i}$ (i=1,2) and the frequency of $\mathrm{cov}(\hat{a}_1, \hat{a}_2) \geq 0.25$ for different sample sizes m to estimate the EOFs.

m	$\mathrm{cov}(\hat{a}_1, \hat{a}_2) > 0.25$	$\overline{\mathrm{Var}(\hat{a}_1) - \hat{f}_1}$	$\overline{\mathrm{Var}(\hat{a}_2) - \hat{f}_2}$
4	62.4 %	-0.69	0.76
6	56.9 %	-0.64	0.59
8	55.3 %	-0.48	0.51
12	47.3 %	-0.32	0.36
50	22.2 %	-0.10	0.08
100	7.7 %	-0.05	0.04

(b) Example n=7
A closer look to the 7-dimensional example of Section 4 shows that instead of (5.2) the relations

$$(5.3.1) \qquad E(\mathrm{Var}(\hat{a}_i)) < r_i < E(\hat{f}_i)$$

for small indices i, and

$$(5.3.2) \qquad E(\mathrm{Var}(\hat{a}_i)) > r_i > E(\hat{f}_i)$$

for large indices i, are valid. As an evidence, the biases $\mathrm{Var}(\hat{a}_i) - \hat{f}_i$, i=1,...7, were estimated by 200 independent Monte Carlo trials (Table 5). As in the preceding example, the respective over- and underestimations of the variances are not negligible.

Table 5
Sample means of $\overline{\mathrm{Var}(\hat{a}_i) - \hat{f}_i}$ for i=1,..,7 for example (b), obtained by a series of 200 Monte Carlo trials. If a value is underlined, the respective random variable is significantly not normally distributed, if it is denoted by a star, it is significantly non-zero (95% level).

m	3.0	2.0	1.4	1.2	1.0	0.8	0.6
4	-3.88*	-1.25*	0.24	0.92*	--	--	--
7	-2.66*	-0.87*	0.01	0.51*	0.88*	1.03*	1.10*
10	-2.02*	-0.69*	0.01	0.43*	0.65*	0.83*	0.98*
15	-1.55*	-0.57*	-0.05*	0.25*	0.55*	0.66*	0.80*
50	-0.59*	-0.21*	-0.12*	0.09*	0.22*	0.31*	0.33*
100	-0.33*	-0.12*	-0.11*	0.03*	0.14*	0.17*	0.18*

Another attempt to estimate the 2nd moment matrix could be to use the data $x_1, \ldots, x_m$ which have been already used for computing the EOFs and for estimating the 2nd moment matrix conventionally, i.e. by the sample 2nd moments. Unfortunately, this proceeding is not successful, because the analysis of Section 2 with finite sums instead of expectations yields that the resulting estimator is just (5.1).

The significance of (5.3.1-2) is that a test basing on these variances is conservative for the weighty components and radical for the less important ones. As a consequence, the statistical test mentioned in the beginning of this Section would hardly find the low indexed (generally: large-scale) structures to be different weighted in the compared climates. Instead, high indexed (generally: small-scale) components will be erroneously diagnosed to differ from each other.

For me, a useful approach seems to be to estimate the 2nd moment matrix conventionally utilizing additional, independent data or to use only a part of the total available data set to estimate the EOFs: In this case, the estimation of the PVs is not optimal, but the systematic deviations (5.3.1-2) may be eliminated.

6. SIGNIFICANCE OF EOFS

The principal vectors are uniquely determined if and only if the corresponding eigenvalues are unique. If an eigenvalue r is, say, double with eigenvectors y_1 and y_2, any linear combination of these two is also an eigenvector:

$$(6.1) \qquad X(y_1 + \alpha y_2) = r(y_1 + \alpha y_2)$$

If the multiplicity of the ith eigenvalue is s_i, $\sum s_i = n$, the basic $\mathbb{R}^n$ may be represented as direct sum of s_i-dimensional subspaces H_i. Each of these H_i is spanned by eigenvectors y_{ij}, $j=1,\ldots s_i$, corresponding to r_i.
If $s_i > 1$, the principal vectors are not uniquely determined und thus called "noisy". If $s_i = 1$, the eigenvector is uniquely (besides of a multiplicative constant as -1) determined, i.e. its pattern, distribution of zeros and of maxima/minima may be interpreted. Therefore, PVs with $s_i = 1$ are called "significant". Note that the usage of the word "significant" here has nothing to do with the statistical problem of the valuation of estimated quantities.

The question arises how to decide by means of a finite sample whether the underlying PVs are significant. A statistical test ("Rule N") for this purpose was proposed by Overland and Preisendorfer (1982): the test-statistic is

$$(6.2) \qquad T_j := \hat{f}_j \Big/ \sum_{i=1}^{n} \hat{f}_i$$

with j=1,..n. For each j, the cumulative distribution function for the T_j of a white-noise-spectrum (i.e. $r_i = 1$ for all i) is estimated by a series of Monte Carlo experiments resulting in an estimated 95%-percentile Q_j.
Then, the decision of Rule N is: If $T_j > Q_j$ y_j is declared to be significant. Otherwise y_j is not shown to be not noisy.

"Rule N" is to be criticised for the following reasons:

a) The actual nullhypothesis is $s_1 = p$ (= all PVs are noisy) and the alternative $s_1 < p$ (= at least one PV is significant).

b) If Rule N has classified an eigenvalue as simple, one has only proved the existence of a significant PV. Nothing is said about the quality of its estimate, the EOF. Especially the stability of the respective EOF (cf. North et al., 1982; Gray, 1981) is uncertain.

Thus, I wonder whether an EOF-pattern is interpretable if the respective eigenvalue-estimate has passed Rule N.

c) An EOF is accepted as significant, if T is large. Thus, in case of the first EOF, the chances are improved if the overestimation (4.1) of the first eigenvalue is strong. This seems not reasonable.

d) To test the merits of Rule N practically, it was subjected to a Monte Carlo test (Assmann, 1982): n = 3, m = 500 yields Q_1= 38.7%, Q_2= 34.5% and Q_3= 31.7%. The (true) covariance matrix is X = diag (1.5, 1.0, 0.5). With m=500 samples, the estimated quantities are:

i	$\hat{r}_i$	T_i	$\hat{y}_1$	$\hat{y}_2$	$\hat{y}_3$
1	1.53	51.3%	1.00	-.05	.04
2	0.99	33.1%	.05	1.00	-.05
3	0.46	15.6%	-.04	.06	1.00

In fact, the estimates are good (even if (4.1) is still current) but according to Rule N only EOF 1 is shown to be not noisy.

I doubt that the study of the eigenvalue spectrum is sufficient to state the reliability of the PV-estimates. Additionally, one should take into account the stability of the EOFs.

References:

Assmann, H., (1983), Ein Vergleich der Januar-Punktklimatologien turbulenter Transporte in der wassernahen Genzschicht zwischen Beobachtung und Modell an ausgewählten Positionen im Nordatlantik, Diploma thesis, Hamburg, FRG

Barnett, T.P. (1977), The principal time and space scales of the Pacific trade wind fields, J. Atm. Sc. 34, 221-236

Conover, W.J. (1971), Practical nonparametric statistics, Wiley New York, London, Sidney, Toronto

ECMWF (1977), Workshop on the use of empirical orthogonal functions in meteorology, European Center for Medium Range Weather Forecast, Reading, Great Britain

Fechner, H. (1975), Die Darstellung des Geopotentials der 500mb-Fläche der winterlichen Nordhalbkugel durch natürliche Orthogonalfunktionen, Ber.Inst.f.Meeresk. Kiel, 5

Gray, M. (1981), On the stability of temperature eigenvector patterns, J. of Climatology 1, 273-281

Hannoschöck, G. (1983), Response of the GISS circulation model to a prescribed sea surface temperature anomaly, PhD thesis, Hamburg, FRG

Hasselmann, K. (1979), On the signal-to-noise problem in atmospheric response studies, Meteorology of Tropical Oceans, Roy.Met.Soc.

Heddingshaus, T.R., Kung, E.C. (1980), An analysis of climatological patterns of the northern hemispheric circulation, Mon.Wea. Rev.108,1-17

Jackson, E. (1981), Principal components and factor analysis: part II - additional topics related to principal components, J. Quality Techn. 13, 46-58

James, A.T. (1960), The distribution of the latent roots of the covariance matrix, Ann.Math.Stat. 31,151-158

Kendall, M. (1980), Multivariate analysis, Charles Griffin Co, London and Wycombe, 2nd edition

Kutzbach, J.E. (1967), Empirical eigenvectors of sea level pressure, surface temperature und precipitation complexes over North America, J. Appl. Met. 6, 791-802

Lemke, P. (1980), Applications of the inverse modelling technique to Arctic and Antarctic Sea ice anomalies, Hamb.Geoph.Einzelschr.49

Lorenz, E.N. (1959), Empirical orthogonal functions and statistical weather prediction, in: Final report statistical forecasting project

North, G.; Bell, T.L.; Cahalan R.F.; Moeng, F.J. (1982), Sampling errors in the estimation of empirical orthogonal functions, Mon.Wea.Rev.110, 699-706

Overland, J. E.; Preisendorfer, R.W. (1982), A significance test for principal components applied to a cyclone climatology, Mon.Wea. Rev.110, 1-4

Peagle, J.N.; Haslam, R.B. (1982a), Empirical orthogonal function estimates of local predictability, J.Appl.Met.21, 117-126

Peagle, J.N.; Haslam, R.B. (1982b), Statistical prediction of 500 mb height field using eigenvectors, J.Appl.Met.21, 127-138

Preisendorfer, R.W., Barnett , T.P. (1977), Significance tests for empirical orthogonal functions, Proceedings Am.Met.Soc. 5th Conf. on Probability and Statistics in Atmospheric Sciences, 169-172

Preisendorfer, R.W.; Barnett, T.P. (1983), Numerical model/reality intercomparisons tests using small-sample statistics, subm. J. Atm.Sc.

Rinne, J., Karhila, V. (1975), A spectral barotropic model in horizontal orthogonal functions, Quart.J.R.Met.Soc.101,365-382

Savijärvi, H. (1978), Verification and storing with empirical orthogonal functions, ECMWF Internal Report 18

Storch, H.v. (1982), Comparison of a sequence of model generated 500mb-topografies with climate, Tellus 34, 89-91

Storch, H.v.; Fischer I. (1983), An analysis of geopotential at 300 mbar in the frequency wavenumber domain along 50N in observed and modelled January climate, Beitr. Phys. d. Atm. 56, (in press)

Storch, H.v.; Roeckner, E. (1983a), A strategy to assess the quality of GCM simulations with applications to a January simulation with the Hamburg University model, submitted to Mon. Wea. Rev.

Storch, H.v.; Roeckner, E. (1983b), On the verification of January GCM simulations, this issue

Weare, B.C. (1982), El Niño and tropical pacific ocean surface temperature, J.Phys.Oceano.12,17-27

METHODOLOGICAL ASPECTS OF PRINCIPAL COMPONENT ANALYSIS IN METEOROLOGICAL FIELDS

Walter Mario Vargas

Facultad de Ciencias Exactas y Naturales, Dpto. de Meteorología
Universidad de Buenos Aires
1428 Nuñez, Buenos Aires, Argentina

Rosa H. Compagnucci

Instituto Argentino de Nivología y Glaciología (IANIGLA)
C.C.330, 5500 Mendoza, Argentina

1. INTRODUCTION

One of the problems climatology must solve is the synthesis of weather types. Subsets of data patterns which occur more frequently are found with map typing, thus summarizing the essential information of a large set.

In order to resolve this problem many ways were used in synoptic-climatology. Several works have been done with this aim specially in the last years; Barry and Perry (1973) give wide examples of appropriate methods.

The multivariate data analysis provides different classification methods for meteorological patterns. McCutchan and Schroeder (1973) used discrimanant analysis of eight meteorological variables to classify the days, Lorenz (1956) states the use of empirical orthogonal functions for statistical weather predictions, and Lund (1963) selected a map-pattern by choosing the most frequent type in the correlations matrix, where he clusters days as opposed to Williams and Terjung (1981) who used eigenvectors to create map types by clustering stations (covariances between the grid-point data).

Through factor analysis -empirical orthogonal functions, principal components or eigenvectors- different internal relations of multivariate samples are synthetized and it is possible to create map types that extract the maximum variance from the data set.

According to Green P.E. (1978) one could consider two ways of resolving the input matrix, correlating pairs of fixed points over occasions: S-Mode; or correlating pairs of occasions over fixed points: T-Mode, in each case this is done over a single variable.

The aim of this paper is to investigate which of the two ways is best for typification of daily pressure map sets in a grid-point where the S-Mode, in our case, clusters stations and the T-Mode clusters days.

The results of applying both options in a meteorological plasmode, or types of sea level pressure with previously known configurations, are obtained. Tests over real samples are done too, showing in this paper just those results that best fit to the problem of obtaining map patterns and also allowing the investigator to recognize the days in the real input sample pointed out by the map type.

2. DATA

2.1 Plasmode

According to Cattell and Sullivan (1962) a plasmode is a set of variables with previously known internal relations that constitutes an elementary sample to which we apply the analysis that we wish to examine.

In order to resolve which of the two options S-Mode or T-Mode, describe more widely, the different input data characteristics, we analyse this by introducing controlled patterns in the input matrix.

The plasmode (Fig.1) contains simple pressure fields which were placed into a 6x6 grid (36 stations). The set is formed by three basic fields one with meridional flow, a second with zonal flow a third with pressure centers (low or high) and their corresponding inverse.

Another four cases differing slightly from its patterns, in a mb range, for each grid-point, were created for each basic field. By these means we intend the introduction of variability in order to avoid a singular matrix.

For each one of the 15 cases, formed by the three basic fields, their inverse is made in such a manner that for each point of the grid the average between the basic field and its corresponding inverse gives a value equal to 1012 mb.

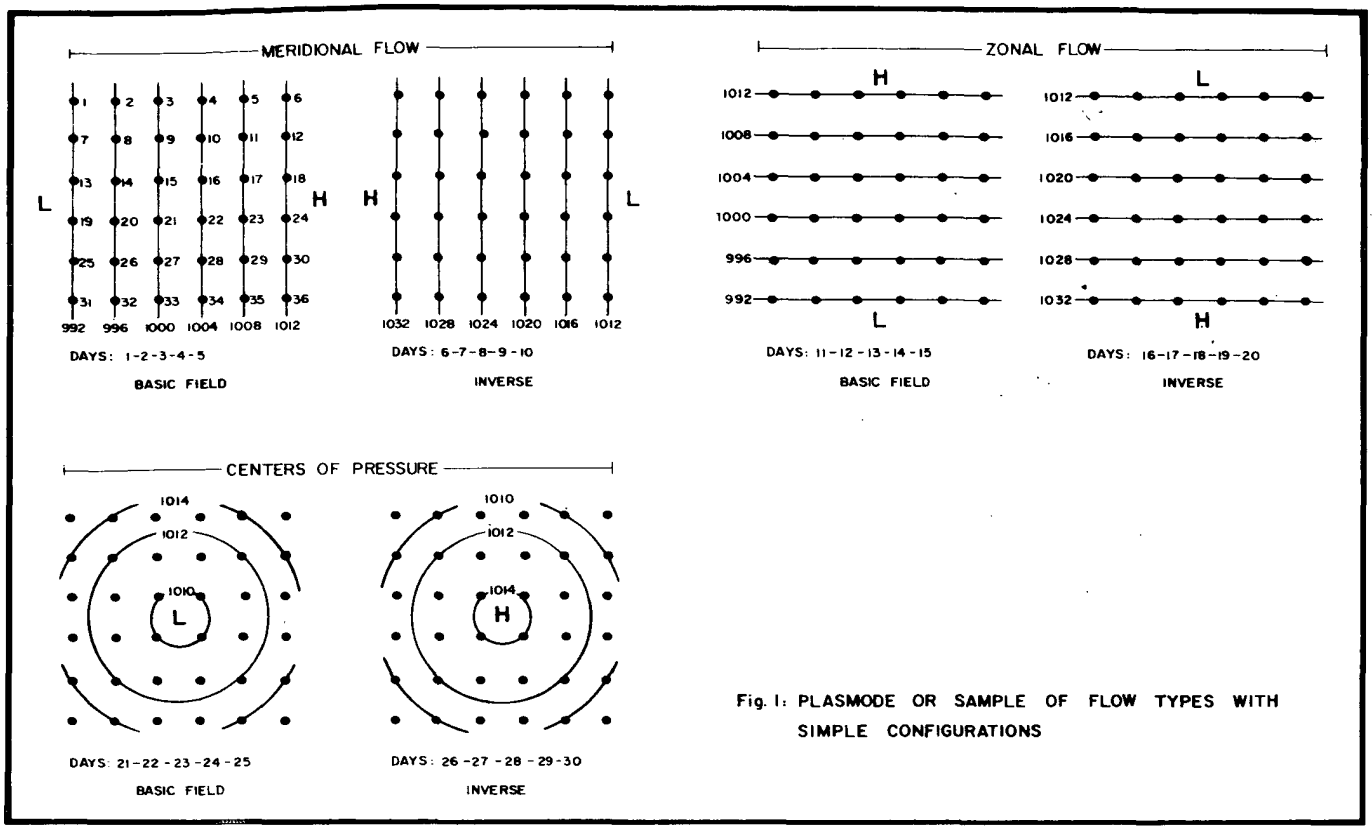


Fig. 1: PLASMODE OR SAMPLE OF FLOW TYPES WITH SIMPLE CONFIGURATIONS

Therefore, the mean pressure level of the plasmode is represented in a flat field of 1012mb. for all grid-point.

In brief, the plasmode is formed by 30 cases (days) containing 36 grid-points each.

2.2 Real Samples

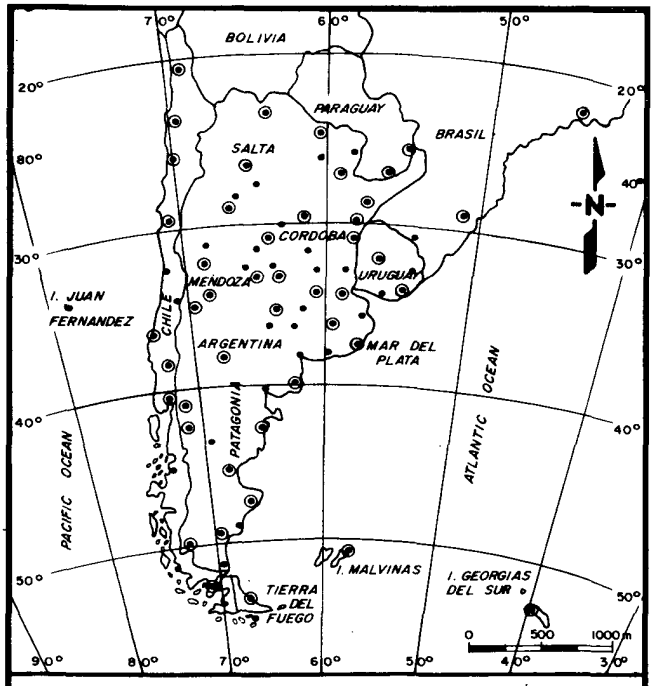
Pressure values of synoptic maps from the National Meteorological Service (Servicio Meteorológico Nacional -S.M.N.- Fuerza Aérea Argentina, El Plumerillo, Mendoza) were used. The stations taken into account are whown on the map.

The samples studied include 15 days of July 1977, 25 days of August 1976 and 27 days of August 1977. In the cases considered, many examples of synoptic situations are presented. Some of them are very important due to the significant associated phenomena in different areas of the zone under survey.

Regarding the file it is necessary to point out that the network stations have incomplete data series. Therefore we may consider the following solutions:

a) To take the days where some data are missing off the sample. Under these circumstances a considerable percentage of cases are eliminated; if one wants to study a special sample to analyse its characteristics in this case, the conclusions would be impoverished.

Generally the less frequent situations are



MAP 1: SOUTH END OF SOUTH AMERICA

the ones that give place to more interesting phenomena for study. We also know by experience that due to different reasons, it is in these cases when data for some stations are lacking.

b) Another possibility is to withdraw stations with incomplete series from the network.

As not always the ones that have lost the information are the same stations, doing the afore said would reduce the grid-density, and therefore risk a subclassification of the types.

Usually the percentage of missing data in a particular day is small and often it is a differente stations the one with lack of data. In this way, the interpolation is done on more or less randomly distributed grid-points.

If we have plenty grid-density giving redundant information, interpolating missing data does not affect the final result.

According to all this we decide to interpolate the missing data from the pressure field and the stations' time series.

3. ANALYSIS TECNIQUE

Let the initial P number of observations by N number of variables data matrix be designated where one of its element is d_{ij} with i= 1....P and j= 1....N.

The deviations matrix **IB** with respect to the mean of each variable can be expressed as:

$$(1)\quad IB = \left\{ b_{ij} = d_{ij} - \bar{d}_{(i)} \right\}$$

$$(2)\quad \bar{d}_{(j)} = \frac{1}{P} \sum_{i=1}^{P} d_{ij}$$

From **IB** the correlation matrix **IR** can be calculated as the following:

$$(3)\quad IR = \sum^{-1} IB^{T}\ IB \sum^{-1}$$

where $\sum$ is a diagonal matrix with dispersions.

Eigenvalues are represented by the diagonal matrix λ from the solution of $\det(IR = \lambda\, II) = 0$, resolved by Jacobi's method (Lorenz,1956). The elements of λ are otherwise known as the scalar characteristics values (λ_i) of linear transformation of **IB**, **II** is the identity matrix.

Eigenvectors matrix **/A** is obtained; it has N eigenvectors columns, each containing N terms.

The obtained eigenvectors are orthogonal and normalized, and arranged according to eigenvalues magnitude order.

$$(4)\quad \sum_{i=1}^{N} \lambda_i = N \quad ; \quad \frac{\lambda_i}{N}$$

{ represented the proportions of variance account for the eigenvector.

The multipliers matrix **IM** is formed by P multipliers row of N elements each.

$$(5)\quad IM = ID \times /A$$

3.1. S-Mode, taking values in the stations as variables and days as observations (clustering stations)

Here d_{ij}, element of the input matrix , represents i^{th}-days for j^{th}-stations. The correlation matrix **IR** yields the interdependence between the stations for different times (days).

The deviation matrix **IB** is expressed as:

$$(6)\quad IB = \left\{ b_{ij} = d_{ij} - \bar{d}_{t\,(j)} \right\}$$

$\bar{d}_t(j)$ is the average for each station over the days.

3.2. T-Mode, taking days as variables and stations values as observations. (clustering days).

Here d_{ij} represents the value in i^{th}-station for j^{th}-day. The interdependence between the days is given by **IR** matrix.

The deviation matrix **IB** is expressed as:

$$(7)\quad IB = \left\{ b_{ij} = d_{ij} - \bar{d}_{e\,(i)} \right\}$$

$\bar{d}_e(j)$ is the average for each day over the stations.

4. RESULTS AND DISCUSSION

4.1. Type patterns for the plasmode through S-Mode

In this case the patterns are drawn by analyzing the values for each grid-point for the corresponding eigenvectors.

The diagram of the j^{th}-element of all the P multipliers corresponding to the j^{th}-eigenvector is the function days-multipliers.

Here the deviation matrix **IB** is expressed by the equation (6) and where $\bar{d}_t(j)$=1012mb. for our plasmode.

In Fig.2 the results of the first three eigenvalues are shown, displaying the patterns of the eigenvectors and multipliers-days functions. These eigenvectors account for 98.77% of the variance which is significant according to Kaiser's test (1959); the only significant eigenvectors are those whose eigenvalues are ≥ 1. In our case the first eigenvalues in order are: 26.35; 7.28; 1.92; 0.14; etc.; where the fourth eigenvector is not significant because its eigenvalue is lower than one.

It can be seen in the Fig.2 that only the 3rd eigenvector shows a pattern similar to the basic field of the input matrix. This type resembles the days 21 to 30 of the plasmode that corresponds to high and low pressure centers, with

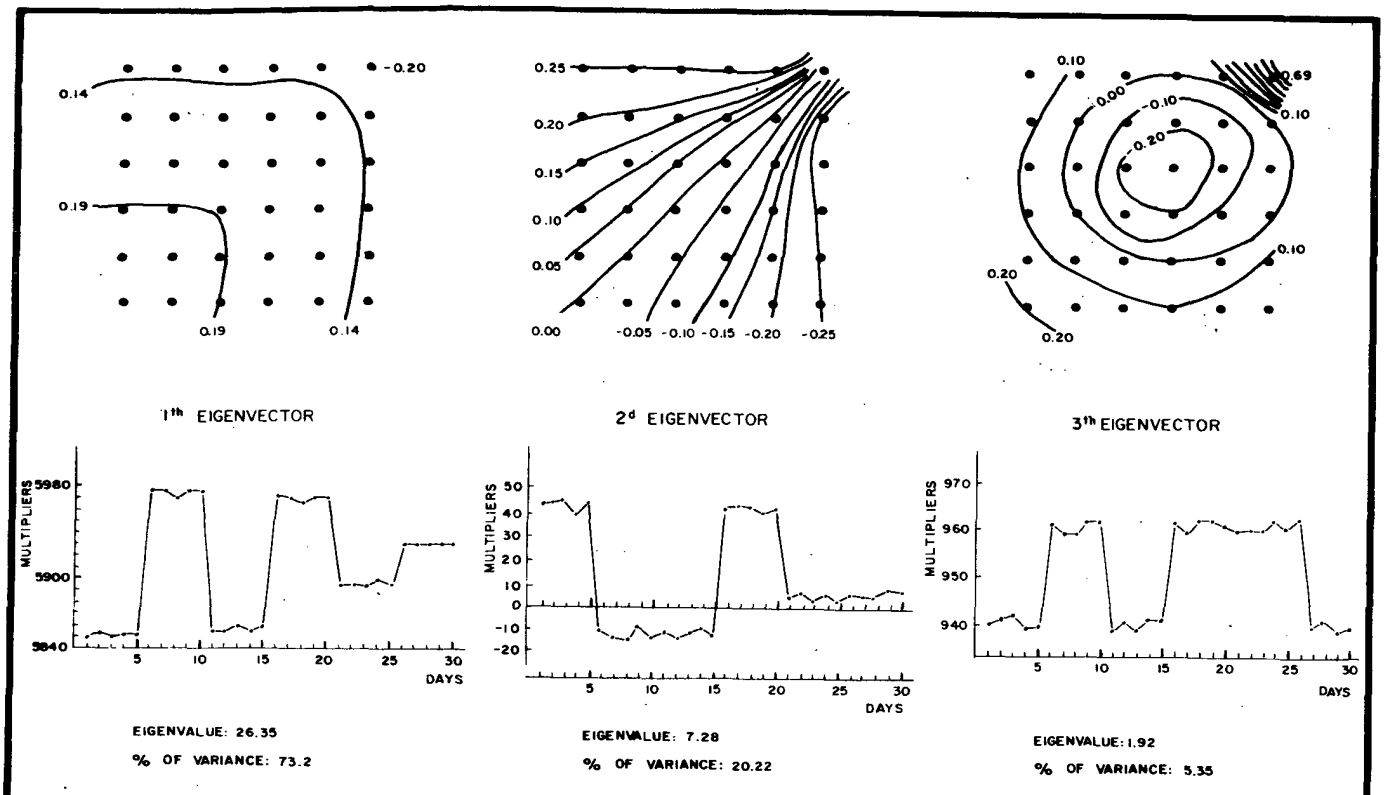


Fig. 2: S-MODE RESULTS; THE PATTERNS (UPPER) ARE GIVEN BY THE EIGENVECTORS, THE TIME SERIES ARE GIVE BY THE FUNCTIONS MULTIPLIERS-DAYS (LOWER).

a very strong gradient difference in the northeast end of the field, unexisting in the input fields.

The patterns given by the 1st and 2nd eigenvectors does not resemble any of the basic fields in the plasmode.

The station entitled '6' (northeast end of the grid-point) presents an anomalous value with respect to the field general configurations in the cases of the 1st and 2nd eigenvectors, thus generating a very strong gradient in that end; the second eigenvector is shown in that area as a strong convergence (divergence).

The functions multipliers-days, in the S-Mode do not indicate a realistic looking time series, and it is impossible to recover the days of the plasmode that belong to the patterns yielded by the eigenvectors.

A last inference is that the variance percentage explained by the eigenvectors, in this case, does not keep a relation with the frequency existing in the initial sample of the patterns.

4.2. Type patterns for the plasmode through T-Mode

The j^{th} pattern in this method is obtained by analyzing in the grid-point, the values for the j^{th}-element of the P multipliers in accordance with the j^{th}-eigenvector, and the time series diagram is drawn from the function eigenvector-days for the j^{th}-eigenvector.

For the deviation matrix the equation (7) is used, where $\bar{d}_e(j)$ is not constant for this plasmode.

The results for the first three eigenvalues are shown in Fig. 3, displaying the patterns and the time series recovered by the function eigenvector-days.

According to Kaiser's test (1959) and taking into account the output patterns configurations, only the first three eigenvectors are significant. The fourth eigenvector whose results are not shown in the paper, displays a completely erratic field.

The first pattern (accounting for 33.46% of the variance) shows a field with a more or less homogeneous gradient and with a strong meridional flow component quite assimilable to the days 1 to 10 of the plasmode.

The second pattern (accounting for 32.89% of the variance) displays a mainly zonal flow more or less homogeneous too. This is quite assimilable to the days 11 to 20.

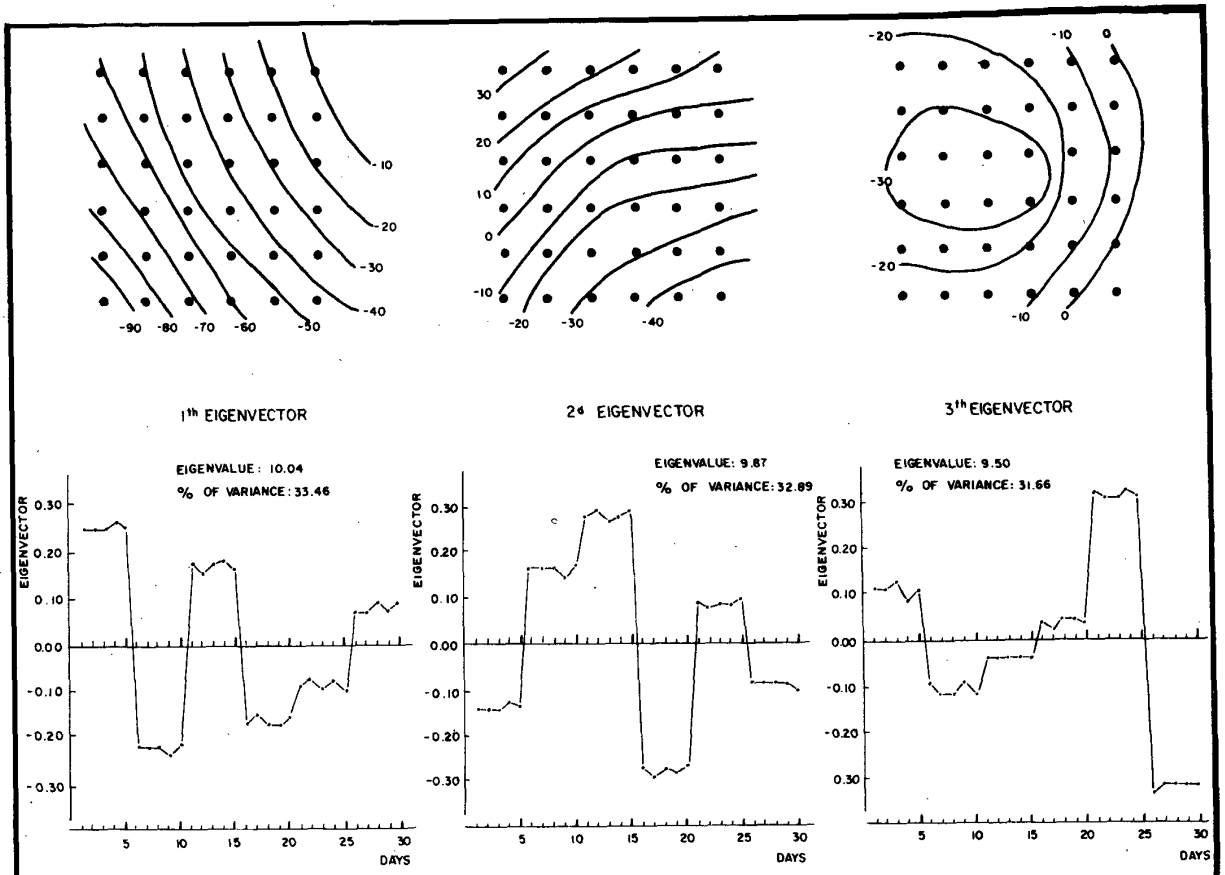


Fig. 3: T-MODE RESULTS; THE PATTERNS (UPPER) ARE GIVEN BY THE MULTIPLIERS, THE TIME SERIES ARE GIVEN BY THE FUNCTIONS EIGENVECTOR-DAYS (LOWER)

The last significant pattern (accounting for 31.66% of the variance) gives a pressure center (low or high)located similarly to those of the days 21 to 30, showing a gradient difference with respect to the basic field in the input matrix.

The functions eigenvector-days do not appear to isolate the time periods very well, but the maximums (minimums) recover the days corresponding to their type patterns.

In the plasmode, each of the three basic fields -meridional flow, zonal flow, pressure centers (low and high)- represents 33.3% of the sample in each case. In the T-Mode the eigenvectors account for 33.41%, 32.89% and 31.66% of the total variance respectively. Therefore the variance percentages with this method represent the frequency of the patterns in the input matrix.

4.3. Real Sample

As the results of the T-Mode proved to be better than the S-Mode to determine map types contained in a sample, only the results of the first method are shown.

4.3.1. Sensitiveness Tests:

a. Sensitiveness of the method studied with respect to the grid density.

The sample used consists of 19 days of July 1977. The eigenvectors are calculated from the grid composed by 77 stations showed on the map(1) with a later extraction of 31 stations, remaining the 46 marked on the map by doble circle.In both cases the results are almost the same for 84.3% of the variance, inferring that the grid of 77 stations give redundant information.

b. Sensitiveness of the T-Mode with respect to the reduction of the area under survey.

The eigenvectors of the grid are calculated for an initial density of 77 stations, and they are compared with the results of removing the information of Chile. The results of the eigenvectors in a situation of minor stations density (47) are also compared with the results of removing information of Chile. In both cases only 50% of the variance give similar results, inferring that only 50% of the stations is defined independently of the information given by the Chilean stations. Therefore, the system is independent of the Chilean field in that percentage, or is taken as a constant by the method.

c. Sensitiviness of the eigenvectors value with respect to different sizes of the samples.

The first sample consists of August 1976,

the second, August 1977, and the third a group of days of both months.

Analyzing the values obtained in the three situations (Fig.4) it can be inferred that for this case the relative value of the eigenvectors is almost independent of the samples size. This occurs for a high percentage of the variance that is the accumulated by the first three eigenvectors as it can be seen by the results.

The value of the first eigenvector is small in all the days, in the total sample. This would be caused by a major size of the sample.

From the fourth eigenvector on, the above is not evident, therefore between 20 and 25% of the variance the independence of the samples size is not valid.

4.3.2. Map type in a real sample

The results obtained for August 1976 are shown in this paper, displaying only the patterns and time series of the significant eigenvectors according to Kaiser's test (1959).

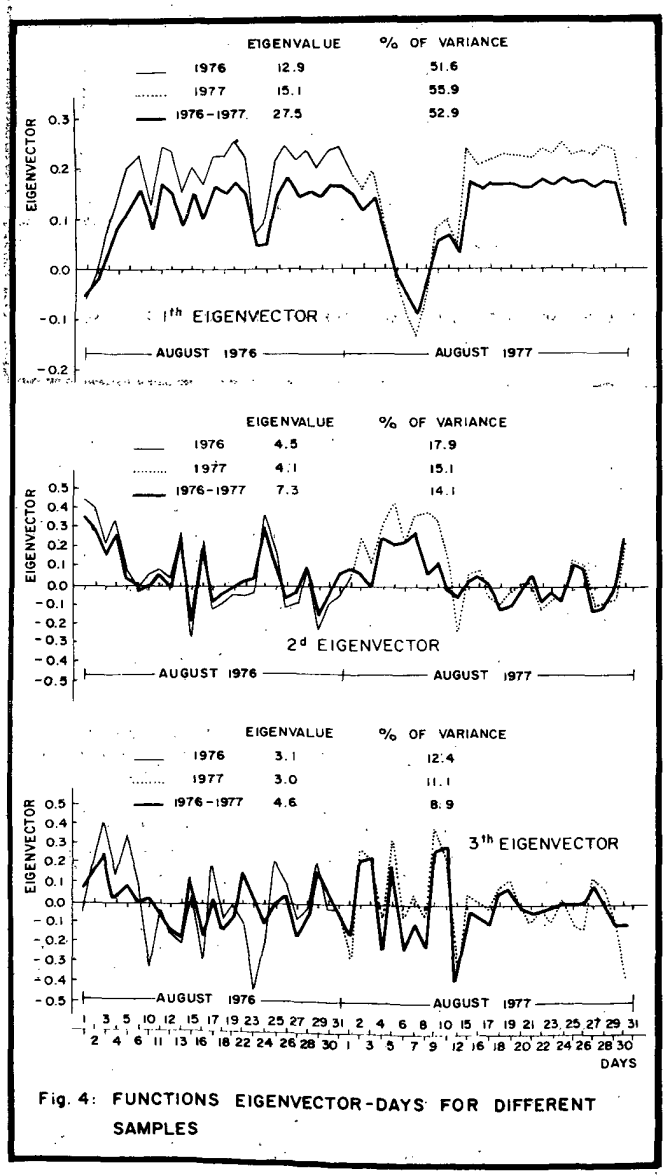


Fig. 4: FUNCTIONS EIGENVECTOR-DAYS FOR DIFFERENT SAMPLES

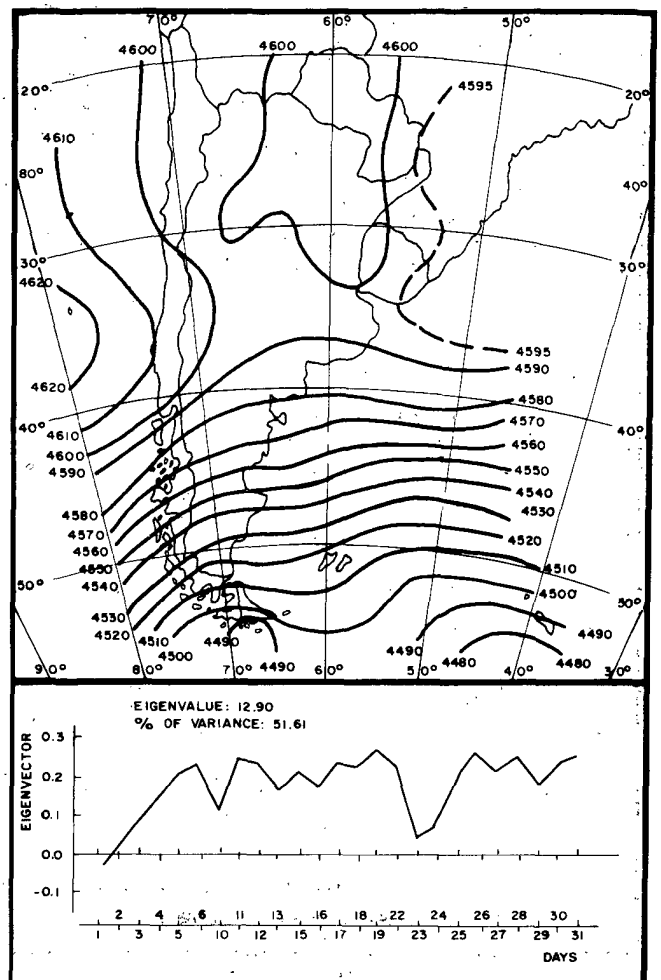


Fig. 5: FIRST PATTERN OF AUGUST 1976 (UPPER) AND THE FUNCTIONS EIGENVECTOR-DAYS (LOWER)

The first pattern (Fig.5) accounts for 51.8% of the variance and represents: a west flow with a considerable gradient south of parallel 40S; a low pressure south of Tierra del Fuego; a low pressure than occupies the north-east of Argentina; the Atlantic and Pacific semipermanet anticyclone north of parallel 40S; the Pacific anticylone entering Mendoza and the Atlantic one softly sketched; a slight gradient north of parallel 40S.

This situation is the most frequent en August 1976. The more similar days to this theoretical map are those corresponding to the maximums of the eigenvector's function with respect to the days (lower end of Fig.1). In this particular case the days 11, 18, 26, 28, 30 and 31, have this property. This was verified consulting the synoptic maps of the S.M.N. (National Meteorological Service).

The inverse situation is described as inprobable: a high pressure center south of Tierra del Fuego; low pressure instead of high semipermanent pressures of the Atlantic and Pacific; high pressure in the north-east and strong gradient in medium latitudes. This is why there are no minimums

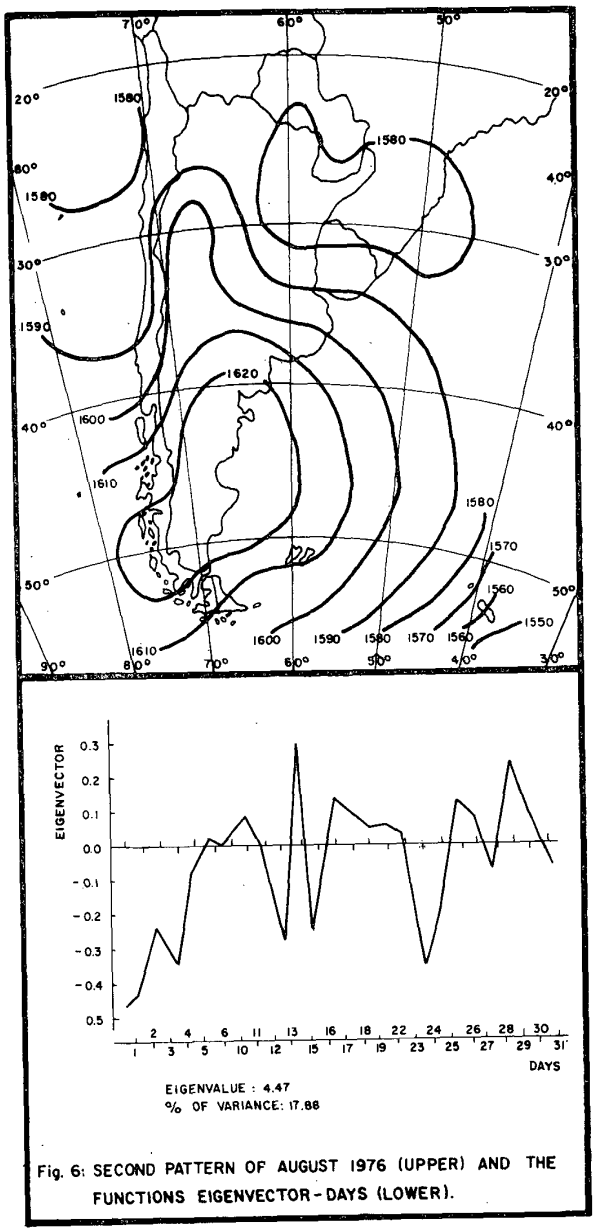


Fig. 6: SECOND PATTERN OF AUGUST 1976 (UPPER) AND THE FUNCTIONS EIGENVECTOR-DAYS (LOWER).

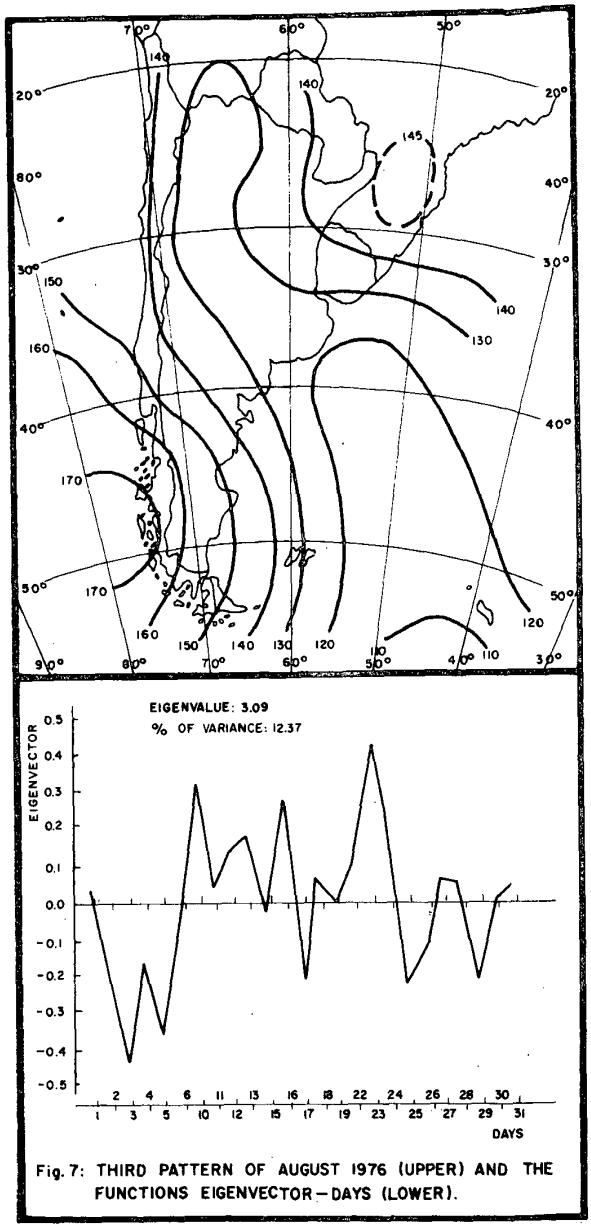


Fig. 7: THIRD PATTERN OF AUGUST 1976 (UPPER) AND THE FUNCTIONS EIGENVECTOR-DAYS (LOWER).

values of equal modulus and different sign as the maximums in the function eigenvector-days.

The second pattern (fig.6) accounts for 17.9 of the variance and represents: a high pressure entering between parallels 40S and 50S that affects all the area south of parallel 30S; a low one south of Brazil and extreme north of Argentina a low one in the west coast north of Chile. This situation corresponds to the minimums of the function eigenvector-days (lower end of Fig.6) and the dates 1, 2, 4 and 24 are the ones that more clearly present this theoretical configuration.

The inverse would be given by the maximums of the function eigenvector-days, in this case. This means a basic field with the following characteristics: a low pressure center in Patagonia; a high one north-east of the area and in the Pacific, north of parallel 30S. The days 15 and 29 are the ones that are most assimilable to this case.

The third pattern (Fig.7) accounts for 12.4% of the variance. The theoretical map represents an axe of low pressures through Salta, Córdoba, Mar del Plata, and 50S-45W; a high pressure entering through the Chilean coast between parallels 40S and 50S; a high pressure center south of Brazil. This corresponds to the maximums of the function eigenvector-days (lower end of Fig.7) and the more assimilable cases are the days 10 and 23.

The inverse: an axe of high pressures, a low pressure south of Brazil, and another entering in Patagonia, corresponding to the minimums of the graphic eigenvector-days, and the more assimilable days are 3 and 5.

The fourth pattern (Fig.8) accounts for 5.6% of the variance and represents: low pressures north of Argentina and south of Tierra del Fuego; the well defined Atlantic anticyclone with its area of influence in all the south Atlantic

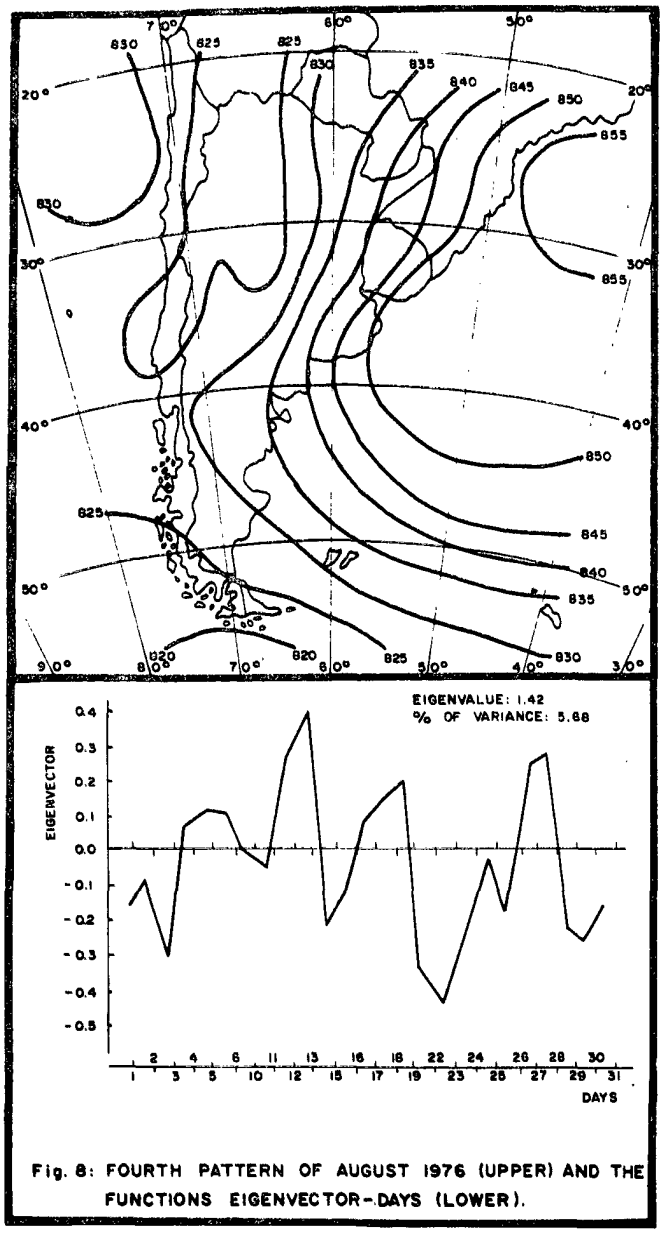


Fig. 8: FOURTH PATTERN OF AUGUST 1976 (UPPER) AND THE FUNCTIONS EIGENVECTOR-DAYS (LOWER).

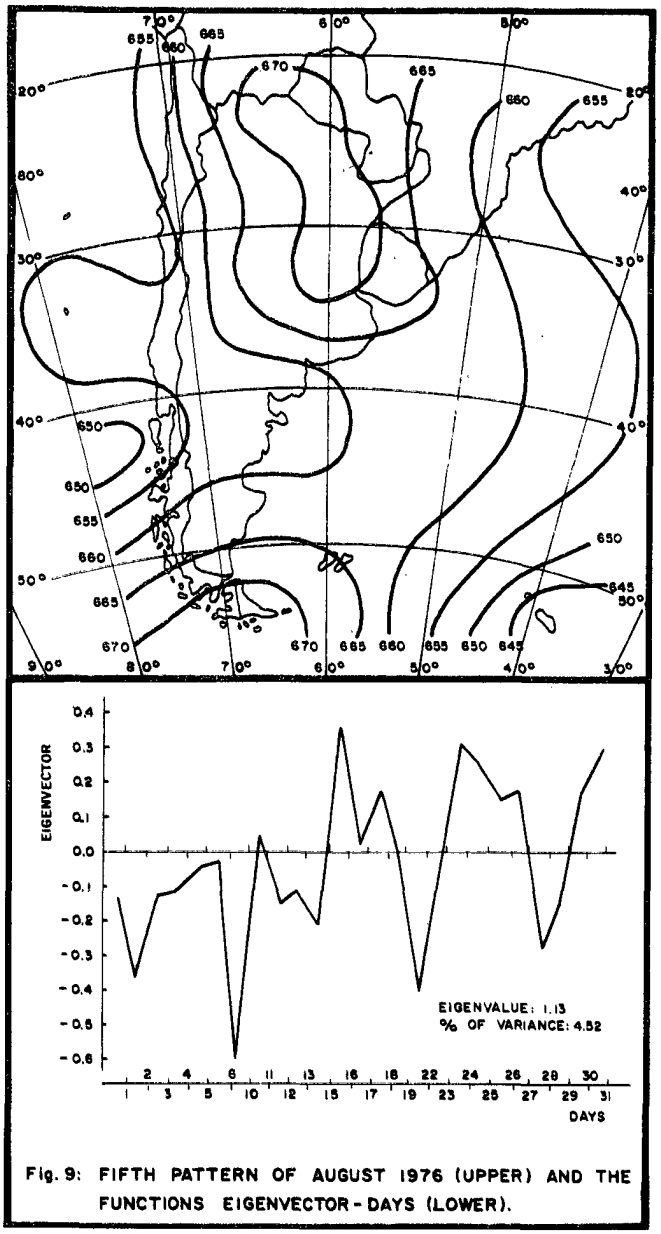


Fig. 9: FIFTH PATTERN OF AUGUST 1976 (UPPER) AND THE FUNCTIONS EIGENVECTOR-DAYS (LOWER).

and east coast of Argentina; the Pacific anticyclone slightly sketched north of parallel 30S giving low pressure in Juan Fernandez Island.

Its characteristics are a strong meridional gradient west of the Atlantic anticyclone, giving a low index of circulation; unlike this the results of the first eigenvectors gives a sample with strong zonal gradient and therefore a high index of circulation.

The fourth sample is represented by the days given in the minimums of the graphic eigenvector-days and the more representantive is the day 23.

Its inverse is in this case the maximums, the most representative being the day 13. The corresponding basic configuration is: an anticyclone entering south of Patagonia; a low pressure in the Atlantic that affects south of Brazil.

The fifth and last significant pattern (Fig. 9) represents a low pressure center in the north-east, a low one south of Patagonia, a relative low one affecting Juan Fernandez Island. It also represents a high pressure entering between parallels 40S and 50S, a high pressure center around the Georgias Island. The Atlantic anticyclone displaced to the east, the Pacific anticyclone located north of parallel 30S. According to the function eigenvector-days, it could be inferred that the dates 16, 24 and 31 would be the best ones assimilable to this sample.

Also the days 10, 22, and 28 would assimilate its inverse. Nevertheless the pressure fields of the dates mentioned are only partially identified with the sample. Maybe this is due to the scarse variance(4.52%) that represents this last type, and to the fact that the method used does not obtain, in this case, a singular matrix. This could be improved according to other metodological

varieties as the orthogonal or oblique rotation of the principal components.

5. CONCLUSIONS

We infer that the results of calculating the eigenvectors, considering in one case days as variables, and in another case stations as variables, are in both cases completely different.

With the S-Mode, it is impossible to recover the characteristics of the fields contained in the input sample. On the other hand, with the T-Mode the samples of the plasmode can be acceptably separated and also the days belonging to the types in the initial sample can be identified from the function eigenvector-days.

It could be concluded from an empirical point of view, that at least in this case, it is unnecessary to exclude days or stations in future analysis, as the errors when interpolating are negligible. There are exceptions to this, these are when it is impossible to define the field due to insufficient information.

From the analysis of the eigenvectors it is demonstrated that the grids density in the area studied can be significantly minor, not altering the results. This is valid at least when the extractions are random.

In the sample analysed almost 50% of the variance of the synoptic fields is independent of the Chilean information.

The function eigenvector-days is practically independent of the sample size. Therefore each day and consequently the month, keeps its charcteristics in any sample size. This implies that the sample and the properties given by the eigen vectors and the multipliers are unvariable in relative terms. If this occurs and a climatic sample is analysed, the metodology would allow to obtain mean and modal characteristics and the anomalies with respect to them.

The patterns given by the multipliers are quite assimilable to the sample fields, as it can be observed in the case of August 1976. Also, it is possible to recover the days corresponding to each sample taking into account the values of the function eigenvector-days. Nevertheless, if the day presents an unfrequent field, it achieves characteristics of more than one pattern among the most probables.

According to Richman's works (1981) it would be possible to obtain a singular type matrix when calculating the eigenvectors and retaining the significant ones in order to rotate them obliquely, by Harris-Kaiser class II. Furthermore with Dr.Richman's personal collaboration, who worked with the results of the plasmode for the T-Mode and S-Mode in the cases of the input correlation matrix and the input covariance matrix, we arrive to the same conclusion.

This will be taken into account in future works for real samples.

AKNOWLEDGMENTS

The authors wish to thank Dr.Richman for his collaboration and discussions on this work.

REFERENCES

Barry R.G. and Perry A.H.,1973: Synoptic Climatology.London,Mathuen Co Ltd,555pp.

Cattel R.B. and Sullivan W.,1962:The Scientific Nature of Factors:A Demostration by Cups of Coffee. Behavioral Sci.,7,184- 193

Green P.E.,1978:Analyzing Multivariate Data.Illinois,The Dryden Press,519pp.

Kaise,1959:The Application of Electronic Computers of Factor Analysis.Symposium on the Application of Computer to Psychological Problems. American Psychologial Association

Mc Cutchan and Schroeder,1973:Classification of Meteorological Patterns in Southern California by Discriminant Analysis.Journal of Applied Meteorology,2,571-577

Lorenz E.N.,1956: Empirical orthogonal functions and statistical weather prediction.Scientific Report N 1,Statistical Forecasting Project, Massachusetts Institute of Technology, Cambridge,49pp.

Lund, 1963:Map-pattern classification methods. J.Appl.Meteorology,2,56-65

Richman,1981:Obliquely rotated principal components:An Improved Meteorological Map Typing technique.J.Appl.Meteorology,20,1145-1159

Williams and Terjung,1981:Eigenvectors Filtering of three-dimensional Pressure Field Data. J.Apll.Meteorology,20,281-287

A THREE-DIMENSIONAL APPROACH TO THE STUDY OF DIURNAL PRECIPITATION PATTERNS IN THE AMERICAN PLAINS

Robert C. Balling, Jr.

Department of Geography
Climatology Program
University of Nebraska-Lincoln
Lincoln, NE 68588-0135, U.S.A.

1. INTRODUCTION

Theoretical and applied studies on the summertime nocturnal rainfall phenomenon of the North American Great Plains continue to appear frequently in the literature. Means (1944) used only one station in his landmark study of the intra-annual structure of diurnal precipitation patterns in the midwestern United States. Other investigators (e.g., Rasmusson, 1971; Wallace, 1975) included a spatial component by using more than one station, but they analyzed only the diurnal variations of a given month or season. A three-dimensional analytical procedure must be considered to effectively disentangle the diurnal, intra-annual, and spatial variance structures found in a precipitation climatology. The results from such a three-dimensional procedure may be useful to (1) various practical problems, including precipitation forecasting; (2) the evaluation of existing theories regarding nocturnal rainfall; and (3) the inductive generation of new explanatory hypotheses.

This study presents a method for operationalizing a three-dimensional procedure in the analysis of the anomalous nocturnal rainfall pattern that characterizes the North American Great Plains in the summer months. A factor analytical method is introduced to the analysis of a supermatrix containing the diurnal, intra-annual, and spatial variance patterns in hourly precipitation data. The application of a three-dimensional factor analysis to supermatrices of precipitation amounts, frequencies, durations, and intensities allows potentially important features of the Great Plains nocturnal rainfall anomaly to be clearly revealed. These same features may be masked when more commonly used methods, such as harmonic analysis (Rasmusson, 1971; Wallace, 1975), are applied to the problem. The three-dimensional method used in this study may prove to have wide applications in the analysis of many other types of meteorological and climatological variance structures.

2. THREE-DIMENSIONAL ANALYSIS

Three-dimensional factor analysis (Horst, 1965; Rummel, 1970) simultaneously disentangles diurnal, intra-annual, and spatial components in the Great Plains precipitation regime. A data cube (Fig. 1) with each cell containing a precipitation index for each of h hours, m months, and s stations is constructed to illustrate the concept of three-dimensional factor analysis. A data cube of this type may be sliced into two-dimensional matrices containing precipitation data by hour for each month for a given station. If all slices are placed in juxtaposition, a supermatrix of h rows and m x s columns is produced.

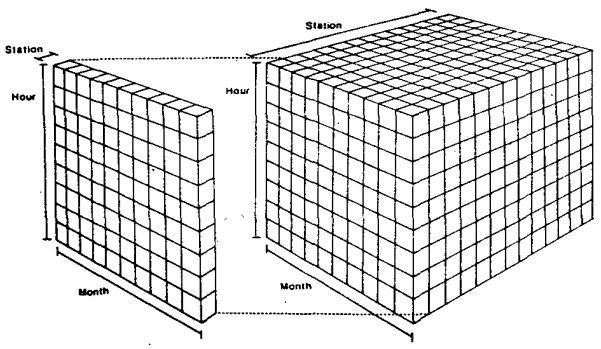


Fig. 1. Hypothetical data cube.

Eigenvectors calculated from such a supermatrix simultaneously reveal three-dimensional components of the original data cube. Factor scores for a particular eigenvector represent standardized indices of a distinctive diurnal variance structure found in the cube. Each factor loading defines a correlation between the diurnal variance pattern for a given month and station and the variance pattern depicted by a particular eigenvector. A map of the loadings displays the spatial structure of an eigenvector when the month is held constant. If only one station is examined, a plot of the factor loadings reveals the annual cycle in the importance of the variance pattern of a given eigenvector. A properly employed three-dimensional factor analysis is a powerful research tool capable of revealing salient features

in the curious precipitation patterns of the Great Plains region.

3. DATA DESCRIPTION

Eleven stations located in and around Kincer's (1916) region of Great Plains nocturnal rainfall were selected for analysis (Fig. 2). Hourly precipitation data from 1948 to 1975 formed the original data base for this study. With more than 100,000 precipitation observations ≥ 0.25 mm in the hourly data, each station passed the significance criterion suggested by Wallace (1975).

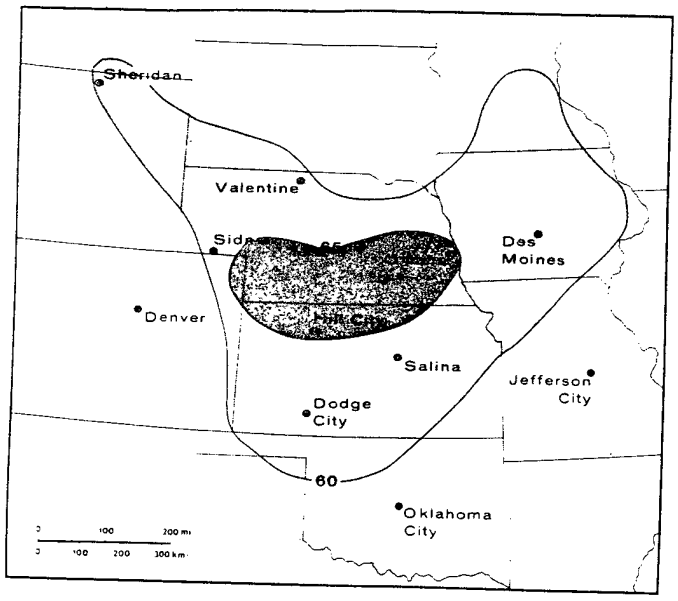


Fig. 2. Kincer's (1916) region of nocturnal rainfall and the eleven-station network for this study. Isolines represent the percentage of warm-season rainfall occurring between 2000 and 0800 LST.

Four separate supermatrices containing precipitation amounts, frequencies, intensities, and durations were constructed from the raw hourly data. Each supermatrix included 24 rows, one for each hour, and 132 columns containing the 12 monthly arrays for the 11 stations. The cells in the matrix of amounts contained the total amount of precipitation recorded for each hourly interval by month for each station. The supermatrix of frequencies similarly contained the total number of precipitation occurrences by hour, month, and station. The mean hourly precipitation intensity (mm/hr) was recorded in a third 24 x 132 matrix. The cells of the duration supermatrix contained mean values of the length of each continuous rain event associated with every hourly measurement. Because all recorded rainfall is not evenly distributed through the entire hour of data collection, the measure of intensity tends to be conservative while the duration values tend to be slightly inflated. Given the long period of record examined in this study, the "signal" in the diurnal variations for these two precipitation indices is adequately captured.

4. RESULTS

A principal components analysis with an orthogonal rotation was applied to each of the four supermatrices. The principal mode of variation in each matrix is clearly related to the summertime nocturnal rainfall pattern, thus suggesting that the mechanisms associated with nighttime rainfall produce the most distinctive diurnal precipitation pattern found in the Great Plains. The primary variance structure in each of the four supermatrices displays a temporal pattern that appears to be associated with the development and maintenance of mesoscale convective complexes (Maddox, 1980).

The first eigenvector extracted from the supermatrix of precipitation amounts explains 20.0% of the total variance in the 132 arrays. The portion of variance directly explained by this first factor (i.e., associated with positive factor loadings) is 13.1%. The diurnal pattern in the factor scores (Fig. 3) shows maximum amounts falling between 1930 and 2330 LST, with minimum values around noon. This diurnal pattern appears to be

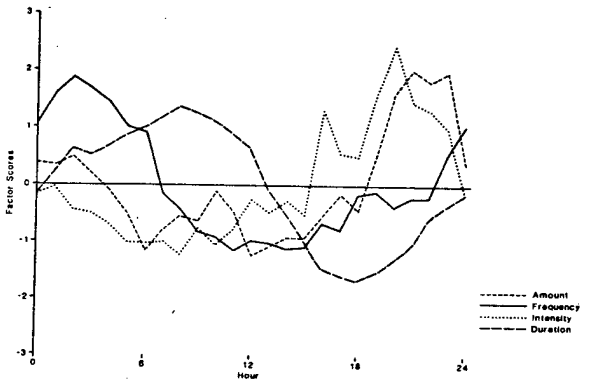


Fig. 3. Factor scores for the principal factor of precipitation amount, frequency, intensity, and duration.

best developed in the central core of the study area (Fig. 4). The five stations in this core all have more than 50% of their warm-season (April through September) rainfall between 2000 and 0800 LST, and at least 20% of their diurnal variance structure in the warm season months is directly explained by the first eigenvector. The highest mean loadings for these five core stations are from May to July, with particularly low values from September through April (Fig. 5). Additional eigenvectors are not associated with patterns that are temporally and spatially homogeneous.

The first eigenvector from the supermatrix of hourly precipitation frequencies accounts for 29.1% of the total variance with 22.3% directly explained. This principal mode of variation clearly suggests a nocturnal maximum in precipitation events in the study area. The diurnal pattern depicts relatively high frequencies occurring from 2345 to 0500 LST and low frequencies in morning and early afternoon (Fig. 3). This finding suggests that the period of maximum precipitation frequency occurs approximately three hours later

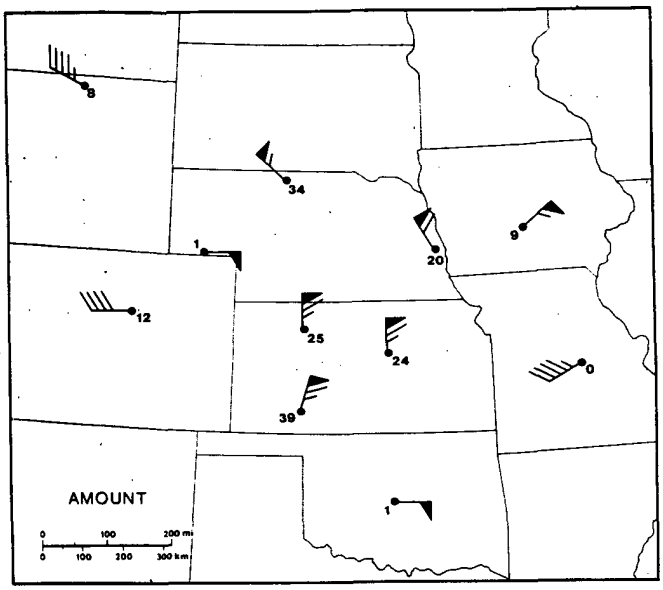


Fig. 4. Spatial components in the "nocturnal" factor in precipitation amount data. The timing of maximum amount is depicted by the orientation of the arrows, with north indicating midnight, east indicating 0600 LST, etc. The percentage of rain amounts at night (2000 to 0800) is indicated by the barbs on the tails of the arrows (half barb = 5%, full barb = 10%, and triangular barb = 50%). The numbers indicate the percent of warm-season diurnal variance explained by the "nocturnal" factor.

than the period of maximum amounts.

The nocturnal maximum in precipitation frequency appears to be best developed in the central section of the study area (Fig. 6). Six stations receive at least 55% of their warm season precipitation events from 2000 to 0800 LST, and at each of these six stations, a minimum of 33% of the warm season diurnal variance is directly explained by the first eigenvector. This nocturnal pattern appears to be best developed at the

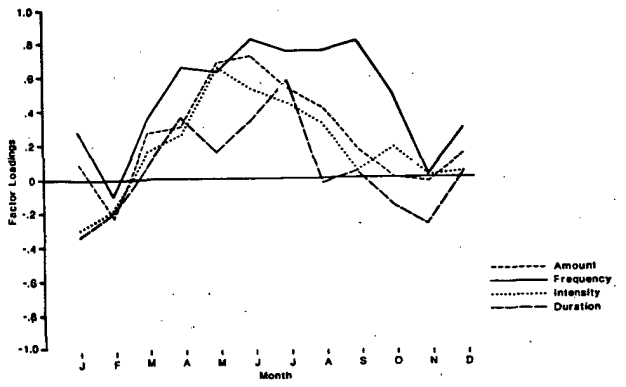


Fig. 5. Mean monthly factor loadings for the selected five station core (Dodge City, Hill City, Salina, Omaha, and Valentine).

three Kansas stations where explained variances and percentages of nighttime frequencies are especially high. Moving northeast of Dodge City, Kansas, the time of maximum frequency appears to consistently occur later in the evening. The spatial and temporal patterns associated with the first eigenvector of precipitation frequency in the core area are consistently important throughout all six warm season months (Fig. 5).

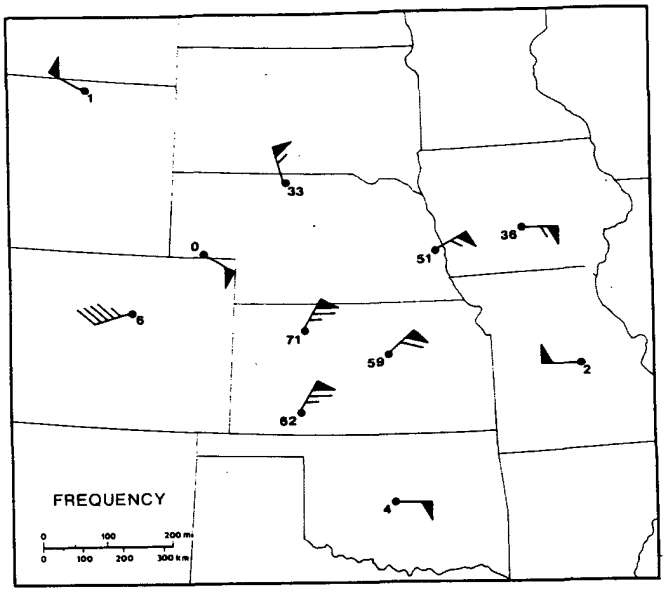


Fig. 6. Spatial components in the "nocturnal" factor in precipitation frequency data. Plotting conventions are identical to those of Fig. 4.

The diurnal pattern of the first factor from the intensity matrix shows particularly intense storms occurring from 1830 to 2300 LST, with a very high intensity value at 2000 LST (Fig. 3). The variance structure depicted by the first factor is important only in the early portion of the warm season (Fig. 5), and no coherent spatial pattern appears to be associated with this principal factor.

The first eigenvector of precipitation duration displays a diurnal pattern of longer events from 0600 to 1030 LST and shorter duration events during the late afternoon and early evening period (Fig. 3). The pattern described by the first eigenvector is important during the warm season prior to August, with July representing the month of maximum explained variance in the nocturnal core region (Fig. 5). The spatial variance pattern of this factor reveals high explained variance levels in most of the nocturnal rainfall area, but Denver, Colorado and Sheridan, Wyoming have the highest values.

The three-dimensional factor analysis produced the following results that must be accounted for by the plausible theories of nocturnal rainfall in the American Great Plains:

(1) The first component of precipitation

amounts clearly displays the propensity for nighttime rainfall in the central United States. Relatively sharp gradients in the importance of such a diurnal pattern appear to extend westward, southward, and southeastward from a core region.

(2) The patterns for precipitation frequencies suggest greater spatial and temporal coherence compared to precipitation amounts. The importance of the first factor remains quite high throughout the warm season, suggesting that studies based upon frequency measures may overestimate the relative importance of the nocturnal rainfall phenomenon. Similar to the spatial pattern associated with the first factor of amounts, steep spatial gradients are present for precipitation frequencies. There appears to be a strong tendency in the warm season for the time of maximum precipitation frequency to occur later as one moves northeast from Dodge City.

(3) Coherent temporal and spatial patterns in hourly precipitation data are exclusive features of the warm season, with a strong tendency for the nocturnal rainfall pattern to be best developed from April to June.

5. THEORETICAL INTERPRETATIONS

The many theories for explaining nighttime rainfall on the Great Plains may be conveniently categorized into four major groupings. These clusters of theories can be reevaluated in light of the preliminary results of the three-dimensional factor analysis.

(1) Some investigators (Hewson, 1937; Hales, 1977) proposed nighttime radiational and/or evaporative cloud-top cooling as a mechanism for steepening the environmental lapse rate in the cloud deck, thus initiating increased nocturnal convection. The time-space components from the analysis in this study reveal steep spatial gradients in the diurnal patterns that are not easily reconciled by the cloud-top cooling theories. The distinctive inter-monthly patterns in the primary eigenvectors are also difficult to explain with this group of theories. Although cloud-top cooling may contribute to increases in nighttime storms, the results of this study suggest that cloud-top cooling is not the primary control of the nocturnal rainfall phenomenon.

(2) Other investigators (Goldie, 1931; Dexter, 1944; Andre, 1949) suggested that rainbelts associated with frontal boundaries expand and contract on some diurnal cycle. Dexter (1944) believed that potentially unstable layers produced near the ground during the afternoon would be lifted slowly along frontal boundaries, thus destablizing the cloud layer at night. Some difficult challenges arise for such theories regardless of the mechanism proposed to force the diurnal cycle in the frontal activity. The steep spatial gradients and the pronounced inter-monthly patterns revealed in this study do not support this category of theories. Hoecker (1965) concluded that frontal passages are not normally associated with the primary diurnal variance patterns found in any of the four precipitation indices.

(3) Bleeker and Andre (1951) developed a dynamical theory that eloquently accounts for both the daytime minimum and the nighttime maximum in precipitation over the Great Plains region in the summer months. They suggested that at night drainage into the Plains caused by downslope movement of radiationally cooled air along the eastern slopes of the Rockies creates low-level convergence. The positive vertical velocities associated with this convergence significantly enhance the probability of precipitation formation in the Great Plains. An opposite process during the daytime hours presumably creates low-level divergence that diminishes rainfall amounts. The fact that no important eigenvectors associated with warm season mid-afternoon rainfall are identified in this study supports Bleeker and Andre's thesis. However, this theory cannot account for the orientation and steepness of the spatial gradients in the diurnal variance patterns. This is especially true in the case of precipitation frequency where a relatively strong southwest to northeast orientation appears to exist.

(4) The fewest problems are encountered in interpreting the findings of this study in terms of the low-level jet explanation. The low-level jet is a narrow band of relatively strong southerly winds that flow at night along an arc from Texas through the Plains states (Bonner, 1968). Thermodynamically, strong low-level advection of warm, moist air promotes instability and precipitation potential. Downwind of the jet speed maximum, low-level convergence promotes positive vertical velocities that enhance nighttime precipitation probabilities (Pitchford and London, 1962). The jet's normal position adequately accounts for the steep spatial gradients found in the frequency and amount data, and the southwest to northeast orientation of the jet accounts for a similar orientation in many of the findings. The increased lag in the time of maximum frequency found from Dodge City, Kansas to Des Moines, Iowa may be explained by the downwind migration of the jet maximum throughout the evening. A more dense spatial network of collection stations is needed for more detailed physical and dynamical interpretations of the results from the statistical analyses.

6. CONCLUSIONS

A three-dimensional factor analytical procedure is introduced to the study of the nocturnal precipitation maximum observed in the Great Plains of the United States. This statistical technique produces a set of results that simultaneously reveal underlying spatial, diurnal, and inter-monthly patterns in hourly precipitation data that may be masked by the use of popular two-dimensional approaches. Three-dimensional factor analysis is shown to be a powerful analytical tool that may be valuably applied in many areas of statistical climatological research.

7. REFERENCES

Andre, M. J., 1949: Distribution of the instability weather types with respect to summer cold fronts in the United States. Bull. Am. Met. Soc., 30, 228-230.

Bleeker, W. and Andre, M. J., 1951: On the diurnal variation of precipitation, particularly over central U.S.A., and its relation to large-scale orographic circulation systems. Quart. J. Roy. Met. Soc., 77, 260-271.

Bonner, W. D., 1968: Climatology of the low-level jet. Mon. Wea. Rev., 96, 833-850.

Dexter, R. V., 1944: The diurnal variation of warm-frontal precipitation and thunderstorms. Quart. J. Roy. Met. Soc., 70, 129-137.

Goldie, A. H. R., 1931: Characteristics of rainfall distribution in homogeneous air currents and at surfaces of discontinuity. London Meteor. Office, Geophys. Mem., 6, No. 53.

Hewson, E. W., 1937: The application of wet-bulb potential temperature to air mass analysis. III. Rainfall in depressions. Quart. J. Roy. Met. Soc., 63, 323-335.

Hales, J. E., Jr., 1977: On the relationship of convective cooling to nocturnal thunderstorms at Phoenix. Mon. Wea. Rev., 105, 1609-1613.

Hoecker, W. J., Jr., 1965: Comparative physical behavior of southerly boundary-layer wind jets. Mon. Wea. Rev., 93, 133-144.

Horst, P., 1965: Factor Analysis of Data Matrices, New York, Holt, Rinehart, and Winston.

Kincer, J. B., 1916: Daytime and nighttime precipitation and their economic significance. Mon. Wea. Rev., 44, 628-633.

Maddox, R. A., 1980: Mesoscale convective complexes. Bull. Am. Met. Soc., 61, 1374-1387.

Means, L. L., 1944: The nocturnal maximum occurrence of thunderstorms in the midwestern states. University of Chicago Masters Thesis.

Pitchford, K. L. and London, J., 1962: The low-level jet as related to nocturnal thunderstorms over midwest United States. J. Appl. Met., 1, 43-47.

Rasmusson, E. M., 1971: Diurnal variation of summertime thunderstorm activity over the U. S., Technical Note 71-4, Washington, U. S. Air Force Environmental Technical Applications Center.

Rummel, R. J., 1970: Applied Factor Analysis. Evanston, Northwestern University Press.

Wallace, J. M., 1975: Diurnal variations in precipitation and thunderstorm frequency over the conterminous United States. Mon. Wea. Rev., 103, 406-419.

PROBABILISTIC MODELING OF THE PRECIPITATION PROCESS

Ignacio Rodríguez-Iturbe

Graduate Program in Hydrology and Water Resources
Universidad Simón Bolívar
Caracas, Venezuela

1. INTRODUCTION

Studies related to the modeling of the precipitation process may be divided between those which focus on the characteristics of the process at a fixed point in space and those requiring an integrated description of temporal and spatial characteristics of rainfall. One important aspect in both cases is that rainfall in time or in space-time will not be treated on an event basis. This aspect is crucial because important dependencies in the rainfall process, e.g., its covariance structure, cannot be comprehensible described if one views rainfall on an event basis. Surprisingly there has been relatively minor emphasis in the analytical study of the covariance structure inplicit in the models used for rainfall representation at a point in time or in time and space. The reason probably lies in the mathematical difficulties encountered in the analysis. Nevertheless there is an urgent need for this kind of research which is of basic importance for rainfall forecasts, design of hydrometeorological networks and the analysis of river floods through any kind of input-output type of models.

We will first study the probabilistic models most commonly used for the representation of rainfall at a point in space, after this we will proceed to a general description of the space-time problem.

2. PUNCTUAL RAINFALL MODELS

2.1 Compound Poisson Model for Total Rainfall Depth

Storm arrivals are considered as a stream of point events occurring accordingly to a Poisson process of intensity $\lambda(t)$. Each point event carries with it an additional random variable called its mark which represents the total depth of precipitation associated with the storm (see Fig.1). The sequence of rainfall depths U(n) is a sequence of independent, identically distributed random variables with common density function $f_u(u)$. The counting process N(t), $t \geq o$, associated with the number of storms up to time t and the mark sequence U(n) are assumed to be mutually independent.

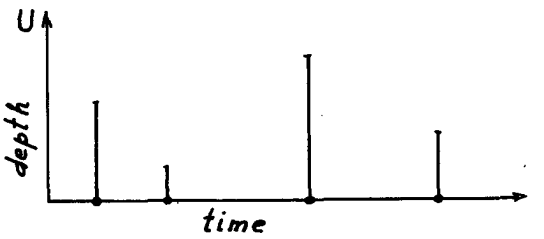


Figure 1. Compound Poisson model for total rainfall depth

The hypothesis specified above define the accumulated rainfall process as a compound Poisson process

$$N_u(t)=U(1)+U(2)+....+U(n)=\sum_{n=o}^{N(t)} U(n) \quad ,t>o \qquad (1)$$

where $N_u(t)$ is the total depth of rainfall accumulated up to time t. The moments of $N_u(t)$ are given by the well known formulas

$$E[N_u(t)]= E[U] \int_o^t \lambda(\tau)d\tau \qquad (2)$$

$$Var[N_u(t)]= E[U^2] \int_o^t \lambda(\tau)d\tau \qquad (3)$$

$$Cov[N_u(t_1), N_u(t_2)] = E[U^2] \int_o^{t_1} \lambda(\tau)d\tau \qquad (4)$$

$$t_1 < t_2$$

The instantaneous rainfall process is defined by $dN_u(t)=N_u(t+dt)-N_u(t)$, which represents an infinitesimal forward increment of the process $N_u(t)$. The moments of the instantaneous rainfall are

$$E[dN_u(t)] = E[U]\lambda(t)dt \quad (5)$$

$$Var[dN_u(t)] = E[U^2]\lambda(t)dt \quad (6)$$

$$Cov[dN_u(t_1),dN_u(t_2)] = E[U^2]\delta(t_2-t_1)\lambda(t_1).\ dt_1\ dt_2 \quad (7)$$

where $\delta(\cdot)$ represents the Dirac delta function.

The data usually available corresponds to accumulated rainfall over non overlapping periods of length T (e.g., hourly or daily rainfall). From this data is then necessary to estimate the parameters λ, $E[U]$ and $E[U^2]$, which characterize the rainfall process. For this one works with hydrologically homogeneous periods throughout the year, say we divide each year into monthly periods. In each of these periods, the mean, variance, skewness and correlation function are estimated from the data; thus we will have, per example, the mean daily (or hourly) rainfall depth for the month of April, the variance of daily rainfall for that month, its skewness and its correlation function.

Accordingly to this model, the theoretical mean and variance of the daily (or hourly) rainfall is given by (2) and (3) with $\lambda(\tau)=\lambda$ and $t=T$:

$$E[N_u(t)] = E[U].\ \lambda.T \quad (8)$$

$$Var[N_u(t)] = E[U^2].\ \lambda.T \quad (9)$$

where T=1 for the case of daily or hourly rainfall. Since we are dealing with rainfall accumulated over non overlapping intervals, the theoretical autocorrelation is zero for all lags different from zero. If this were not the case in the estimated correlation function, then there is evidence the present model is not an appropriate representation of the process under analysis. Making (8) and (9) equal to the estimated mean and variance we then have two equations to estimate the parameters $E[U]$, $E[U^2]$ and λ. It is frequently assumed that the random variable U is exponentially distributed, in this case $E[U^2] = 2E^2[U]$ and (8) and (9) are sufficient for estimation purposes. If no assumption is made regarding the distribution of the sequence of rainfall depths, U, we need one more equation. Taking U to be gamma distributed with probability density function

$$f_u(u)=\frac{\alpha^\nu u^{\alpha-1}}{\Gamma(\nu)} e^{-\alpha u}, \quad u>o \quad (10)$$

with $E[U]=\nu/\alpha$ and $Var[U]=\nu/\alpha^2$, then the skewness coefficient of the accumulated rainfall process, $N_u(t)$, is equal to (Rodríguez-Iturbe et al.,1983):

$$\gamma_1=\frac{\nu+2}{[\lambda t\nu(\nu+1)]^{1/2}} \quad (11)$$

Equation (11) is the additional relationship we need to estimate the parameters $E[U]$, $E[U^2]$ and λ.

2.2 Rectangular Pulses Rainfall Model

Rainfall is taken as a sequence of rectangular pulses of random height (rainfall intensity) and random length (rainfall duration), positioned in time accordingly to a Poisson process with rate $\lambda(t)$. Overlapping of the pulses is permitted (Figure 2). Two general cases may be distinguished: one in which the parameters of the process are taken as constant throughout time and the analysis should be performed in homogeneous time periods (e.g., monthly basis); the other in which the parameters are considered function of time and the analysis is carried out in a continuous manner. Rodríguez-Iturbe et al, (1983) have studied in detail both cases considering in the second case a general description of $\lambda(t)$ by means of a Fourier representation

$$\lambda(t)=\sum_{k=o}^{N} \lambda_{1k}\cos k\omega t + \lambda_{2k}\sin k\omega t \quad (12)$$

This paper will only discuss the first case in which the parameters of the rainfall intensity process are taken as constant throughout time.

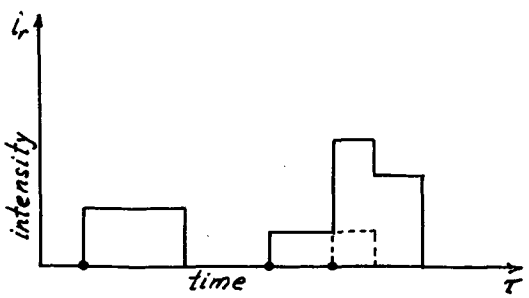


Figure 2. Rectangular pulses rainfall model

There is enough statistical evidence to allow the consideration of the rainfall intensity, i_r, and the rainfall duration, t_r, as uncorrelated random variables (e.g., Eagleson, 1970). Moreover, although the analysis can be carried out for any distribution, we will make the usual assumption that i_r and t_r are exponentially distributed random variables with mean values $\bar{i}_r$ and $\bar{t}_r$ respectively.

Calling Y (t) the rainfall intensity process it can be shown the moments of Y(t) are given by

$$E[Y(t)] = \lambda E[i_r] E[t_r] \quad (13)$$

$$Var[Y(t)] = \frac{\lambda E[i_r^2]}{\eta} \quad (14)$$

$$Corr[Y(t_1), Y(t_2)] = e^{-\eta|\tau|} \quad (15)$$

where $\eta = 1/\bar{t}_r$.

Notice that the rate of storm occurrences, λ, and the intensity of the storms, i_r, do not play any role in the correlation structure of the process which only depends on the expected value of the storm duration, $\bar{t}_r$. The coefficient of variation is in this model

$$C.V.[Y(t)] = \left(\frac{2}{\lambda \bar{t}_r} \right)^{1/2} \quad (16)$$

Equation (16) shows that humid climates with large λ and $\bar{t}_r$ will yield a much lower coefficient of variation than arid or semi-arid climates. The same is true for those months with a larger λ in comparison with the other months in the same geographic location. The skewness coefficient of the rainfall intensity process is for this model

$$\gamma_1 = \frac{3}{(2\lambda \bar{t}_r)^{1/2}} \quad (17)$$

which shows that when λ tends to be very large the normal approximation may be an acceptable one for Y(t).

Rainfall is commonly measured in an accumulative manner throughout fixed and non overlapping time intervals of lenght T. The data gathering process transform the rainfall intensity process Y(t) to a rainfall integrated process

$$Y_I(t) = \int_{t-T}^{t} Y(s)ds$$

Rodríguez-Iturbe et al., (1983) gives the moments of the integrated process $Y_I(t)$ as:

$$E[Y_I(t)] = \lambda \bar{i}_r \bar{t}_r T \quad (18)$$

$$Var[Y_I(t)] = \frac{2\lambda E[i_r^2]}{\eta^3}(\eta T - 1 + e^{-\eta T}) \quad (19)$$

$$Cov[Y_I(t), Y_I(t+\tau)] = \frac{\lambda E[i_r^2]}{\eta^3} \cdot (e^{\eta T}-1)(1-e^{-\eta T})e^{-\eta\tau}, \quad \tau > T \quad (20)$$

Notice equation (20) is valid for $\tau > T$, thus it gives the covariance for lags equal to $\tau = hT$ with $h \geq 1$. Making T=1 in (18), (19) and (20) one obtains the mean, variance and covariance of the daily or hourly rainfall process. Equalizing these expressions to the estimated mean, variance and lag-one autocorrelation coefficient yields an estimation of λ, $E[i_r]$ and $E[t_r]$ since because of the exponential assumption, $E[i_r^2] = 2E^2[i_r]$.

For T=1 (19) can be written as

$$Var[Y_I(t)] = Var[Y(t)] \cdot \frac{2}{\eta^2}(\eta - 1 + e^{-\eta}) \quad (21)$$

which gives the variance of the integrated process as the variance of the rainfall intensity process multiplied by a reduction factor depending on $\eta = 1/\bar{t}_r$. Similarly for T=1 the lag-one autocorrelation coefficient of the integrated rainfall process is

$$\rho_I(1) = \frac{1}{2} \frac{(1-e^{-\eta})^2}{(\eta - 1 + e^{-\eta})} \quad (22)$$

The correlation function of $Y_I(t)$ depends only on $\bar{t}_r$ and can not take negative values. Sometimes $\hat{\rho}_I(1)$ is negative most probably due to sampling variations and uncertainties inherent to the estimation process; nevertheless it is conceivable, depending on the integration interval T, that $\rho_I(1)$ may take a value of zero. This implies in the present model that $\bar{t}_r \to 0$ and $\bar{i}_r \to \infty$, with the product $\bar{i}_r \cdot \bar{t}_r$ remaining constant and equal to the expected value of storm rainfall depth. In this case the rectangular pulses model collapses into the compound Poisson model for total rainfall depth studied in section 2.1 (Rodríguez-Iturbe et al., 1983).

2.3 Neyman-Scott Rainfall Representation.

Each storm is conceptualized as a cluster of instantaneous bursts, the "cluster centers" are point events positioned in time accordingly to a Poisson process with rate $\lambda(t)$. The number of bursts in each cluster are independent and identically distributed random variables, ν.

Each burst carries with it an amount of rain X and is placed randomly from its corresponding cluster center accordingly to a probability density function f(t) (Figure 3). We further assume that the sequence of X_i constitute an i.i.d. random sequence and X_i is independent of the other random variables in the model.

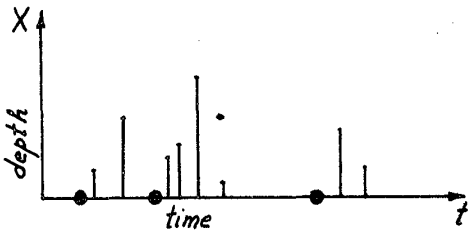


Figure 3. Neyman-Scott rainfall representation

Notice that there are two possible interpretations of the Neyman-Scott rainfall representation. One was just described before in which the bursts constitute the internal structure of the process, the second one is to interpret each X_i as the total depth of a storm in which case we are just modelling the storm occurrences as a cluster type of process. This interpretation is in the same spirit of the compound Poisson model for total rainfall depth in section 2.1. Now we have the great advantage of allowing serial correlation between rainfall accumulated in non-overlapping periods.

Let N(t) denote the total number of rainfall bursts in time (0,t). Let a process N_t be defined as follows:

N_t = 1 if there is a burst in dt around t

N_t = 0 otherwise (23)

Mathematically N(t) and $\{N_s, 0<s<t\}$ are related by

$$N(t) = \int_0^t N_s \, ds \qquad (24)$$

The covariance structure of N(t) and N_s has been studied in a pioneering paper by Waymire and Gupta (1981). The important fact here is that the rainfall intensity process Y(t) is now represented by

$Y(t) = X(t) \cdot N_t$ if a burst occurs in dt around t

$Y(t) = 0$ otherwise (25)

The structure of Y(t) is studied by Rodríguez-Iturbe et al. (1983), the moments for the case $\lambda(t) = \lambda$ are:

$$E[Y(t)] = \lambda . E[X] . E[\nu] \qquad (26)$$

$$Cov[Y(t_1),Y(t_2)] = E[X^2] . \lambda . E[\nu] . \delta(|\tau|) + E^2[X] . E[\nu^2-\nu] . \lambda . \int_0^\infty f(u) . f(|\tau|+u) du \quad -\infty<\tau<\infty \qquad (27)$$

We shall assume following Kavvas and Delleur (1981) that f(t) has an exponential form

$$f(t) = \beta e^{-\beta t} \quad , \; t>0, \beta>0 \qquad (28)$$

In this case the covariance of the rainfall intensity process reduces to

$$Cov[Y(t_1),Y(t_2)] = \theta_1 \delta(|\tau|) + \theta_2 e^{-\beta|\tau|}, \quad -\infty<\tau<\infty \qquad (29)$$

where

$$\theta_1 = E[X^2] . \lambda . E[\nu] \qquad (30)$$

$$\theta_2 = E^2[X] . \lambda . E[\nu^2-\nu] . \frac{\beta}{2} \qquad (31)$$

Because of the same motivations explained in sections 2.1 and 2.2 it is important to know the mathematical characteristics of the integrated rainfall process under a Neyman-Scott representation. Assuming again that rainfall is integrated throughout equal and non overlapping periods of length T, the moments of the integrated process $Y_I(t)$ are given by Rodríguez-Iturbe et al. (1983) as:

$$E[Y_I(t)] = \lambda . E[X] . E[\nu] . T \qquad (32)$$

$$Var[Y_I(t)] = \theta_1 T + \frac{2\theta_2}{\beta^2}[T - \frac{(1-e^{-\beta T})}{\beta}] \qquad (33)$$

$$Cov[Y_I(t_1),Y_I(t_2)] = \frac{\theta_2}{\beta^2} e^{-\beta\tau}(e^{\beta T}-1) . (1-e^{-\beta T}), \tau>T \qquad (34)$$

2.4 Time Scale Considerations

The time scale at which rainfall needs to be represented is of crucial importance in the selection of the appropriate probabilistic representation. At one extreme, there are time scales, say of the order of one week, for which in some regions of the world there does not seem to exist a correlation structure in the rainfall amounts. Nevertheless there are many climates for which the previous statement is obviously false;

the correlation structure also becomes stronger with the shortening of the time scale and this makes necessary the consideration of models like the rectangular pulses or Neyman-Scott representation which allow such a structure among non overlapping periods.

The Neyman-Scott type of model is more flexible than the rectangular pulses scheme in what concerns the correlation structure of the process. In the later one, the autocorrelation only depends on the mean of the storms duration, $E[t_r]$ but in the Neyman-Scott all the parameters involved in the process play a role in the correlation structure. This richness of the structure has its counterpart in the number of parameters to estimate which is larger than in the rectangular pulses case.

The importance of the time scale on the validity of any particular representation is also reflected in the fact that one may implement a model like those previously described at a certain time scale estimating the parameters from the integrated rainfall process at that particular scale; these parameters may nevertheless be quite different from those obtained when doing the analysis at a different time scale. Rodríguez-Iturbe et al., (1983) present some examples of this situation, more experience with different kinds of precipitation data is necessary in this particular aspect.

3. SPATIAL RAINFALL MODELS

As mentioned in section 1 of this paper when studying the precipitation process one deals with rainfall in time or in space-time. Nevertheless as an introduction to the space-time study of the rainfall field it is of interest in its own right a brief look to the spatial structure of a rainfall event and the probabilistic modelling of such an structure. After doing this we will proceed to make a qualitative description of the space-time modelling and its areas of major interest.

3.1 Static Spatial Modeling of Total Depth of a Rainfall Event

Figure 4 shows the typical structure found in precipitation fields. Rainfall is occurring at a large mesoscale level, LMSA, where we distinguish a background rainfall intensity accompanied by small mesoscale areas, SMSA, where precipitation is more intense due to the clustering of rain cells.

Suposse cluster centers defining the occurrence of SMSA are distributed in space in accordance with a two dimensional Poisson process with rate parameters λ :

X(S) = number of cluster centers in region S

S = region of the LMSA

$$P[X(S)=k] = e^{-\lambda A(S)} \frac{[\lambda A(S)]^k}{k!} \quad \text{for } k=0,1,2,\ldots \quad (35)$$

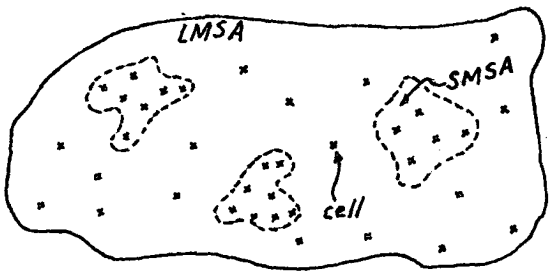


Figure 4. Rainfall process in space

Each cluster center is associated, independently of other cluster centers, to a random number of rain cells, ν . Thus, the spatial pattern of rain cells, is described by a two dimensional Neyman-Scott process in the plane. For simplicity in the analysis we may consider a static model in which there is no consideration of the movement of rain cells. Also, since we are not describing the temporal structure of the storm, there is no provision for the fact that cells are born continuously throughout the storm life time, have a period of growth and then die continuously during the storm. For effects of total rainfall depth at any point we need to consider the rain contributions to this point from the different cells in the storm. Each cell is taken to spread its rain in space accordingly to a function which may have random parameters on its structure.

The total rainfall depth $\zeta(r,\theta)$ at a point of coordinates (r,θ) is then:

$$\zeta(r,\theta) = \int_S g(r,\theta)\, dN(r,\theta) \quad (36)$$

where g (.) is the rain spread function of a single cell. A reasonable structure for g (.) is

$$g(r,\theta) = g(r) = M_o e^{-br} \quad (37)$$

where Mo is a random variable standing for the total rainfall depth produced by

a cell at the cell's center and b is a decay parameter assumed to be the same for all cells (although it could also have been considered to be a random variable). The term $dN(r,\theta)$ is defined through the relationship

$$N(R) = \int_R dN(r,\theta) \qquad (38)$$

where N(R) is the number of cells in a region R. Taking Mo to be independent and identically distributed random variables uncorrelated with the spatial Neyman-Scott process N(R) one obtains the mean value of the rainfall depth at a point as:

$$E[\zeta(r,\theta)] = \int_S E[g(r,\theta)].E[dN(r,\theta)] \qquad (39)$$

Usually the LMSA is large enough so that S may be taken as the whole two dimensional space. In this case

$$E[\zeta(r,\theta)] = \frac{2\pi E[Mo]\ \lambda\ E[\nu]}{b^2} \qquad (40)$$

Rodríguez-Iturbe et al., (1983) also study the variance and covariance structure of $\zeta(r,\theta)$.

3.2 Space-Time Modeling of the Rainfall Intensity Process

Waymire et al., (1983) have recently developed a model for the representation of the space-time covariance structure of the rainfall intensity field. We will limit ourselves in this paper to a qualitative description of the main features of this model. The first important thing now is that rainfall is treated as a dynamic process both in what concerns the arrival of storms and the growth and development of each event.

Storm arrivals are considered as a stream of point events occurring accordingly to a Poisson process. The spatial pattern of storms is described by a two-dimensional Neyman-Scott process along the same ideas presented in section 3.1. Both the LMSA and the cells are treated dynamically with the incorporation of their velocities into the model. The process of cells birth, growth and death is explicitly incorporated into the model with the appropriate probabilistic considerations. There is now a spread function in time and space for the dissemination of the rainfall intensity field of a cell, this spread function may incorporate random parameters.

The space-time covariance structure obtained in this model has been extremely useful for the study of the conditions under which Taylor's hypothesis of fluid turbulence can be extended to the space-time precipitation field. Work now in process is addressed to the study of the relative importance of the different parameters for an adequate representation of the rainfall intensity field. A parallel line of research deals with the design of rainfall networks, topic for which it is of crucial importance the covariance structure of the rainfall intensity field (Rodríguez-Iturbe and Mejía, 1974).

4. FINAL COMENTS

The probabilistic modeling of the precipitation process besides of its own fundamental relevance in the study of rainfall phenomena has also a crucial importance in the related area of river flow analysis. In effect, the streamflow process is the output from the watershed when the rainfall process is inputed into the system. The covariance structure of the streamflows is a key factor in the study of their extreme values, crossing properties and sampling requirements. Rodríguez-Iturbe et al., (1983*) derive the covariance structure of the streamflow process under different representations of the rainfall process. Thus, the models described in section 2 of this paper where used as inputs into a kernel representing the main physical features of a basin in order to derive the corresponding properties of the instantaneous and integrated streamflow processes.

Future work will deal with the use of space-time covariance structures of the rainfall field as input into kernel representations of a watershed.

5. REFERENCES

Eagleson, P.S., 1970: Dynamic Hydrology. McGraw-Hill, 462p.

Kavvas, M.L., and J.W. Delleur, 1981: A stochastic cluster model of daily rainfall sequences. Water Resources Research, 17 (4), 1151-1160.

Rodríguez-Iturbe, I., and J.M. Mejía,1974: The design of rainfall networks in time and space. Water Resources Research, 10 (4), 713-728.

Rodríguez-Iturbe, I., V.K. Gupta, and E. C. Waymire, 1983: A theoretical framework for rainfall-runoff analysis at the basin scale, Part I: the rainfall process. To be published, Water Resources Research.

Rodríguez-Iturbe, I., V.K. Gupta, and E. C. Waymire, 1983*: A theoretical framework for rainfall-runoff analysis at

the basin scale, Part II: The run-off process. To be published, Water Resources Research.

Waymire, E.C., and V.K. Gupta, 1981: The mathematical structure of rainfall representations. 3: Some applications of the point processes theory to rainfall processes. Water Resources Research, 17 (5), 1287-1294.

Waymire, E.C., V.K. Gupta, and I. Rodríguez-Iturbe, 1983: A representation of a rainfall intensity field at the mesoscale and the structure of its space-time covariance. To be published, Journal of Atmospheric Sciences.

6. ACKNOWLEDGEMENTS

The research presented in this paper is part of a joint investigation with Prof. Vijay Gupta (University of Mississippi) and Prof. Ed Waymire (Oregon State University). I also thank the Venezuelan Council for Scientific and Technological Research (CONICIT) for financial support. The author is also a member of the Instituto Internacional de Estudios Avanzados (I.D.E.A.).

STOCHASTIC MODELING OF RAINFALL IN SPACE AND TIME

Vijay K. Gupta

Department of Civil Engineering
The University of Mississippi
University, Mississippi 38677

1. INTRODUCTION

Precipitation is the end result of complex interactions between dynamic, thermodynamic and cloud microphysical processes at different scales in space and time. This makes the modeling and prediction of precipitation a very complicated and difficult problem. Borrowing the widely used terminology of Orlanski (1975) for meteorological scale classification, precipitation at the meso-β scale (20-200 Km) is always observed to exhibit significant variability in space and time. Although the physics of precipitation at this scale is not well understood, it is a scale of great significance from the viewpoint of hydrologic modeling (Gupta et al., 1983), for parameterization of general circulation in tropical climates (Arakawa and Schubert, 1974; Cho, 1978) and for meteorological forecasting and prediction of precipitation in tropical and extratropical climatic regions (Leary et al., 1983). Therefore, one of the major goals of any research in predicting precipitation is to be able to model the observed variability at the meso-β scale in a mathematically general and a physically meaningful manner. This paper summarizes results of a recent study (Waymire et al., 1983) and identifies some future research directions toward solving this very important problem.

Past attempts on the statistical modeling of rainfall have been largely confined to modeling its temporal structure at fixed spatial locations, primarily because of an easy availability of rainfall data from raingages. However, these models have been developed on a case by case basis in a mathematically restrictive and physically ad hoc manner; see Waymire and Gupta (1981, Part I) and Katz (1983) for some recent reviews. Since the spatial scale of a raingage in general is not suitable for specifying precipitation inputs into hydrologic models, e.g., in regional water balance studies, in rainfall-runoff modeling at the basin scale, etc., (see, e.g., Gupta et al., 1983) it becomes necessary to obtain quantitative information about rainfall at appropriate spatial and temporal scales. The available models on the temporal evolution of rainfall at a fixed spatial point do not extend to a space-time framework in a mathematically general and physically meaningful manner. Although the importance of scales in hydrologic modeling and understanding has not been given much attention in the past, it is our conviction that it is destined to play a fundamental role in future development of hydrologic theories (see, e.g., Rodriguez-Iturbe and Gupta, 1983; Fiering, 1982). Since precipitation is one of the most basic components of the hydrologic cycle, it is rather obvious that a very serious effort is necessary to make significant progress in understanding and modeling precipitation at appropriate scales of practical interest.

Within the last decade one of the major advances in diagnostic meteorology has been the recognition that precipitation at the meso-β scale organizes itself into distinct patterns. This organization exhibits the fact that a precipitation region at one scale contains one or several regions of more intense precipitation embedded within it. In an unstable environment, the precipitation intensities at the meso-β scale can be highly variable (10-100 mm h^{-1}); see, e.g., Houze and Hobbs (1982); Houze (1981); Hobbs and Locatelli (1978); Harrold and Austin (1974); etc., for excellent descriptions of this organization. This preferred organization is observed in diverse storm types in extratropical as well as in tropical climates. In this sense, a certain universality exists in the structure of space-time precipitation at the meso-β scale and this provides an avenue along which stochastic models can be developed to model this variability. The general stochastic assumptions that go into this type of modeling of precipitation at the meso-β scale are briefly described in Section 2.

In order to obtain specific information about the rainfall intensity field from its general representation described in Section 2, it becomes necessary to introduce specific assumptions and approximations into this representation. One such scheme has been carried out recently by Waymire et al. (1983) to derive an analytical expression for the covariance function of the rainfall intensity field. This expression correctly explains the empirical findings of Zawadski (1973) on a rainfall intensity covariance function which include the physical conditions under which the covariance satisfies Taylor's hypothesis on fluid turbulence (see, e.g., Hinze, 1975, pp. 46-47) and when it does not. These results are summarized in Section 3.

The model assumptions in the work of Waymire et al. (1983) only capture the observed variability

in rainfall at the meso-β scale in a qualitative manner. No considerations involving the physics of precipitation are included in this model. Studies of meso-β scale precipitation in the tropical Atlantic have shown that their existence is modulated by large-scale air motions in the form of macro-β scale (2000-10,000 Km) waves in the middle troposphere. This dynamic control on precipitation at a larger scale is accompanied by statistical fluctuations in its occurrence at the meso-β scale (see, e.g., Houze and Betts, 1981). Such large scale dynamic controls are also observed to be present in the extratropical climates (Houze and Hobbs, 1982). Therefore it is an important avenue to investigate the manner in which the dynamical aspects of air motions at large scales constrain and modulate the statistical structure of precipitation at the meso-β scale. In this sense the modeling of space-time rainfall is a problem of coupling large scale dynamics with its stochastic description at smaller scales. Section 4 gives a brief discussion to illustrate this point. It is precisely this aspect of the statistical modeling of meso-β scale precipitation fields that needs major emphasis in future researches. Some new ideas along these lines are discussed by Leary et al. (1983) in this symposium.

2. A MATHEMATICAL REPRESENTATION OF THE RAINFALL INTENSITY FIELD AT THE MESO-β SCALE

The general mathematical representation given in this section is an extension and a generalization of the previous work by Gupta and Waymire (1979) to allow for temporal occurrences of multiple rainbands over a fixed but arbitrary geographic region. Physically, these occurrences can be due to a single synoptic scale disturbance passing over that region or due to distinct synoptic fronts. The observed embedding of regions of more intense precipitation within a rainband is mathematically incorporated in this description in the form of spatial clustering of occurrences of rain cells within a rainband.

In this article the terms 'stochastic point process' and 'point random field' designate the random occurrences of points in a one and a higher dimensional space respectively; the word 'random' does not refer to uniform and independent occurrences in the present context. Geometrically, each occurrence is represented by a point in this space. For example, the random cell occurrences in space-time are denoted by the vector $(\tau,\underset{\rightarrow}{y})$ where $\tau \in (-\infty,\infty)$ denotes the time of birth of a cell and $\underset{\rightarrow}{y} \in R^2$ its spatial location at the time of birth within the two dimensional plane. If B is a subregion of this three dimensional space, then by a point random field $X(\cdot)$ we also mean that $X(B)$ is the random number of cells in the subregion B. Mathematically, the two modes of characterizing a stochastic point process or a point random field either by the coordinates of occurrences of random events or by the random number of such occurrences in one or more subregions are equivalent; see Waymire and Gupta (1981, Part II) for a further discussion.

For specificity we will motivate the following mathematical representation in terms of extratropical cyclonic storms. The mathematical ideas will, however, apply to many other storm types as well, for example, cloud clusters in the tropics.

Fix a spatial origin at some geographic location within the range of extratropical cyclonic storms. The location of rainbands are identified relative to this spatial origin and the times at which rainbands arrive are measured with respect to a fixed but arbitrary time origin. A general mathematical structure of space-time rainfall will now be summarized.

(I) Rainband arrivals

The arrival times of rainbands, say, $\ldots,-s_1,s_0,s_1,\ldots$, represent the random occurence of a stochastic point process $M(\cdot)$ in the time interval $(-\infty,\infty)$.

For the purpose of representing a precipitation cloud field inside a rainband we identify a precipitating cloud, i.e., a rain cell, by its random time of birth τ and its location $\underset{\rightarrow}{y}$ at the time of birth inside a moving rainband. Physically, the point $(\tau,\underset{\rightarrow}{y})$ in space-time can be identified with the point of maximum rainfall intensity which we assume to be unique. It is called a 'birth-center point'.

(II) Occurrences of cluster potentials within a band

The occurrence of cloud clusters or cluster potentials in a rainband which arrives at time s_i, are regarded as point occurrences of a point random field $L_{s_i}(\cdot)$ in the two-dimensional plane region $B(s_i)$ denoting the extent of the rainband. Each occurrence of a cluster potential is denoted by a random vector $\underset{\rightarrow}{x} \in B(s_i)$.

(III) Occurrences of raincells within a cluster potential

The configuration of cell birth-center points inside a cluster potential region located at $\underset{\rightarrow}{x}$ within a rainband arriving at time s_i are the random point occurrences of a point random field $N_{s_i,\underset{\rightarrow}{x}}(\cdot)$ in space-time.

(IV) Occurrences of raincells within a rainband

The birth center points of raincells due to a rainband arrival at a random time s_i are the random occurrences of a stochastic point random field $N_{s_i}(\cdot)$ in three-dimensional space-time. Given the rainband arrivals at times $\ldots,-s_1,s_0,s_1,\ldots$, the random fields $N_{-s_1}(\cdot),N_{s_0}(\cdot),\ldots$, etc., are assumed to be statistically independent of one another. The point random field $N_{s_i}(\cdot)$ is itself a cluster field in space obtained by combining (II) and (III) as,

$$N_{s_i}(\cdot) = \sum_{\underset{\sim}{x}_j \in L_{s_i}} N_{s_i,\underset{\sim}{x}_j}(\cdot)$$

$$= \int N_{s_i,\underset{\sim}{x}}(\cdot)\, L_{s_i}(d\underset{\sim}{x}) \quad . \qquad (2.1)$$

(V) Totality of rain cells from all rainbands

The birth-center points of raincells due to multiple rainbands are the occurrences of a point random field $X(\cdot)$. This represents a cluster point random field in space-time given by combining (I) and (IV) as,

$$X(\cdot) = \sum_{s_i \in M} N_{s_i}(\cdot) = \int_{-\infty}^{\infty} N_s(\cdot)\, M(ds) \qquad (2.2)$$

Substituting (2.1) into (2.2),

$$X(\cdot) = \sum_{s_i \in M} \sum_{\underset{\sim}{x}_j \in L_{s_i}} N_{s_i,\underset{\sim}{x}_j}(\cdot)$$

$$= \int \left\{ \int N_{s,\underset{\sim}{x}}(\cdot)\, L_s(d\underset{\sim}{x}) \right\} M(ds) \ , \qquad (2.3)$$

According to (2.3), a precipitating cloud field is represented as a two stage clustering field. The first stage clustering is in space and the second stage clustering is in time.

(VI) Representation of rainfall intensity at ground level

In order to obtain a representation of the rainfall intensity at ground level, we introduce the cell rainfall intensity $g(a,r)$ denoting the rainfall from a cell of age a at distance r from the cell center. Let $\underset{\sim}{U}_b = (U_{b_1}, U_{b_2})$ denote the rainband velocity relative to the fixed coordinates on the ground and let $\underset{\sim}{U}_c = (U_{c_1}, U_{c_2})$ denote the common cell velocity relative to the band. Define,

$$\underset{\sim}{U} = \underset{\sim}{U}_b + \underset{\sim}{U}_c \qquad (2.4)$$

Now for a cell with birth-center point $(\tau, \underset{\sim}{y})$, the rainfall contribution at a ground point at $\underset{\sim}{z}$ at time t is $g(t-\tau, \| \underset{\sim}{z} - (\underset{\sim}{y} + \underset{\sim}{U}(t-\tau)) \|)$ where $\| \cdot \|$ denotes the length of a vector. By adding rainfall contributions from all cells whose occurrences are governed by the point random field $X(\cdot)$ constructed in (2.3), we arrive at the representation,

$$\xi(t,\underset{\sim}{z}) = \sum_{(\tau_i, \underset{\sim}{y}_i) \in X} g(t-\tau_i, \| \underset{\sim}{z} - (\underset{\sim}{y}_i - \underset{\sim}{U}(t-\tau_i)) \|)$$

$$= \int g(t-\tau, \| \underset{\sim}{z} - (\underset{\sim}{y} - \underset{\sim}{U}(t-\tau)) \|) \cdot X(d\tau, d\underset{\sim}{y}) \qquad (2.5)$$

where integration in (2.5) is over $(-\infty, \infty)$ in time and over the entire two-dimensional space. It provides the mathematical representation that we sought to obtain.

3. SPACE-TIME COVARIANCE OF RAINFALL INTENSITY AND ITS CONNECTION WITH THE TAYLOR'S HYPOTHESIS ON FLUID TURBULENCE

Some experimental results on the covariance function of a precipitation field are reported in the work of Zawadski (1973). However, because it is difficult to separate out the components of a storm system, Zawadski (1973) did not attempt a theoretical analysis at the scale of turbulence. Instead he assumed that the instantaneous rainfall intensity field $\xi(t,\underset{\sim}{z})$ at the ground level is a random field and then undertook an experimental determination of the covariance structure of this field in space-time. He observed that during the development and peak periods of storms the cells' speed $(|\underset{\sim}{U}_c|)$ relative to the storm speed $(|\underset{\sim}{U}_b|)$ is essentially zero, i.e., the cells and the storm move essentially with the same speed (Zawadski, 1973, p. 465, Table 2). However, during the dissipation period of the storm, the cells continued to move with essentially the same speed whereas the storm speed became nearly zero. The data in Table 2 implies that $|\underset{\sim}{U}_b| = |\underset{\sim}{U}|$ for the first 110 minutes because $|\underset{\sim}{U}_c| = 0$. Zawadski observed that during the first 40 minutes, the covariance function of the rainfall intensity field,

$$\mathrm{Cov}(\xi(t_1,\underset{\sim}{z}_1), \xi(t_2,\underset{\sim}{z}_2)) = \rho(|t_1 - t_2|, \|\underset{\sim}{z}_1 - \underset{\sim}{z}_2\|) \ , \qquad (3.1)$$

satisfies Taylor's hypothesis on fluid turbulence (Hinze, 1975, p. 46-47) which may be expressed as,

$$\rho(|t_1 - t_2|, 0) = \rho(0, \| \underset{\sim}{U}(t_1 - t_2) \|) \qquad (3.2)$$

In words, (3.2) says that the autocovariance at some lag $(t_1 - t_2)$ at a fixed spatial point is the same as the spatial covariance at some fixed time, say zero, between two spatial points separated by $\underset{\sim}{U}(t_1 - t_2)$. For time lags larger than 40 minutes, the difference between $\rho(|t_1 - t_2|, 0)$ and $\rho(0, \| \underset{\sim}{U}(t_1 - t_2) \|)$ was observed to become significant (Zawadski, 1973, Fig. 15).

In order to obtain an expression for the rainfall covariance function from the representation (2.5), Waymire et al. (1983) introduced specific assumptions on different stochastic components which make up the representation. This was necessary because (2.5) is too general in that it only captures the broad features of embedding of cells within cluster potentials in space and that of the latter in time within rainbands. The assumptions were intended to incorporate the observed features of the components appearing in (2.5) only in a qualitative manner. Then an analytical expression was derived for the space-time covariance of the

ground level rainfall field. The main features of this covariance function are (i) when $\underset{\rightarrow}{U}_b = \underset{\rightarrow}{U}$, then for $\tau \sim T$, where T denotes the mean life time of cells,

$$\rho(\tau,0) \simeq \rho(\underset{\rightarrow}{0}, \| \underset{\rightarrow}{U}\tau \|), \quad \underset{\rightarrow}{U} = \underset{\rightarrow}{U}_b, \quad \tau < T \tag{3.3}$$

and (ii) for $\tau > T$,

$$\rho(\tau,0) < \rho(0, \| \underset{\rightarrow}{U}\tau \|), \quad \underset{\rightarrow}{U} = \underset{\rightarrow}{U}_b, \quad \tau > T \tag{3.4}$$

Equation (3.3) is the same as (3.2) for $\tau < T$. Therefore the analytical expression derived by Waymire et al. (1983) indeed satisfies Taylor's hypothesis under similar physical conditions as were observed by Zawadski (1973), because T, the mean life span of rain cells, is approximately 40 minutes to an hour. For $\tau > T$, the inequality given by (3.4) also agrees with the observations of Zawadski; in this range the rainfall covariance does not satisfy Taylor's hypothesis. Therefore qualitative features of the covariance function obtained by Waymire et al. (1983) theoretically explain the observed features of space-time rainfall covariance shown in Fig. 15 in Zawadski (1973). In this sense, one gets some confidence in the general ideas and approximations underlying the stochastic representations of rainfall fields in (2.5) as representing physical reality to a good first approximation. However, the main drawback of this theoretical work lies in the fact that the mathematical assumptions do not explicitly include precipitation dynamics.

4. ON THE INCLUSION OF DYNAMICS IN THE STATISTICAL MODELING OF SPACE-TIME PRECIPITATION

The dynamics of precipitation at the meso-β scale is not understood at present. However it is clearly recognized in the current literature that the air motions at the synoptic scale or meso-α scale (200-2000 Km) and macro-β scale (2000-10,000 Km) exert a strong modulating influence on the observed precipitation at the meso-β scale in extratropical storms (Houze and Hobbs, 1982) as well as in tropical storms (Houze and Betts, 1981). In fact, in the parameterization schemes of general circulation in the tropics some preliminary attempts have already been made to include the effect of a large number of cloud clusters in a statistical manner at the meso-β scale (Arakawa and Schubert, 1974; Ogura and Cho, 1973). Physically speaking, this interaction between dynamics at a large scale and statistical fluctuations at a smaller scale in the tropics comes from the fact that the cumulus clouds transform latent heat into sensible heat through condensation of water vapor. The released latent heat drives the vertical circulation of cumulus clouds and this in turn produces a mixture of moisture and sensible heat contents in the vertical direction (Cho, 1978). Some of the condensed water is released in the form of precipitation. Therefore, development of cloud fields and generation of precipitation at the meso-β scale in the tropics depends critically on the large scale moisture and heat balances. Consequently, the collective or mean behavior of the shallow non-precipitating cloud fields as well as those of the precipitating cloud fields is constrained by the large scale heat and moisture balance equations (Cho, 1978).

Cho's investigation dealt with determining the dynamical constraints imposed by large scale heat and moisture balances on a statistical population of shallow non-precipitating cumulus cloud clusters in the tropics. He proposed that the spatial distribution of the cloud groups can be described by a Poisson field and made some statistical tests to verify this assumption with the radar imagery of observed cloud groups reported by Lopez (1976). The agreement was found to be quite good. He then derived an expression for the probability density of the areas or sizes of cloud groups and showed it to follow an exponential density under certain dynamical conditions. More significantly, an expression for the mean cloud size was obtained in terms of the net moisture flux at the cloud base. This physical constraint was deduced from the moisture and heat balance equations at the synoptic scale. Once again good verification of this density was obtained with size distribution of cloud radar echoes reported by Lopez (1976). Though Cho's (1978) investigation did not include the size distribution of precipitating cloud groups, an extension of his results to precipitating clouds can be incorporated into the purely statistical representation of the precipitation field described in Section 2. We will briefly illustrate this important point below because it clearly shows the manner in which dynamics at the synoptic scale can constrain the statistical descriptions of precipitation at the meso-β scale.

Let us assume for illustrative purposes that the point occurrences of $N_{s,\underset{\rightarrow}{x}}(\cdot)$ in (III) in Section 2 are a random number ν of points, independently and identically distributed within a cluster potential with center at $\underset{\rightarrow}{x}$ within a rainband arriving at time s. Let,

$$E[N_{s,\underset{\rightarrow}{x}}(d\tau,d\underset{\rightarrow}{y})] = p_{N_{s,\underset{\rightarrow}{x}}}(\tau,\underset{\rightarrow}{y})\, d\tau d\underset{\rightarrow}{y}$$

$$= \langle\nu\rangle\, f_{s,\underset{\rightarrow}{x}}(\tau,\underset{\rightarrow}{y})\, d\tau d\underset{\rightarrow}{y} \quad , \tag{4.1}$$

where $E[\cdot]$ denotes mathematical expectation and $\langle\nu\rangle$ is the mean of ν. $p_{N_{s,\underset{\rightarrow}{x}}}(\tau,\underset{\rightarrow}{y})$ is also called the first order product density of the random field $N_{s,\underset{\rightarrow}{x}}(\cdot)$. Let us further assume that (also see Waymire et al., 1983),

$$f_{s,\underset{\rightarrow}{x}}(\tau,\underset{\rightarrow}{y}) = f_1(\tau-s) f_2(\| \underset{\rightarrow}{y}-\underset{\rightarrow}{x} \|) \quad , \tag{4.2}$$

where $f_2(\cdot)$ may be interpreted as the probability of occurrence of a raincall at a distance $\| \underset{\rightarrow}{y}-\underset{\rightarrow}{x} \|$ from the cluster potential center at $\underset{\rightarrow}{x}$. It is physically natural to suppose that $f_2(\cdot)$ is proportional to the spatial area which a cluster potential occupies. Following Waymire at al. (1983), is we assume that,

$$f_2(\| \underline{y} \|) = \frac{1}{2\pi\sigma^2} e^{-(y_1^2 + y_2^2)/2\sigma^2} , \quad (4.3)$$

where $\underline{y} = (y_1, y_2)$, then the transformation,

$$R^2 = Y_1^2 + Y_2^2 , \quad (4.4)$$

along with (4.3), gives (Parzen, 1960, p. 321),

$$f_2(r^2) = \frac{1}{2\sigma^2} e^{-r^2/2\sigma^2} . \quad (4.5)$$

The exponential density in (4.5) then corresponds to the size distribution of the cluster potential regions. This is precisely the expression derived by Cho (1978) for the size distribution of cloud groups. More importantly, he obtained an expres- expression for σ^2 in (4.5) in terms of the mean moisture flux in each cloud group as imposed by the moisture and heat budget equations at the synoptic scale.

Cho's derivation depends critically on the assumption that the cloud groups consist of shallow non-precipitating clouds, whereas (4.5) refers to the size distribution of precipitating cloud groups. Therefore, on physical grounds, the expression for σ^2 derived by Cho will not apply to (4.5). Nevertheless, the above discussion clearly illustrates the manner in which the synoptic scale dynamics affects statistical description of a precipitation field at the meso-β scale. Incidentally, the assumptions (4.2) and (4.3) were also introduced by Waymire et al. (1983) for computing the expression for the precipitation covariance function discussed in Section 3.

5. FINAL REMARKS

The discussion in Section 4 only illustrates some ideas on one possible avenue along which the synoptic scale dynamics can be coupled with a statistical description of precipitation at the meso-β scale. On physical grounds it may be more fruitful to investigate the statistical structure of those physical parameters which control precipitation rates provided the mechanisms relating these parameters to the precipitation rates can be adequately identified (see, e.g., Leary et al., 1983). We think that these various avenues on coupling the synoptic scale dynamics with the statistical fluctuations in precipitation at the meso-β scale need to be explored in an imaginative manner because they lie at the heart of making basic advances in our understanding of precipitation phenomena and in modeling and prediction. Since statistical descriptions of rainfall have always been of key importance in various hydrologic studies, basic advances in this area will undoubtedly be of immense value in diverse investigations ranging from statistical modeling of hydrologic processes from physical considerations to design of data collection networks for rainfall and other hydrologic variables (Rodriguez-Iturbe, 1983).

We think that solutions of many of the problems discussed above will require a close interaction between hydrologists, meteorologists, atmospheric physicists, climatologists and mathematicians. One of the major objectives of the Precipitation Committee of the Hydrology Section, American Geophysical Union, is to promote such an interaction in the United States and Canada. An article detailing important research problems related to precipitation will be forthcoming on behalf of this committee in EOS, an AGU weekly publication. We hope that scientists from other countries will also take an initiative in promoting a close collaboration among different disciplines toward solving this very important problem.

6. ACKNOWLEDGEMENTS

Several fruitful discussions with Han Ru Cho and Ed Waymire and some input from Colleen Leary were very helpful in formulating some ideas detailed in this paper. To them goes my sincere appreciation. This research was supported by the grant CME-7907793 from the National Science Foundation.

7. REFERENCES

Arakawa, A. and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. J. Atmos Sci., 31, 674-701.

Cho, H. R., 1978: Some statistical properties of a homogeneous and stationary shallow cumulus cloud field. J. Atmos. Sci., 35(1), 125-138.

Fiering, R. (Chairman), 1982: Scientific Basis of Water Resources Management. National Academy Press, Washington, D. C.

Gupta, V. K. and E. C. Waymire, 1979: A stochastic kinematic study of subsynoptic space-time rainfall. Water Resour. Res., 15(3), 637-644.

Gupta, V. K., I. Rodriguez-Iturbe and E. C. Waymire, 1983: Towards a unified statistical theory of rainfall-runoff at the basin scale. (to appear)

Harrold, T. W. and P. M. Austin, 1974: The structure of precipitation systems - A review. J. de Rech. Atmos., 8(1-2), 41-57.

Hinze, J. O., 1975: Turbulence, 2nd Edition. McGraw-Hill, New York.

Hobbs, P. V. and J. D. Locatelli, 1978: Rainbands, precipitation cores and generating cells in a cyclonic storm. J. Atmos. Sci., 35, 230-241.

Houze, R. A., 1981: Structures of atmospheric precipitation systems: A global survey. Radio Sci., 16(5), 671-689.

Houze, R. A., Jr. and A. K. Betts, 1981: Convection in GATE. In Reviews of Geophysics and Space Physics, 19, 541-576.

Houze, R. A., Jr. and P. V. Hobbs, 1982: Organization and structures of precipitating cloud systems. Adv. in Geophys., 24, 225-315.

Katz, R. W., 1983: Probabilistic Models. In Probability Statistics and Decision Making in the Atmospheric Sciences. (eds. A. H. Murphy and R. W. Katz), West View Press, Boulder, Colorado.

Leary, C. A., D. R. Haragan and H. R. Cho, 1983: Precipitation Dynamics as a Stochastic Process. Proceed 2nd Int. Symp. on Statistical Climatology, Lisbon, Portugal, Sep., 1983 (this issue).

Lopez, R. E., 1976: Radar characteristics of the cloud populations of tropical disturbances in the northwest Atlantic. Mon. Wea. Rev., 104, 268-283.

Ogura, Y. and H. R. Cho, 1973: Diagnostic determination of cumulus cloud populations from observed large-scale variables. J. Atmos. Sci., 30, 1276-1286.

Orlanski, I., 1975: A rational subdivision for atmospheric processes. Bul. Amer. Meteor. Soc., 52, 1186-1188.

Parzen, E., 1960: Modern Probability Theory and Its Applications. J. Wiley, New York.

Rodriguez-Iturbe, I., 1983: Probabilistic Modeling of the Precipitation Process. Proceed. 2nd Int. Symp. on Statistical Climatology, Lisbon, Portugal, Sep., 1983 (this issue).

Rodriguez-Iturbe, I. and V. K. Gupta (Guest Editors), 1983: Scale Problems in Hydrology. J. Hydrol., special issue.

Waymire, E. C. and V. K. Gupta, 1981: The mathematical structure of rainfall representations. I. A review of the stochastic rainfall models. Water Resour. Res., 17(5), 1261-1272.

Waymire, E. C. and V. K. Gupta, 1981: The mathematical structure of rainfall representations. II. A review of the theory of point processes. Water Resour. Res., 17(5), 1273-1285.

Waymire, E. C., V. K. Gupta and I. Rodriguez-Iturbe, 1983: Spectral theory of a rainfall intensity field at the mesoscale. (to appear)

Zawadski, I. I., 1973: Statistical properties of precipitation patterns. J. Appl. Meteorol., 12, 459-472.

TREND EFFECTS ON LONG TIME SCALES

Ed Waymire*

Department of Mathematics
Oregon State University
Corvallis, Oregon 97331 U.S.A.

1. INTRODUCTION

The nature of climatic variation is a topic which crosses the boundaries of many disciplines. As a result the very meaning of 'climatic variation' is enriched by a diversity in scientific perspectives. In the present paper we come to this topic from the point of review of hydrology. The particular climatic considerations focused on here originate in the very practical requirements for long term reservior storage design. However the implications of these considerations extend far beyond the domain of engineering into the heart of statistical climatology.

Let $X_1, X_2, \ldots$ be a sequence of random variables. The partial sum sequence Q_t, $t = 1,2,\ldots$ is defined by

$$Q_0 = 0 \tag{1}$$

$$Q_t = X_1 + X_2 + \cdots + X_t, \quad x \geq 1. \tag{2}$$

The (location) 'adjusted partial sum' sequence $Q_{t,N}$ for the period N, $t = 0,1,\ldots,N$, is defined by

$$Q_{t,N} = Q_t - \frac{t}{N} Q_N, \quad t = 0,1,\ldots,N. \tag{3}$$

Finally the 'adjusted range' statistic is then defined by

$$R_N = \max_{0 \leq t \leq N} Q_{t,N} - \min_{0 \leq t \leq N} Q_{t,N}, \tag{4}$$

and the 'rescaled adjusted range' statistic is the ratio R_N/D_N where D_N is the sample standard deviation of $X_1, X_2, \ldots, X_N$.

Although hydrologists were prompted by reservior design problems to study the statistic R_N in the case when $X_1, X_2, \ldots, X_N$ represent annual (volumetric) streamflows over a large design span N, the lack of a wide availability of records led Harold Edwin Hurst (1951) to study the statistic R_N/D_N for other variables indicative of climatic patterns such as tree rings, mud varves, precipitation, temperatures, atmospheric pressures, etc., along with his own extensive data collection of streamflows on the Nile river. The basic discovery reported by Hurst, now known as the 'Hurst Phenomena' is that R_N/D_N behaves like N^H, for large values of N, where H is strictly larger than .5. Typical values of H are scattered around .7.

* Partially supported by a grant from the National Science Foundation, U.S.A.

2. SOME THEORETICAL RESULTS

In this section results which are known rigorously in connection with the Hurst phenomena are highlighted. The summary is by no means exhaustive as the literature on this topic is quite vast. The intent is only to indicate the various directions which have been taken in this connection.

In the statements of the following results the Hurst exponent H refers to a real number $H \geq .5$ such that $N^{-H} \cdot R_N/D_N$ converges in distribution as $N \to \infty$ to a (possibly degenerate) distribution not concentrated at 0. The first theoretical result to be motivated by the emperical work of Hurst is the following theorem of Feller (1951).

Theorem 1. If $X_1, X_2, \ldots$ is a sequence of independent and identically distributed random variables with finite second moment then $H = .5$.

Much of the theoretical work of the period subsequenct to Feller concerns relaxation of independence. However since Feller's result remains true under various forms of weak dependence (for example Markov dependence, ϕ-mixing, etc.), the Hurst effect ($H > \frac{1}{2}$) requires long-range correlations if $X_1, X_2, \ldots$ is to be kept stationary with finite second moments. The following two results are indicative of the departures from Feller's (weak dependence, stationary, finite second moment) framework. These together with a more comprehensive overview can be found in Mandelbrot and Taqqu (1979).

Theorem 2. If $X_1, X_2, \ldots$ is a stationary Gaussian sequence with correlations given by $\rho(n) = \frac{1}{2}\{|n+1|^\theta - 2|n|^\theta + |n-1|^\theta\}$, where $0 \leq \theta \leq 2$, then $H = .5\theta$.

Theorem 3. If $X_1, X_2, \ldots$ is a sequence of independent and identically distributed random variables having a stable distribution with exponent α, $0 < \alpha < 2$, then $H = .5$.

The first rigorous result leading to the Hurst effect due to non-stationarity is illustrated by the following theorem of Bhattacharya, Gupta, and Waymire (1983).

Theorem 4. If $Y_1, Y_2, \ldots$ is a sequence of independent and identically distributed random variables with finite second moment and $X_n = Y_n + c \cdot (n+m)^\beta + d$, $n \geq 1$, where $c > 0$, $m \geq 0$, $d \in (-\infty,\infty)$, and $\beta \in (-\infty,\infty)$ then

$$H = \begin{cases} .5 & \text{if } \beta = 0 \text{ or if } \beta \leq -.5 \\ \beta + 1 & \text{if } -.5 \leq H < 0 \\ 1 & \text{if } \beta > 0. \end{cases}$$

Each of the above special results reflects the nature of conditions under which one can expect various exponents H. However it should be emphasized that these are only special cases of results which are known more robustly; see the references cited above for the broader framework. Roughly speaking, the result of Theorem 1,

$H = .5$, refers to the case when $X_1, X_2, \ldots$ is a stationary weakly dependent sequence with finite second moment. Theorem 2, $H = -.5\theta > .5$ for $1 < \theta < 2$, refers to a case when $X_1, X_2, \ldots$ is stationary with finite second moment but with long range dependence. Notice that $\rho(n) \sim (\text{const.})n^{\theta-2}$ as $n \to \infty$ so that in particular $\sum_{n=0}^{\infty} \rho(n) = +\infty$ for $1 < \theta < 2$, in contrast, for example, to the case of weak dependence where the decay of correlations is exponentially fast. Theorem 3, $H = .5$, refers to a case in which $X_1, X_2, \ldots$ is stationary weakly dependent sequence with no second moment. Finally Theorem 4, $H = \beta + 1$, $-.5 < \beta < 0$, refers to a weakly dependent non-stationary case having finite second moments with the non-stationary due to a small decreasing trend.

3. IMPLICATIONS FOR STATISTICAL CLIMATOLOGY

Intrinsic to the most primitive definitions of climatic change, based on deterministic considerations of averages over blocks of time, are both the notions of statistical non-stationarity and that of statistical dependence as soon as one admits to naturally occurring random fluctuations. The role of random fluctuations in climatic variables is addressed in Curry (1962) and Jacchia (1975) at the foundational level of physical theory. It seems quite clear from the points raised by these authors that deeper statistical considerations are required to resolve even the simplest issues. For example both Curry and Jacchia suggest that such notions as cycles and periodicities observed for various climatic variables are perhaps only 'apparent' (virtual effects). Such hypothesis are ventured on the basis of the return time behavior of the simple random walks as furnished by arc-sine laws (see Feller, 1968).

It is precisely in connection with considerations such as those of Curry and Jacchia that more definitive statistical analysis is required. Along these lines the analysis of phenomena such as that observed by Hurst has quite transparent implications. After all, in view of Theorem 1 the Hurst effect would not be reflected in the statistical models examined by these authors. This is not meant to suggest that the Hurst effect should be used as the sole definitive criterion, but only that such technical results can play some part in the decisive procedures. Since climatologist's are more interested in the maximal fluctuations that are likely to occur in the long run than in short term fluctuations, the law of the iterated logarithm serves also as a more refined criterion in this connection.

Physical theories on the nature of climatic variation are currently being debated and re-evaluated at various fundamental levels. For example, the presumed links between climatic variation and estimates of oceanic cooling trends based on standard oxygen isotope analysis for deep sea microfossils are now being challenged by the recent results of Killingley (1983). While we do not pretend to be able to give a physical theory which will support conditions such as those indicated by Theorem 4, such conditions do seem to be fairly representative of current observation. Apart from this, though, there remains an important purely statistical issue. Namely, there is the issue of designing a statistical test which will discriminate between the alternative sets of conditions required by Theorems 2 and 4. Efforts in this direction are currently being pursued by Bhattacharya, et. al.

4. REFERENCES

Bhattacharya, R. N., V. K. Gupta, and Ed Waymire, 1983: The Hurst effect under trends, J. Appl. Probab., to appear.

Curry, Leslie, 1962: Climatic change as a random series, Ann. Assoc, Amer. Geograph., 52, 21-31.

Feller, W., 1951: The asymptotic distribution of the range of sums of independent random variables, Ann. Math. Stat., 22, 427-432.

Feller, W., 1968: An introduction to probability theory and its applications, Vol. I, Third Edition, Wiley, New York.

Hurst, H. E., 1951: Long term storage capacity of reserviors, Trans. Amer. Soc. Civil Engrs., 116, 770-779

Jacchia, L. G., 1975: Some thoughts about randomness, Sky and Telescope, 50, 371-374.

Killingley, J. S., 1983: Effects of diagnostic recrystallization on $^{18}O/^{16}O$ values of deep sea sediments, Nature, Vol. 301, 594-597.

Mandelbrot, B. B. and M. S. Taqqu, 1979: Robust R/S anslysis of long run serial correlation, Proc. 42nd Intnal., Stat. Inst., I.P. 19.3, Manila, 1-38.

PRECIPITATION DYNAMICS AS A STOCHASTIC PROCESS

Colleen A. Leary
Donald R. Haragan

Atmospheric Science Group
Department of Geosciences
Texas Tech University
Lubbock, Texas, USA

Han-Ru Cho

Department of Physics
University of Toronto
Toronto, Ontario, Canada

1. INTRODUCTION

Precipitation is one of the most difficult meteorological quantities to forecast, because it is the end result of a combination of dynamic, thermodynamic, and cloud microphysical processes. An important emphasis in meteorology has been the identification of meteorological phenomena, including precipitation patterns, with the interactions of physical processes operating on a variety of scales in time and space. An assortment of terminologies have been used to define these scales, but Orlanski's (1975) has been most widely adopted, and is used below.

A major discovery of the last ten years has been the recognition of the importance of organization of precipitation patterns at the meso-β scale (20-200 km). In an unstable environment, meso-γ scale (2-20 km in horizontal dimension) cumulonimbus clouds are often organized into characteristic meso-β scale precipitation patterns. The internal structure of these precipitation patterns includes a leading edge of intense convective cells of about 25 km in width and 200 km in length. This leading edge is frequently arc-shaped. Rainfall rates in this portion of the precipitation pattern are highly variable (10-100 mm h^{-1}). Behind the leading edge in the mature precipitation feature, a meso-β scale area of rainfall is characterized by rather steady, light to moderate precipitation rates (1-10 mm h^{-1}).

In a stable environment, band-shaped areas of moderate to heavy precipitation at the meso-β scale (called rainbands) are often embedded in a large area of stratiform cloud at the meso-α or macro-β (2000-10000 km) scale. These rainbands are typically 5-50 km in width and 100-200 km in length. The dynamics involved in the formation of rainbands are not yet understood. A more detailed description of the precipitation and cloud structures at the various scales can be found in the review papers by Houze (1981) and Houze and Hobbs (1982).

Since the scales of the processes which produce precipitation constitute a continuum, all of them cannot be resolved in one single model. For both theoretical and practical reasons, a model describing or predicting precipitation can have only limited temporal and spatial resolution. At the scales resolved by the model, atmospheric variables and processes must be determined either by the model itself or from meteorological observations. These processes can be treated mainly as being deterministic because of the deterministic nature of the governing laws. Atmospheric processes at the unresolvable scales have to be treated only statistically because the information needed to determine these processes is missing.

A major task in precipitation modeling is then to determine the statistical structure of the unresolved processes, and to relate the statistical structure to the processes at the resolved scales. It is a problem of statistics as well as dynamics.

Because of the strong control and feedback mechanisms between scales involved in producing precipitation, it is a reasonable assumption that a link does exist between the statistics of the smaller scale and the dynamics of the larger scale. Observations appear to support such an assumption. For example, studies of meso-β scale precipitation features in the tropical Atlantic (see the review by Houze and Betts, 1981) have shown that their existence is modulated by the larger-scale air motions in the form of macro-β or synoptic scale waves in the middle troposphere. The cloud clusters preferentially occur in and ahead of the

wave trough. This deterministic aspect of the occurrence of the meso-β scale precipitation pattern is accompanied by an uncertainty in the exact position within a large meso-α (200-2000 km) scale area where the smaller meso-β scale precipitation pattern will occur (Fig. 1.1). This uncertainty makes the difference between the occurrence and non-occurrence of rainfall at each point due to a particular precipitation pattern. Such large-scale dynamical control over the existence of much smaller-scale precipitation patterns is not true only in the tropics, but in middle latitudes as well (Houze and Hobbs, 1982). Hydrological modeling can be improved if the two deterministic aspects of precipitation patterns can somehow be taken into account. First, the modulating effect of larger-scale weather phenomena like atmospheric waves on the existence of precipitation patterns provides both useful information for establishing the frequency of occurrence of precipitation patterns for a particular area and for forecasting the occurrence of a precipitation pattern on a particular day. Second, the similarity in internal structure of meso-β scale precipitation patterns can be useful in defining with greater precision the structure of individual precipitation events in hydrological models.

This paper presents two examples of methods by which the deterministic aspects of precipitation patterns can be linked with the statistical structure of the unresolved processes. In the first example, bands of heavy rain embedded within a larger area of less intense, more uniform precipitation rates are modeled by allowing the specified large-scale temperature field (the component determined by large-scale dynamics and left out of present stochastic models of precipitation) to possess statistical fluctuations that control the placement and intensity of the rainbands. In the second example, a specified time and space dependent precipitation pattern typical of convection organized on the meso-β scale is simulated to illustrate how measurements by a network of raingages contain both a deterministic component (the presence of the pattern as a whole) and a random component (the exact location of individual cells) in a combination that greatly obscures the nature of the originally specified precipitation pattern.

2. INCORPORATING ATMOSPHERIC DYNAMICS IN STOCHASTIC PRECIPITATION MODELS

In stochastic precipitation modeling for hydrological purposes, assumptions are often made about the stochastic behavior of the precipitation field itself. But precipitation is the end product of many physical processes. A model that attempts to describe the precipitation field directly excludes the natural evolution of the precipitation field dictated by the laws of physics. On the other hand, if the physical mechanisms controlling the rate of precipitation are understood, a stochastic model of precipitation may be constructed by examining the stochastic behavior of the controlling physical parameters. Such a combination of physical laws with the laws of statistics is quite common in the study of stochastic processes. One famous example is the study of Brownian motion (Doob, 1942).

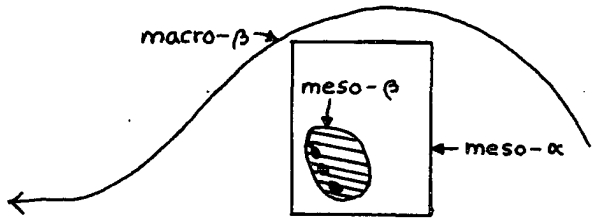


Fig. 1.1 Meso-β scale precipitation pattern (shaded) embedded in a meso-α scale region (rectangle) within a macro-β scale wave.

In this section, an example will be used to illustrate how this unification of dynamics and statistics may be achieved in the study of precipitation distribution. The example is concerned with the precipitation produced by clouds near a front. The assumption will be made that the vertical stratification of the atmosphere is stable. The cloud system we are dealing with is a widespread stratiform cloud. Due to the internal dynamics of the atmosphere near a front, there are natural variations of precipitation even in such a stratiform cloud system.

2.1 The Rates of Condensation and Precipitation in a Stratiform Cloud

Before the precipitation rate may be determined, one needs first to determine the rate of condensation of water vapor. In a stratiform cloud, the water vapor mixing ratio is maintained at the level of saturation

$$q = q_s(T,p) \tag{2.1}$$

$q_s(T,p)$ here is the saturation mixing ratio. It is a function of temperature T and pressure p. Following the motion of an air parcel, the rate of condensation is equal to the rate of decrease of water vapor:

$$C = -\frac{Dq_s}{Dt} \tag{2.2}$$

Since q_s is a function of temperature and pressure, we let

$$F = \frac{\partial q_s}{\partial T} \quad \text{and} \quad G = \frac{\partial q_s}{\partial p} \quad . \tag{2.3}$$

The rate of condensation can be written as

$$C = -F\frac{DT}{Dt} - G\frac{Dp}{Dt} \quad . \tag{2.4}$$

Here $\omega = Dp/Dt$ is the rate of change of pressure following the motion of an air parcel. It is the measure of vertical velocity when the pressure is used as the vertical coordinate. To a high degree of approximation, it is related to the regular vertical velocity w through

$$\omega = -\rho g w \tag{2.5}$$

where ρ is air density and g is the gravitational acceleration.

Including the heating effect due to condensation of water vapor, the first law of thermodynamics requires that

$$\frac{DT}{Dt} = \Gamma_d \omega + \frac{L}{C_p} C \qquad (2.6)$$

$\Gamma_d = (C_p \rho)^{-1}$ is the dry adiabatic lapse rate in p-coordinates. C_p is the specific heat at constant pressure and L is the latent heat of vaporization. This leads to an explicit expression for the rate of condensation:

$$C = -(F \Gamma_d + G) \omega/(1 + FL/C_p) \quad . \qquad (2.7)$$

It is directly proportional to the vertical velocity.

The conversion of the condensed water vapor into precipitation particles is determined by cloud microphysical processes. It is often expressed in terms of a precipitation efficiency ε. The rate of generation of precipitation can be written as

$$r = \varepsilon C \quad . \qquad (2.8)$$

The total precipitation produced by a vertical air column is obtained by vertically integrating equation (2.8):

$$P = \int_{p_T}^{p_B} \varepsilon C \, dp/g = - \int_{p_T}^{p_B} \varepsilon (F\Gamma_d + G)\omega/(1 + FL/C_p) \, dp/g \qquad (2.9)$$

where p_B and p_T are the pressures at the cloud base and the cloud top respectively.

2.2 The Vertical Velocity Field in the Neighborhood of an Atmospheric Front

To determine the distribution of precipitation near a front, one needs to determine the vertical velocity field which is controlled by dynamic processes.

It is well known that the formation of a front is caused by a deformation wind field acting on an inhomogeneous temperature field. This process is illustrated schematically in Fig. 2.1. It shows a horizontal cross-section of the atmosphere. The solid curves show the deformation velocity field with the y-axis being the axis of dilation and x-axis the axis of contraction. The dashed lines are the isotherms. They are parallel to the y-axis so that the temperature gradient is in the +x-direction. The discussions presented here will be based on the configuration shown in this figure. The rate of deformation $\alpha = \frac{1}{2}(\frac{\partial v}{\partial y} - \frac{\partial u}{\partial x})$ will be assumed to be constant in height and in time. Here u and v are the x- and y-components of the wind vector. Other variables needed in the discussion include the potential temperature θ and the equivalent potential temperature θ_e defined by

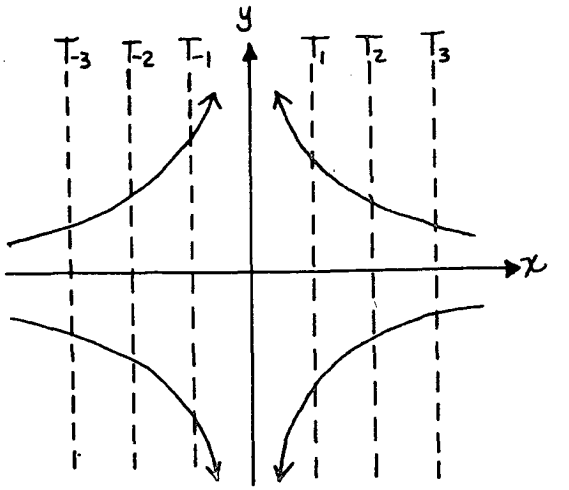


Fig. 2.1 Deformation field associated with frontogenesis. Dashed lines are isotherms.

$$\theta = T \, (p/p_o)^{-R/C_p} \qquad (2.10)$$

and

$$\theta_e = \theta \exp \, (Lq_s/C_pT) \qquad (2.11)$$

where R is the gas constant.

The dynamics of frontogenetic processes including latent heating in general is a problem which is difficult to treat analytically. However, some approximations can be made which lead to very useful simplifications. As noted by Hoskins and Bretherton (1972), one of the observed features of well developed fronts and their associated cloud systems is that θ_e is approximately constant up the frontal slope while the absolute vorticity vector ξ is usually oriented approximately along the frontal slope. There is little stability for motion along vortex lines. This effect of latent heat to reduce stability can be crudely modeled by assuming that the potential vorticity q given by

$$q = \underset{\sim}{\eta} \cdot \underset{\sim}{\nabla}_3 \theta \qquad (2.12)$$

is approximately zero. Here η is the 3-D absolute vorticity vector and ∇_3 the 3-D gradient operator. The results of the zero potential vorticity model proposed by Hoskins and Bretherton (1972) then become relevant here.

The initial state of the atmosphere that is consistent with the zero potential vorticity approximation is a potential temperature field which is independent of height:

$$\theta = \theta_o + \theta_i(x) \quad \text{at } t=0 \qquad (2.13)$$

θ_o here is the horizontal mean potential temperature and θ_i the initial deviation from the mean. The nonlinear dynamics in the frontogenetic process may be simplified by introducing the geostrophic coordinate

$$X = x - v/f \qquad (2.14)$$

where f is the Coriolis parameter. Initially,

when the horizontal temperature gradient is very weak,

$$X_o = x \quad . \tag{2.15}$$

Under the influence of the deformation field, the horizontal scale along the x-axis contracts. If the rate of deformation α is a constant, then following the motion of an air parcel

$$X = X_o\, e^{-\alpha t} \tag{2.16}$$

At time t, the potential temperature field is given by

$$\theta(x,t) = \theta_o + \theta_i(Xe^{\alpha t}) = \theta_o + \Theta(X,t) \tag{2.17}$$

The wind field for the zero potential vorticity model has been determined by Hoskins and Bretherton (1972). In obtaining the solutions the Boussinesq approximation was made and the tropopause was assumed to be rigid and maintained at height z = H. These approximations will not seriously affect our results because precipitation is produced mostly in the lower troposphere. The y-velocity which is needed in the coordinate transformation in equation (2.14) is found to be

$$v = \frac{g}{f\theta_o}\left(z - \frac{H}{2}\right)\frac{\partial\Theta}{\partial X} \quad , \tag{2.18}$$

and the vertical velocity is given by

$$w = -\frac{\alpha g}{f^2\theta_o}\,\frac{H^2/4 - \left(z - \frac{H}{2}\right)^2}{1 - (g/f^2\theta_o)\left(z - \frac{H}{2}\right)\partial^2\Theta/\partial X^2}\,\frac{\partial^2\Theta}{\partial X^2} \tag{2.19}$$

2.3 The Distribution of Precipitation near a Front

The distribution of precipitation near a front can now be examined using equation (2.9) and the vertical velocity given by equation (2.19).

We may assume, for example, that the initial broad baroclinic field is described by the arctangent function.

$$\theta_i(x) = \Delta\theta\frac{2}{\pi}\tan^{-1}\frac{x}{L_o} \tag{2.20}$$

At time t, we obtain from equation (2.17) that

$$\Theta(X,t) = \theta_o + \Delta\theta\frac{2}{\pi}\tan^{-1}\frac{X}{L} \tag{2.21}$$

where

$$L = L_o e^{-\alpha t} \tag{2.22}$$

The vertical velocity as a function of space X and t can also be determined:

$$w = +\frac{\alpha g}{f^2\theta_o}\left(\frac{\Delta\theta}{L^2}\frac{2}{\pi}\frac{2(X/L)}{(1 + (X/L)^2)^2}\right) \cdot \tag{2.23}$$

$$\frac{\left(\frac{H}{4}\right)^2 - \left(z - \frac{H}{2}\right)^2}{1 + \frac{g}{f^2\theta_o}\left(z - \frac{H}{2}\right)\frac{\Delta\theta}{L^2}\frac{2}{\pi}\frac{2(X/L)}{(1 + (X/L)^2)^2}}$$

The time and place of front formation at the ground will depend on the maximum value of $2(X/L)/(1 + (X/L)^2)^2$:

$$A = \text{max.}\ \{2(X/L)/(1 + (X/L)^2)^2\} \tag{2.24}$$

A surface temperature discontinuity will form at time $t = t_s$ when

$$\frac{g}{f^2\theta_o}\frac{H}{2}\frac{\Delta\theta}{L_o}\frac{2}{\pi}A\, e^{2\alpha t_s} = 1 \tag{2.25}$$

The position of the front is at $X = X_s$ where

$$2(X_s/L)/(1 + (X_s/L)^2)^2 = A \tag{2.26}$$

The solution, of course, is valid only for $t \leq t_s$. Equation (2.23) shows that during the processes of front formation, the vertical velocity is the strongest near the frontal zone. This is the region where the heaviest precipitation will fall. The region of heavy precipitation becomes narrower as the frontal structure becomes sharper when time increases.

If initially there are some temperature perturbations at the mesoscale embedded within the broad baroclinic field, then these initial temperature perturbations may lead to enhanced precipitation in band shaped areas parallel to the front. This can be shown easily by describing the initial mesoscale temperature perturbation θ_{mi} in terms of its Fourier spectrum:

$$\theta_{mi}(x) = \sum_j \Delta\theta_j \sin\frac{2\pi}{\lambda_{jo}}x \tag{2.27}$$

The wavelength λ_{jo}, being of the mesoscale, should be much smaller than L_o, the scale of the synoptic scale baroclinic field:

$$\lambda_{jo} << L_o$$

The total initial temperature field can then be written as

$$\theta_i = \theta_o + \Delta\theta\frac{2}{\pi}\tan^{-1}\frac{x}{L_o} + \sum_j \Delta\theta_j \sin\frac{2\pi}{\lambda_{jo}}x \qquad \text{at } t=0 \tag{2.29}$$

At time t, the temperature field becomes

$$\theta(X,t) = \theta_o + \Delta\theta\frac{2}{\pi}\tan^{-1}\frac{X}{L} + \sum_j \Delta\theta_j \sin\frac{2\pi}{\lambda_j}X \tag{2.30}$$

where

$$L = L_o e^{-\alpha t}$$

$$\lambda_j = \lambda_{jo} e^{-\alpha t} \tag{2.31}$$

The vertical velocity field can be obtained from equations (2.23) and (2.30).

$$w = \frac{\alpha g}{f^2\theta_o}\left(\frac{\Delta\theta}{L^2}\frac{2}{\pi}\frac{2(X/L)}{(1+(X/L)^2)^2} + (2\pi)^2 \sum_j \frac{\Delta\theta_j}{\lambda_j^2} \sin\frac{2\pi}{\lambda_j} X\right) \cdot$$

$$\frac{\left(\frac{H}{4}\right)^2 - \left(z - \frac{H}{2}\right)^2}{1+\frac{g}{f^2\theta_o}\left(z-\frac{H}{2}\right)\left(\frac{\Delta\theta}{L^2}\frac{2}{\pi}\frac{2(X/L)}{(1+(X/L)^2)^2} + (2\pi)^2 \sum_j \frac{\Delta\theta_j}{\lambda_j^2} \sin\frac{2\pi}{\lambda_j} X\right)} \qquad (2.32)$$

Equation (2.32) shows that the initial temperature perturbations at the mesoscale can lead to band-shaped areas of enhanced vertical motion. This could be one possible explanation for the formation of rainbands in extratropical cyclones. The vertical velocity, and therefore the rate of precipitation, caused by the mesoscale perturbations is comparable to those near the main front if

$$\frac{\Delta\theta}{L^2} \sim \frac{\Delta\theta_j}{\lambda_j^2} \qquad (2.33)$$

Since $\lambda_j << L$, even very weak perturbations in temperature may lead to very large spatial fluctuations in precipitation.

2.4 Precipitation as a Stochastic Dynamic Process

Equations (2.9) and (2.32) clearly indicate that the initial temperature field is a dominant parameter controlling the precipitation distribution in a frontal cloud system. The model indicates that weak temperature fluctuations at small spatial scales may induce large fluctuations in precipitation. A deterministic prediction or description of precipitation then requires a very precise definition of the initial temperature field.

There are many processes which may produce small scale inhomogeneities in an air mass. Without very precise information about the initial states of the participating air masses, the distribution of precipitation in the frontal cloud system of an extratropical cyclone may appear to be a highly random process. The statistical structure of the precipitation field depends very much on that of the initial temperature field. Therefore, the time and spatial evolutions of precipitation in a frontal cloud system may be considered as a dynamic process subject to random initial conditions.

3. INCORPORATING THE OBSERVED STRUCTURE OF PRECIPITATION PATTERNS INTO STOCHASTIC PRECIPITATION MODELS

3.1 Background

Hydrological analyses usually begin with observed precipitation data for a collection of raingages. Each raingage observation represents the integration over time of the rainfall at a point in space. Spatial analysis of such data produces isohyetal patterns, contoured in units of length-time^{-1} (e.g., mm of rain per 15 minutes). The time interval for the raingage data can be as short as the collection and measurement capabilities of the raingages permit, or as long as the lifetime of the storm which produced the rainfall (or even many such events, if the data are to be used for climatological purposes).

Silverman, et al. (1981) have noted that estimates of precipitation volume from isohyetal analyses obtained from raingage networks are subject to a component of variability that is caused by the random direction of movement and orientation of storms, and variations in storm structure. To study this sampling variance of raingage observations, they specified isohyetal patterns for convective raincells using continuous functions which could be integrated to determine true values of total raincell precipitation. They compared these values to the results obtained when the same simulated cells were observed by a variety of raingage networks differing from one another in density and placement of raingages and isohyetal patterns were constructed from these modeled observations.

In attempting to link the deterministic aspects of storm structure to the statistical structure of individual precipitation patterns, it seems appropriate to specify not the resultant isohyetal pattern, but the precipitation pattern, which, when integrated over time, itself produces an isohyetal pattern. Hammerstrand (1980) designed a model to determine the probability of a given duration or for accumulation of a given amount of precipitation from convective showers during a given period at a specified point, assuming that mean values of the precipitation rate and wind speed are known over a grid square from a large-scale numerical forecasting model. His probability model uses for precipitating elements circular showers of constant radius and lifetime which move with the large-scale wind and are randomly distributed within a grid square. Their rainfall rate, which is constant with radius, time, and from one shower to another, is not calculated explicitly by the probability model.

In this section, a time and space dependent precipitation pattern including both the convective cells (showers) and more widespread uniform (stratiform) precipitation typical of meso-β scale precipitation patterns is used to generate isohyetal patterns. This approach has been chosen to allow the inclusion of both types of precipitation and to illustrate how the temporal integration inherent in isohyetal patterns obscures the rather deterministic structural features of the precipitation pattern. The precipitation pattern specified here is quite a simple one. The method can, however, be generalized to produce isohyets for any precipitation. In particular, it could be used to produce isohyets for precipitation patterns generated using Eqs. 2.9 and 2.32.

3.2 The Basic Components of a Specified Precipitation Pattern

3.2.1 The cell

The precipitation patterns of convective rainshowers can be depicted in a variety of ways. The example chosen here was selected to illustrate the generality of the method while not sacrificing ease of computation. Let an individual rainshower be represented by a circular cell of radius r_c which persists for a lifetime L, and moves at a

constant velocity u. The rainfall rate within the cell depends upon position with respect to the center of the cell (r), and time (t) elapsed during the lifetime of the cell. In addition, the radius of the cell varies with time.

In functional form, the rainfall rate R(r,t) decreases linearly from the center of the cell where the rainfall rate is R_o:

$$R(r,t) = R_o \left(1 - \frac{r}{r_c}\right) \text{ for } r \leq r_c \quad . \qquad (3.1)$$

The rainfall rate at the center of the cell, R_o, increases linearly with time from zero at the beginning of the cell's lifetime to a maximum value R* at time L/4. R_o then decreases linearly to zero at time L. The radius of the cell, r_c, increases linearly with time from 0.75 r_{cmax} at the beginning of the cell's lifetime to a maximum value r_{cmax} at time L/4, whereupon r_c remains constant for the remainder of the lifetime of the cell.

Figs. 3.1 and 3.2 present the results of computations for such a cell where the adjustable parameters are chosen as follows:

L = 60 minutes u = 5 m sec^{-1}

r_{cmax} = 10 km R* = 50 mm hour^{-1}

Fig. 3.1 shows the instantaneous rainfall pattern at t = 30 minutes (t=L/2). Fig. 3.2 depicts the isohyetal pattern over the lifetime of the cell. Fig. 3.2 was obtained by integrating equation (3.1) numerically using 2-minute centered difference time steps at points separated by 2.5 km or less, then contouring the resultant field of simulated raingage observations.

The isohyetal pattern in Fig. 3.2 is elongated in the direction of motion with respect to Fig. 3.1. Perhaps less easily anticipated are the weaker gradients in accumulated precipitation where rainfall amounts are low and the stronger gradients where rainfall amounts are large. This is a feature of many observed isohyetal patterns (Silverman, et al., 1981). Of course, a realistic rainfall pattern would contain a variety of such cells in different stages of their life cycles, with some degree of randomness influencing their positions. The statistics of a collection of cells could be specified using methods like that presented in Section 2.

3.2.2 Larger areas of more uniform precipitation

As cells dissipate and meso-β scale dynamical processes become important in convective precipitation systems, areas of rather uniform precipitation rates develop. Let such an area be rectangular in shape (width w and length ℓ) with a lifetime L_m, and move at a constant velocity u. The rainfall rate within this area depends on position with respect to the center of the area and elapsed time.

In functional form, the rainfall rate R_m decreases linearly from the center of the area where the rainfall rate is R_{om}:

$$R_m = R_{om}\left(1 - \frac{|x-x_o+w/2|}{w/2}\right)\left(1 - \frac{|y-y_o|}{\ell/2}\right)$$

$$x_o - w \leq x \leq x_o$$

$$y_o - \ell/2 \leq y \leq y_o + \ell/2 \quad , \qquad (3.2)$$

where (x_o, y_o) is the position of a reference point on the leading edge of the area (see Fig. 3.3). The width of the area, w, increases linearly with time from a value of 0.2w* at t=0 to w* at $t=L_m/4$, after which w remains constant at its maximum value of w*. R_{om} increases linearly from 0 at t=0 to its maximum value of R* at $t=L_m/4$. R_{om} remains constant at R* until $t=3L_m/4$, whereupon it decreases linearly to zero at t=L. Fig. 3.3 shows the instantaneous rainfall pattern at t=L/2 when the adjustable parameters are:

R* = 16 mm hour^{-1} ℓ = 200 km

w* = 100 km.

If L and u are also specified, equation (3.2) can be integrated with time to produce an isohyetal pattern (not shown). A complete storm simulation requires specification of both cells and an area (or areas) of noncellular rainfall. Such a simulation can be completely deterministic, or some of the adjustable parameters as well as positions and times for the appearance of cells may be chosen using random variables. Thus the deterministic and random components of the stochastic nature of precipitation patterns can both be taken into account in simulating precipitation for hydrological models.

3.3 Applying Specified Precipitation Patterns to Hydrological Models

Specified precipitation patterns have several advantages over other methods of providing input to hydrological models. First, the patterns themselves can be related in space, time, and behavior (both deterministic and random) to dynamical forcing on larger scales using methods like that described in Section 2. Second, the specified patterns can be chosen to resemble those actually observed (usually using radar) over a particular watershed. Third, specified precipitation patterns can be used to generate isohyetal patterns of any desired resolution in space and time. High-resolution simulated isohyetal patterns can be used to compare the total simulated rainfall to that predicted using dynamical methods.

ACKNOWLEDGMENTS

Prof. Leary's research is supported by the National Science Foundation under Grant No. ATM-8205781. Prof. Cho's research is supported by the Canadian Natural Sciences and Engineering Research Council. Ms. Denise Bentley typed the manuscript.

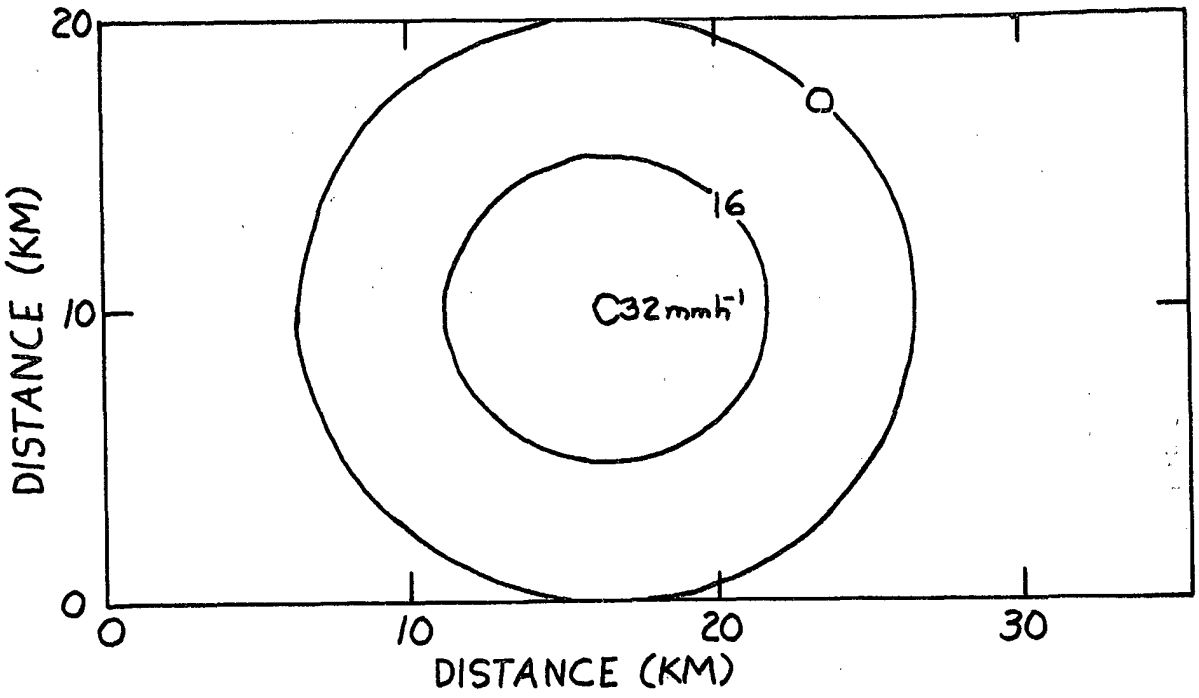


Fig. 3.1 Instantaneous precipitation rates (mm h^{-1}) for the cell described in Section 3.2.1 at t=30 minutes.

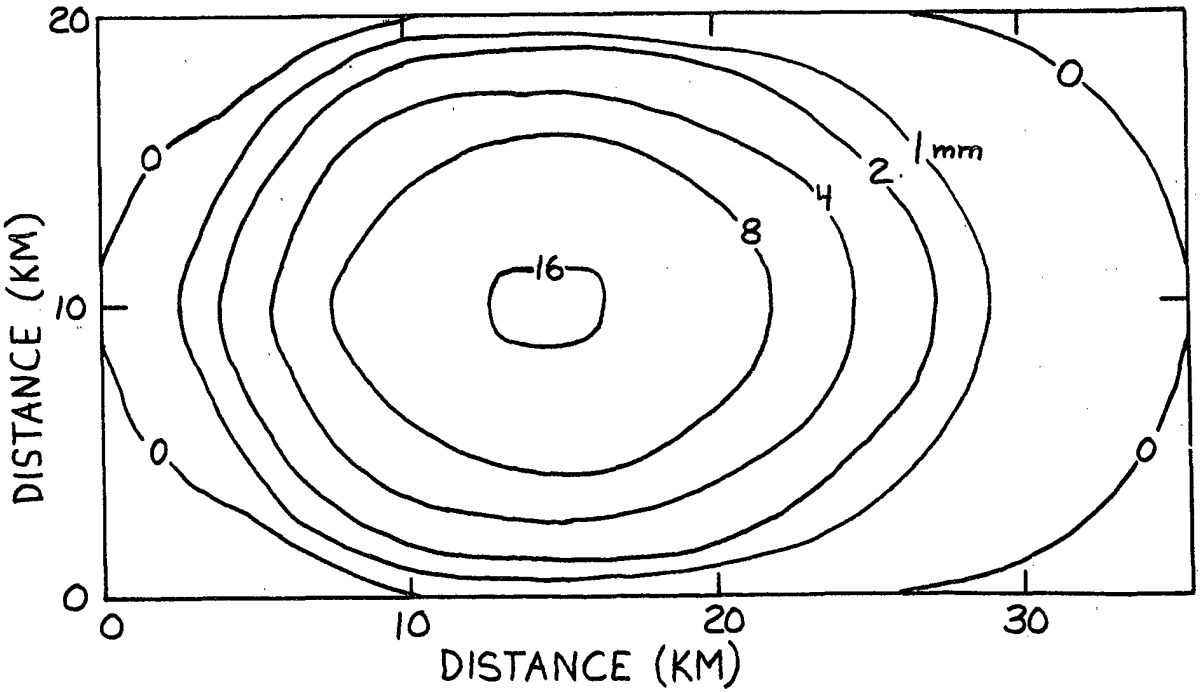


Fig. 3.2 Isohyets of total rain (mm) produced by the cell described in Section 3.2.1.

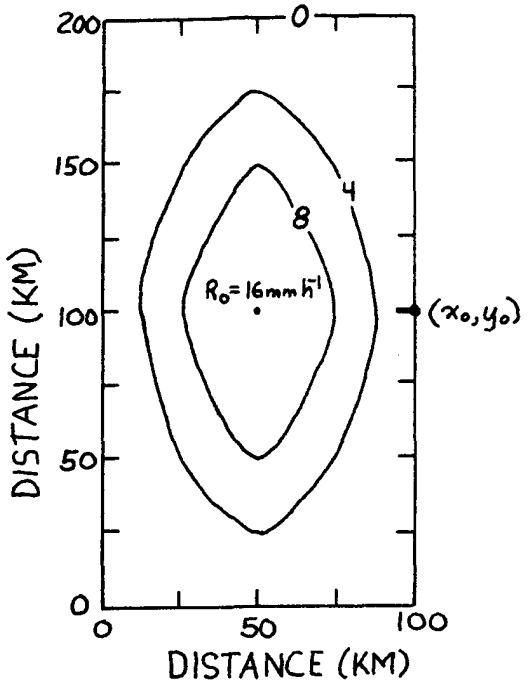


Fig. 3.3 Instantaneous rainfall rates (mm hour^{-1}) for the area of precipitation described in Section 3.2.2 at $t=L_m/2$. Outer boundary is the 0 mm h^{-1} contour. R_o, x_o, and y_o are defined in the text.

REFERENCES

Doob, J. L., 1942: "The Brownian Movement and Stochastic Equations", Annals of Mathematics, 43, No. 2, 351-369.

Hammarstrand, U., 1980: A model to predict the probability of precipitation. Mon. Wea. Rev., 108, 793-803.

Hoskins, B. J., and F. P. Bretherton, 1972: "Atmospheric Frontogenesis Models: Mathematical Formulation and Solution", J. Atmos. Sci., 29, 11-37.

Houze, R. A., Jr., 1981: Structure of atmospheric precipitation systems. A global survey. Radio Science, 16, 671-689.

Houze, R. A., Jr., and A. K. Betts, 1981: Convection in GATE. Reviews of Geophysics and Space Physics, 19, 541-576.

Houze, R. A., Jr., and P. V. Hobbs, 1982: Organization and structures of precipitating cloud systems. Advances in Geophysics, 24, 225-315.

Orlanski, I., 1957: A rational subdivision of scales for atmospheric processes. Bull. Amer. Meteor. Soc., 6, 527-530.

Silverman, B. A., K. K. Rogers, and D. Dahl, 1981: A model to predict the probability of precipitation. J. Appl. Met., 20, 1468-1478.

ARE "SCATTERED" THUNDERSHOWERS RANDOMLY DISTRIBUTED IN SPACE? AN APPRAISAL BASED ON CORRELATION ANALYSIS

David Sharon

Institute of Earth Sciences
Hebrew University of Jerusalem
Jerusalem 91904, Israel

and

Blaustein Institute for Desert Research
Ben Gurion University of the Negev
Sede Boker 88040, Israel

1. INTRODUCTION

Localized convective rainstorms are a major source of rainfall in many areas, and because of their characteristically high intensity, they have attracted special interest in hydrology, soil erosion studies and ecology.

Areal patterns of rainfall within individual storm cells have been extensively treated by meteorologists and statistical hydrologists. But little is known on the spatial distribution of the cells among themselves, i.e., the pattern of their locations relative to each other. In most cases this distribution has been described simply as sporadic, or random in space, or it has at least implicitly been treated as such for modelling purposes and the like. This is the case both in basic meteorological studies and in applicative ones, such as in hydrology.

The application of standard procedures to test the postulated randomness directly is inhibited by the existing scantiness of rainfall data, which is especially severe as far as patchy rainfall is concerned. In the present study the problem is being dealt with indirectly, by pointing to statistical evidence on the existence of systematic patterns in the distribution of rain patches in space. Data from flat, essentially uniform areas have been used, so that results do not reflect localized topographical effects or the like.

2. METHOD

The spatial patterns of rainfall areas have been studied by means of spatial correlation analysis of daily rainfall fields. Spatial correlation analysis is based on linear correlation coefficients r between daily rainfall totals in all pairs of measuring points within an appropriately delimited region. High values of r are interpreted as indicating a frequent coincidence of rainfall, i.e., that in individual events respective points are either exposed to one and the same precipitating cloud or to separate clouds but with a fixed position in space relative to each other. r values from all point pairs in the region defined are used to estimate a spatial correlation function, which is a two dimensional synthesis in which each r value appears - relative to a common origin - at the true direction and distance that exists between respective points on a map. Fig. 2 may serve as an example; it shows all r estimates obtained in a given area, through which the surface of the correlation function may be derived. The technique is described in more detail in the studies referred to below.

In case of isotropic fields, unidimensional functions may also be derived, based on distance alone. A very thorough analytical treatment of unidimensional correlation functions has been recently published by Stol (1981) and will have to be considered in any future use of correlation functions.

Results from all regions summarized below have been obtained on the basis of large data sets covering hundreds of rain events each, over areas of 100 - 150 km in diameter. The size of study areas turned out to be of critical importance for the study of the phenomenon dealt with.

3. RESULTS

Two different types of patterns have been identified hitherto in the areal distribution of rain patches in different areas. Only the main results of the studies cited are presented below.

The first type consists of an isotropic equidistant arrangement of rainfall areas, with a typical spacing of 40 - 50 km between them. This arrangement is reflected in the unidimensional correlation-distance relationship shown in Fig.1. Results of this type have been obtained in a number of low-latitude areas as well as in an extreme subtropical desert environment (Henry 1974, Sharon 1974, 1981). In all above areas rainfall originates predominantly from relatively large convective cells, quasi-stationary throughout their lifespan. There is some evidence to the effect that systematic spacings of storm cells may also prevail in summer storms in mid-latitudes, such as in some northern parts of the USA (Owens, 1966, Poellot and Brady, 1980).

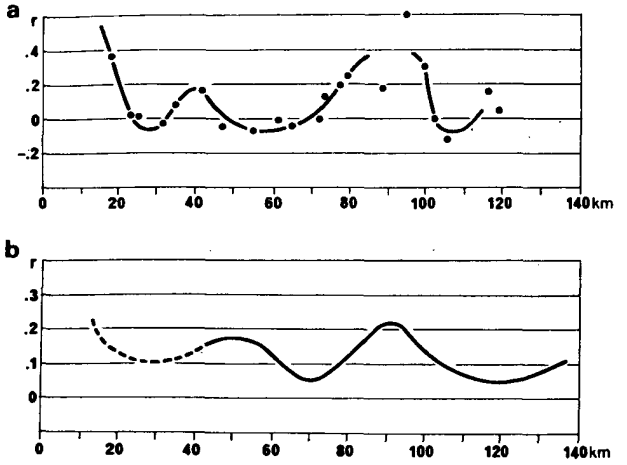


Fig. 1. Unidimensional projection (from polar coordinates) of an itsotropic correlation surface pertaining to daila rainfall fields in: (a) the Namib Desert, (b) Sukumaland, Tanzania. The graphs show peak correlations between points separated 40 - 50 km and also at a double distance (from Sharon 1974, 1981).

The above finding concerning a typical minimum distance between cells has been explained in terms of the radius around a storm, from which moisture and energy converge into the active core of a convective cell, with no additional cells developing in the clear, subsiding air within that radius. This fixed distance between cells, as well as the isotropy of the rainfall field concerned may be manisfestations of an organizing system on a greater scale in space, such as large Benard or Rayleigh cell patterns forming over relatively uniformly heated surfaces.

A quite different pattern, showing a linear arrangement of rainfall areas has been identified in the summer rainfall region of Southern Africa (Sharon 1983). The respective correlation surface reflecting this is shown in Figs. 2 and 3 (see captions for explanation). Fig. 3 also shows a sine curfe fitted to the projected r values themselves (not to the running averages). An appropriately defined objective function related to that sine curve was used to measure the pronouncedness and statistical significance of the linear trend. Using this function and changing the orientation of the projection to cover the entire range of azimuths, it was shown that a linear arrangement can be found only at the above orientation, shown in Fig. 2. The same trend was found independently in distinct sequences of years and in three separate parts of the study area, with only little seasonal variation. No conclusive explanation has been offered so far for the physical phenomenon reflected in these results.

In the above arrangement localized showers are strung out along equidistant parallel lines, the relative positions of which within the system is shown in Fig. 2. The exact geographical position of the system of lines certainly varies from day to day, as in the long range all points in the area receive comparable amounts of rain. At the same time, however, the orientation and spacing within the system can vary within narrow limits only, otherwise they would not come out in the correlation function.

A similar reasoning applies to the system of equidistant cells presented above.

4. CONCLUSIONS

The results of the analysis presented above point to the existence of meso-scale systems that impose relatively rigid systematic arrangements of localized storm cells in certain areas and/or certain seasons. In all these cases cells developing on the same day are certainly not randomly distributed in space.

The above result does not imply that cells necessarily develop at any geographically fixed points. On the contrary, correlation functions of the above type can be derived only in regard to areas within which statistical stationarity can be assumed. Hence, the above results imply that within such areas storm cells occur equally frequently at any point, with only their relative position within the system being fixed. This means that while the relative location of storm cells within a system essentially follows systematic patterns, the system as a whole can develop at varying positions over an area. Hence, when looking at single cells only, or at small areas, randomness may be inferred. Here, as in some other atmospheric research problems, "the choice of the horizontal extent of an observational experiment plays an important role.... in order .... to generate meaningful results" (Orlansky, 1975).

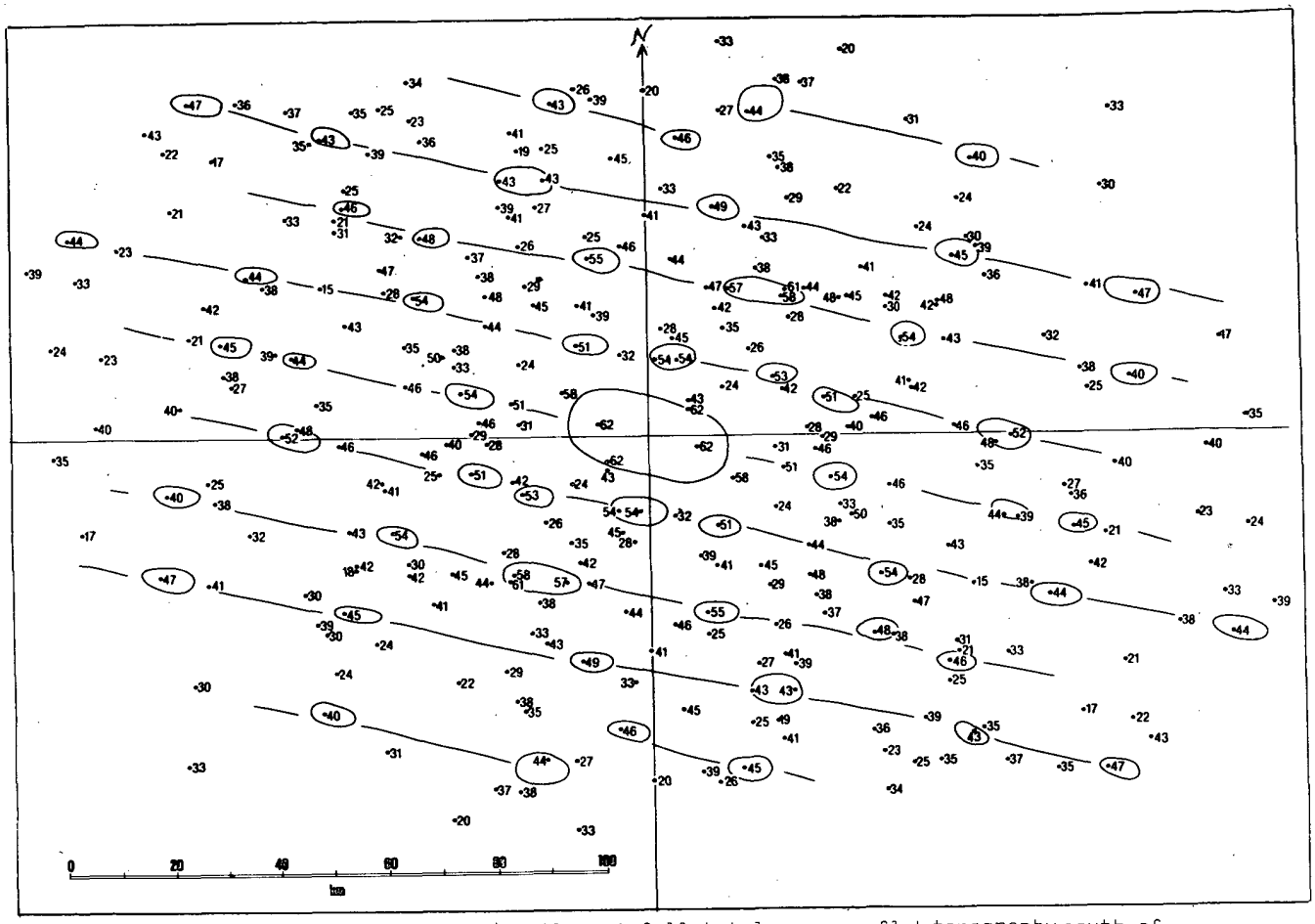


Fig. 2. A correlation function of daily rainfall totals over a flat topography south of Johannesburg. Correlation peaks (in circles) reflect relative positions of points with the most frequent simultaneous incidence of rainfall. Locations are shown relative to the origin (in the centre), preserving the real direction and distance between corresponding stations. The correlation surface shows "ridges" of peaks along lines running approx. from WNW to ESE.

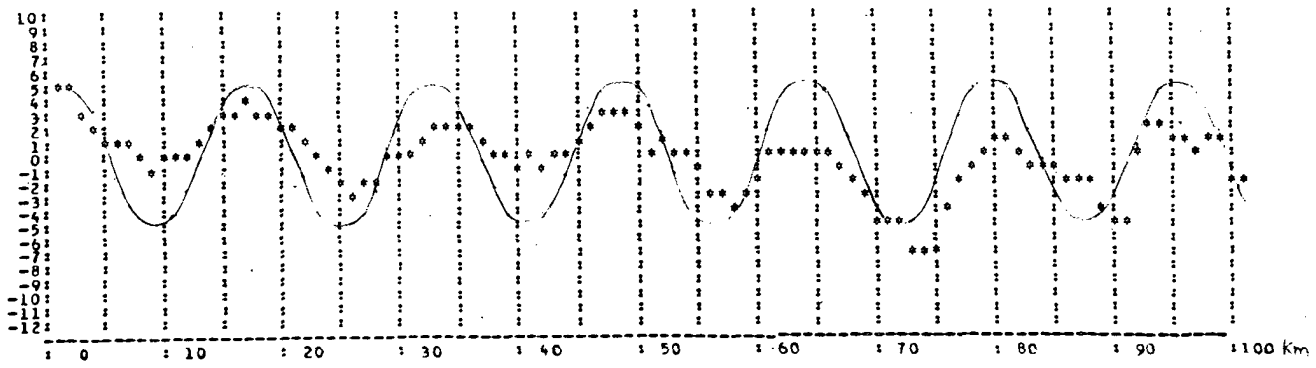


Fig. 3. Two smoothed projections of the surface in Fig. 2 on an (upright) plane, positioned perpendicularly to the storm lines: 5 km running averages of r values (asteriks), and a sine curve fitted to the r values; projected r values themselves are not shown. The 5 or 6 subsequent peaks reflect the respective "ridges" lying symmetrically north and south of the origin in the centre of Fig. 2 above. Values on the ordinate are deviations from the distance-dependent mean r, in hundredths (from Sharon, 1983).

From the applicational point of view, features such as the spacing of storms, preferred directions between them etc. are meaningful for probabilistic modelling, such as of the rainfall input over a watershed or over the catchment area of an urban sewage system.

5. REFERENCES

Henry, W.K. 1974: The tropical rainstorm. Mo. Wea. Rev. 102, 717-725.

Orlansky, I. 1975: A rational subdivision of scales for atmospheric processes. Bull. Am. Met. Soc. 56, 527-530.

Owens, Jr., J.J., 1966: The distribution of air mass thunderstorms in New England. M.Sc. thesis, Dept. of Meteorology, Mass. Inst. Techn., Cambridge, Mass.; cited in Sharon, 1974 (below), p. 203 and 214.

Poellot, M.R. and Brady, P.J., 1980: The spacing of radar echoes over the Montana HIPLEX area. Amer. Met. Soc., 19th Conf. on Radar Meteor., April 15-18, 1980, Miami, Fla., p. 398-400.

Sharon, D. 1974: The spatial pattern of convective rainfall in Sukumaland, Tanzania. Arch. Meteor. Geophys. Bioklim. 22B, 201-218.

Sharon, D. 1981: The distribution in space of local rainfall in the Namib Desert. Jour. Climatol. 1, 69-75.

Sharon, D. 1983: The linear organization of localized storms in the summer rainfall zone of South Africa. Mo. Wea. Rev. 111, No. 3.

Stol, Ph. Th., 1981: Rainfall interstation correlation functions, Pts. I-II, J. Hydrol. 50, 45-71 and 73-104; and Pt. III, J. Hydrol. 52, 269-289.

STATISTICAL METHODS FOR THE ANALYSIS OF CLIMATIC DATA

Roger D. Stern

Department of Applied Statistics
University of Reading, Whiteknights
Reading RG6 2AN, England

1. INTRODUCTION

This review gives a personal view of a range of topics in statistical climatology. Generalized linear models are presented as an effective framework for the analysis of data from individual sites. This is followed by a more general discussion on methods of spatial analysis and on the use of multiple regression in studying crop weather relationships. The intention is to provoke discussion in each of these areas, rather to give a comparative review of the merits of particular methods.

Much of the literature in statistical climatology is broadly one of two types. The first consists of research papers, mainly by statisticians, suggesting new distributions, methods or models. This research is largely ignored in the second category, which includes a wide variety of papers and projects that have not involved statisticians, and have usually used only relatively simple statistical methods. The challenge is for more statisticians to try to occupy the middle ground, producing genuinely usable models and becoming involved in project-work to make constructive use of standard statistical methods when they are adequate and to encourage the adoption of new techniques when necessary.

It is with this challenge in mind that generalized linear models are presented as an area of statistical research that is too important to be ignored in applied projects. Some further research and case studies are needed to compare methods of spatial analysis while the priority in the use of multiple regression to study crop-weather relationships is for standard techniques to be used more effectively.

2. GENERALIZED LINEAR MODELS

The scope of generalized linear models, described by Nelder and Wedderburn (1972), can be seen most clearly when compared with linear models used for the analysis of normally distributed data. Very broadly it is now possible to analyse data from any distribution within the exponential family with the same flexibility that has previously been reserved for normally distributed data. Classical linear model theory can be summarised as follows.

1) The observations X_i are independent.

2) They are normally distributed with mean μ_i and common variance σ^2.

3) The mean $\underline{\mu}=T\underline{\beta}$, where T is a matrix of "explanatory variables" and $\underline{\beta}$ is a vector of unknown parameters.

The maximum likelihood estimates of $\underline{\beta}$ can then be found by solving the least squares equations,

$$T'T\underline{\beta}=T'\underline{X}.$$

For generalized linear models, 2) and 3) change as follows.

2) The observations X_i are from a distribution which is a member of the exponential family.

3) The expected value of X_i can be related to the explanatory variableon h, so that $\underline{\mu}=h(T\underline{\beta})$.

Nelder and Wedderburn (1972) show that the maximum likelihood estimates of β may still be found by solving a set of equations of the type

$$A\underline{\beta}=\underline{r}.$$

The structure of both A and $\underline{r}$ depend on the distribution of X and on the function h. In general, both also involve $\underline{\beta}$ so the method of solving the equations is iterative. When a model is fitted, a quantity called the "deviance" is calculated. This is a generalization of the residual sum of squares, which is calculated when fitting models to normally distributed data. The fact that maximum likelihood estimation is being used allows the relative importance of fitting different numbers of parameters to be assessed with likelihood ratio tests. This is done by comparing the deviances for models with and without the particular parameters.

The process of fitting generalized linear models is straightforward using the computer package GLIM, (Baker and Nelder, 1978). In the examples given below, the technical details of the fitting have largely been omitted in order to emphasise the wide range of problems that can be considered within this single framework. In addition, the language facilities within GLIM allow the class of models to be widened even further. This is exploited in the examples on the extreme value distribution and on the negative binomial distribution, neither of which is a member of the exponential family.

2.1 Fitting Markov Chain Models to Daily Rainfall Occurrence

There is a large literature on this subject which has recently been reviewed by Waymire and Gupta (1981). Gabriel and Neumann (1962) used a first order, two-state Markov chain. The many extensions to their model include the fitting of higher order chains (Gates and Tong, 1976), the fitting of chains with more than two states (Haan et al, 1976) and the fitting of non-stationary chains. Non-stationarity has been considered by fitting separate chains to different periods of the year (Jackson, 1981), and by fitting continuous curves to the transition probabilities (Woolhiser and Pegram, 1979). Many of these studies have involved the construction of special computer programs. The use of GLIM to fit these models is described by Coe and Stern (1982) and the fact that it is no longer necessary to write special programs should encourage more people to consider fitting, testing and using such models.

At most sites that we have considered, a first or second order chain has provided an adequate model. However, a site at which a high order chain is required was chosen deliberately for the example in Fig. 1 in order to illustrate the range of models that can be considered. The model is a fifth order chain in which each of the curves of the probability of rain against the time of year is a fourth order Fourier series (on the logit scale). The 'full' model for a fifth order chain requires 32 separate curves. However the prohibitive number of parameters that this would have involved were not necessary for two reasons.

1) Although the probability of rain on any day was less after increasingly long dry spells, there was no evidence of any difference in the probabilities after rainy days. Six curves were found to be sufficient, denoted $p_0(t)$ ... $p_5(t)$ where $p_i(t)$ $i = 0,...4$ is the probability of rain following exactly i dry days and $p_5(t)$ is the probability of rain after a dry spell of 5 or more days.

2) The further constraint that all the curves except $p_0(t)$ were parallel, on the logit scale, also proved to be reasonable.

Thus, in total, the final model required only 22 parameters to be fitted, i.e. nine parameters for $p_0(t)$ and $p_1(t)$ plus one extra parameter for each of $p_2(t)$ ... $p_5(t)$. It therefore provided a very consise summary of the chance of rain on a daily basis.

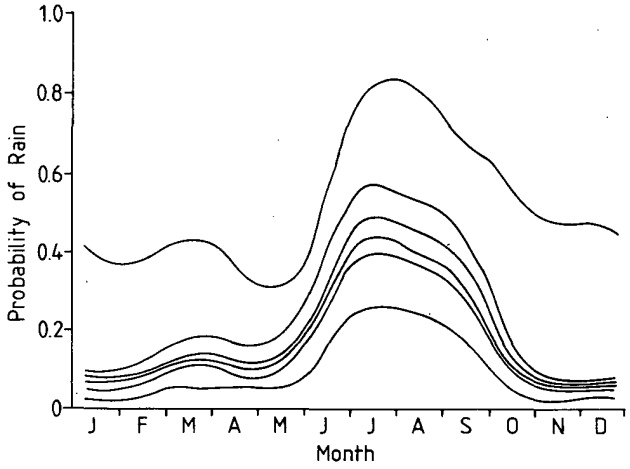


Fig. 1. Nagpur, India (1901-76). Estimated proportions of rainy days for a restricted fifth order Markov chain. From the top down the curves estimate $p_0(t)$ ... $p_5(t)$ respectively.

2.2 Loglinear models

The analysis of data for multidimensional contingency tables is a very large subject in itself, e.g. Fienberg (1980). Much of the recent work has emphasised the use of loglinear models. This is another instance of generalised linear models and a simple example is presented. Following the approach used by Mooley (1971), monthly rainfall totals for Hyderabad, India, are summarised in a 4 way table, Table 1. The data for each month are coded as 'low' if they are below the median value for that month and as 'high' otherwise. There were, for example, 6 years in which the rainfall was 'low' in each of the 4 months and just 1 year in which June rainfall was 'high' and July to September were all 'low'. It is then a simple routine to fit a model to the data, using either GLIM or any other package that has a procedure for loglinear models. The 'full' model contains terms which correspond to interactions between the different months. In this case, no interaction terms were required, i.e. the data showed no evidence of any dependence between successive months, (Stern and Patel, 1980).

Table 1. Relationship between values below and above the median values for Hyderabad for June, July, August and September. Seventy eight years of data 1893-1970.

	July L	July H	July L	July H	July L	July H	July L	July H
June Low	6	4	2	6	9	5	5	2
June High	1	5	8	7	4	5	4	5
	Aug.Low	Aug.High	Aug.Low	Aug.High	Aug.Low	Aug.High	Aug.Low	Aug.High
	September Low				September High			

2.3 Fitting Extreme Value Distributions

It is often important to consider events such as maximum wave heights, wind speeds or rainfalls. Extreme value distributions are often fitted. The probability density function for the Type 1 extreme value distribution can be written,

$$p(x)=(1/\alpha)\exp[-(x-\theta(t))/\alpha-\exp-(x-\theta(t))/\alpha]$$
$$-\infty<x,\theta(t)<\infty,\ \alpha>0.$$

This distribution is not a member of the exponential family. However, for any given value of the scale parameter ,α, the distribution of $z=\exp(-x/\alpha)$ is exponential with mean $\exp(-\theta(t)/\alpha)$. Thus the fitting could proceed simply by considering different values of α iteratively, using GLIM for each, with an exponential error structure and a log link. A more general approach is considered by Aitken and Clayton (1980). The process of fitting such models has been simplified by Roger and Peacock (1982) and their method was used to give the results below.

The annual flood data, 1909-1960 for the Bow River at Banff, Alberta, given by Bruce and Clark (1966), are used as an example. The values range from 4290 cfs in 1949 to 13300 cfs in 1923. Three alternative models are considered. The first fits a single distribution to the data. The second allows for the possibility of a linear trend in the parameter θ through the period of the record. The third fits a different value of θ to the first and second half of the series. The results are given in Table 2. In this case neither refinement improves the fit.

Table 2. Estimate of parameters for alternative models based on the type 1 extreme value distribution.

1. Single Distribution $\theta(t)=6830$ $\alpha=1801$

2. Linear Trend $\theta(t)=7820-35.3t$ $t=1...53$ $\alpha=1730$

3. Step Function $\theta_0=7426$, $\theta_1=6349$, $\alpha=1723$ (1908-34) (1935-60)

From the differences in deviance, neither the linear trend nor the step function improved the fit of the model

2.4 Distribution of Dry Spell Lengths

The negative binomial distribution is widely used, (Johnson and Kotz, 1968). In the analysis of climatological data it is a popular distribution for modelling spell lengths. The distribution considered here is the truncated negative binomial, (TNBD),

$$P(x(t)=k) = {}_{k+r-1}C_k\, p(t)^r(1-p(t))^k/(1-p(t)^r),$$

$$k=1,2,...$$
$$r>0, p(t)>0, t=1,...366.$$

In many studies the observed distribution of dry spell lengths is considered separately for the different months and a single distribution is fitted to each set of data. In some studies, e.g. Buishand (1977) the sets of fitted parameters through the year are then smoothed. Buishand also makes the assumption, used here, that the parameter r remains constant through the year while p(t) is allowed to vary.

Daily rainfall data from Lunuwila, Sri Lanka, were used to illustrate the analysis. A dry day was defined to be one with less than 1mm of rain. Dry spells ending in successive 10 day periods were associated with the middle day number number of the period. A first stage in the analysis was to fit separate TNBD's to the dry spell lengths in each 10 day group. The initial value of r was taken as the mean of these individual estimates. For the Lunuwila data this gave a value of r=0.42, corresponding to a distribution 'between' the log-series, for which r=0 and the truncated geometric, for which r=1. With this value of r, GLIM was then used to estimate p(t). Wth r fixed, the TNBD is a member of the exponential family, but not one of the standard options in GLIM. However there are facilities within the package for user defined models. The model adopted involved fitting Fourier series with three harmonics to $\theta(t)$,

where $p(t)=1-\exp(\theta(t))$,

Strictly the fitting process should proceed iteratively, fitting a new value of r for the given p(t).

The mean spell length from the model is given by

$$r(1-p(t))/(p(t)(1-p(t)))$$

and is plotted in Fig. 2 together with the observed means.

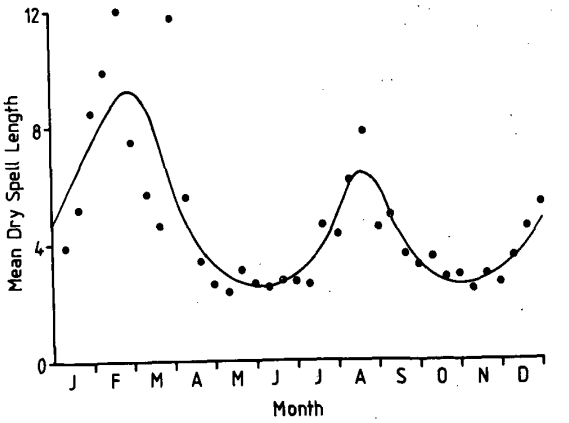


Fig.2 Lunuwila, Sri Lanka (1950-79). Observed and fitted values of the mean dry spell length.

2.5 Summary

In addition to general statistical research into generalized linear models (e.g. Gilchrist, 1982), some publications are now demonstrating the importance of this subject within specific areas of application, (e.g. Adena and Wilson, 1982). In statistical climatology it is hoped that the methodology will at least cause a decline in the number of studies that

arbitrarily split the year into different periods and then compare, in great detail, the fit of single distributions to climatic variables separately in each period. In addition, many studies emphasise the non-normality of the climatic variables being considered. The concept of generalized linear models facilitates the easy analysis of many non-normally distributed variables, and so is of considerable importance.

3. AREAL ANALYSIS

For many applications the analysis of climatic data from individual sites is a preliminary to the assessment of the corresponding parameters for given areas. It is convenient to discriminate broadly between methods that are used to classify areas, and the methods used to construct maps. These are considered in turn.

3.1 Classification

The value of being able to classify areas that have similar climates is obvious. A variety of multivariate methods have been used to "cluster" sites. For example Ayoade (1977) used both cluster and factor analysis on Nigerian data. Gadgil and Joshi (1980) applied principal component analysis to the mean pentad rainfall of 145 stations in India. Other studies include McBoyle (1971, 1972) on European and Australian data, and Gregory (1965) on data from Sierra Leone.

In general these classification methods have two main limitations. The first is that they do not use the spatial information, i.e. the relative positions of the observations, and often re-discover that sites close together geographically also have a similar climate. The second limitation is that, although the analysis can consider a range of important climatic variables together, at some stage a single "index" of distance between each pair of stations is produced. Great care is needed to ensure that this index is informative, because the subsequent classification is a summary of this set of distances or correlations.

Two further techniques that may be useful are canonical correlations (Glahn, 1968) and multi-dimensional scaling (Coxon, 1982; Davies and Coxon, 1982). However, one common feature is the large amount of both research and computing time that each of the multivariate studies has involved. The payoff from such studies is not obvious when compared to much simpler methods using some of the many climatic indices that have been proposed (e.g. ICRISAT 1980).

3.2 Mapping

It is important to distinguish between mapping a variable such as height, where the measurements are made precisely and the problem is therefore one of interpolation, and the mapping of indices based on measurements of random variables. Many of the simple maps that are drawn routinely, based on for example mean annual rainfall, treat the observations as though they are fixed. These maps often provide all that is required though in some cases the years for which data are available differ for the different sites (e.g. Kowal and Knable, 1972) and it is not obvious that any adjustment has been made for this. Several studies, (e.g. Rainbird, 1967; Craddock, 1979) describe methods for adjusting records to a common base period.

Where the network of observations is relatively sparse, methods of mapping which ignore the spatial correlations between sites may be adequate (Davis, 1973). Coote and Cornish (1958) and Clarke and Edwards (1972) discuss the use of the analysis of variance for data from a number of sites. In such studies, a non-orthogonal analysis, which provides estimates adjusted for the different years of record from the different sites, is now straightforward.

Many general techniques for mapping have been developed (Ripley 1981). Creutin and Obled (1982) evaluate a number of these for the mapping of rainfall patterns and conclude by recommending Gandin's optimal interpolation method (WMO, 1970). Their evaluation is based largely on a detailed case study and demonstrates very clearly the value of such case studies in comparing different methods.

Further theoretical work and case studies, could, in addition to continuing the process of evaluating different methods of mapping, also consider the size of standard errors of, for instance, contour levels.

4. CROP WEATHER RELATIONSHIPS

This is a huge topic. The methods of analysis differ, depending on the scale at which the crop and weather data have been collected. Crop measurements may be made on individual leaves or given for a district. The discussion here is limited to the larger scales, those for which multiple regression is usually used to study the relationships between crop yield and climatic variables.

There is a large literature on this subject and Baier (1977) gives a good review of some of the many equations that are now in use. While there are obvious dangers in being too general, there are however also many studies that have consumed large amounts of both research and computing time, but have achieved little, due often to an uncritical use of regression techniques. As an example it is instructive to read Fisher (1924), Runge (1968) and Huda et al.(1976) in that order. These studies are on 60, 54, and 12 years of data respectively. The methods used are similar. However the later papers have used the methods less effectively in parallel with the progressively fewer years of data that were considered.

In many studies a number of stages in the analysis can be distinguished. There is usually an initial time series analysis particularly to model the trend in the yield data. Then there are often far too many climatic variables to include in a regression study and they have

therefore to be "screened" in some way. There are also, at this stage, considerable problems in deciding how to aggregate climatic data if the yield figures are given at, say, a district level. It would be valuable to have more studies such as Cornish (1950) which are conducted at different levels (Cornish considers results both for individual research stations and on a larger scale). This would clarify what is lost by the aggregation of both the crop and weather data. The final stage in the analysis is the multiple regression study itself.

Overall this is an area in which many researchers may not realise how many traps there are for the unwary. It is clearly important that statisticians are more closely involved in such studies despite the fact that there will often be little scope for new statistical methods. There are only a few problems about which statisticians might argue. For example Cornia and Pochop (1975) conduct a principal component analysis on the climatic variables and use the principal components in the multiple regression study. While this is a well recognised technique, I continue to be sceptical about its practical value. However the limitations in technique in many papers are standard problems about which all statisticians would agree. Thus the excellent review by Pittock (1977) gives a number of 'ground rules', three examples of which are as follows.

1) Test the results on one or more independent data sets or subsets of the original data.

2) Estimate the practical significance of the result, for example, the fraction of the relevant total variance which can be predicted or "explained".

3) Carefully explain the properties and limitations of the data, the statistical methods used (including data selection and smoothing), and any assumptions, reservations or doubts.

Pittock's paper is actually a discusson of methods of assessing the possible relationships between weather and sunspots but a review of the statistical methods used in studying crop weather relationships would certainly emphasise the same points.

Within the multiple regression studies, many of the problems are standard. Authors are often over-simplistic in their interpretation of the results. A common mistake is to take multiple correlation coefficients at their face values, and this ignores the fact that they are from an equation that has been selected as the best. This is related to the tendency to fit too many parameters, an extreme example of which is Welch (1964) who fits 27 parameters to 28 observations. One topic that is not considered in enough detail in many text books (or computer packages) is the comparison of regression equations. This is also a much neglected subject in many crop weather studies where data are available for a number of neighbouring sites or regions.

Overall the statistician surveying the literature on crop-weather relations is likely to be left with the impression that the subject is a difficult one, but that, even allowing for this, many studies could be substantially improved if standard statistical methods were applied more effectively. It is however very easy for statisticians to criticise. Such criticism can be interpreted as reflecting very badly on the way that we teach the subject. Perhaps more generally it reflects a lack of willingness among many statisticians to translate theory into practice and to communicate these practical skills effectively to non-statisticians.

5. CONCLUSION

Before the next International Meeting on Statistical Climatology most of the standard statistical packages will be widely available on relatively inexpensive computers. This will present tremendous opportunities for more effective analysis of climatological data and the production of constructive case studies. It is also very important to adapt our methods of teaching statistics to the availability of this new technology. A great deal of the teaching of the mechanics of statistical calculations can be replaced by a direct access to the real problems that occur in the analysis of climatological data. Many teachers of statistics will see obvious dangers in this approach. However these dangers must be compared to those of existing courses. Many basic text books and presumably the corresponding service courses in statistics, concentrate almost exclusively on the methods of fitting models with only a few unknown parameters. It should be emphasised that many applications of statistical methods to climatic data present difficult problems. It is perhaps not surprising that some researchers are unable to unravel all the complexities of conducting a multiple regression study if their statistics courses have only involved fitting very simple correlation and regression models to artificially small data sets.

6. ACKNOWLEDGMENT

I am very grateful to colleagues in the Department of Applied Statistics at Reading for their comments and assistance.

7. REFERENCES

Adena, M.A., and Wilson, S.R., 1982: Generalised linear models in epidemiological research. Report published by The Instat Foundation, Sydney, Australia.

Aitken M., and Clayton, D., 1980: The fitting of exponential, Weibull and extreme value distributions to complex censored survival data using GLIM. Appl.Statis. 29, 156-163.

Ayoade, J.O., 1977: On the use of multivariate techniques in climatic classification and regionalization. Arch.Met.Geoph.Biokl. Series B, 24, 257-267.

Baier, W., 1977: Crop-Weather models and their use in yield assessments. WMO Technical Note No.151 (WMO No.458).

Baker, R.J., and Nelder, J.A. 1978: GLIM Manual. Release 3. Oxford: Numerical Algorithms Group.

Bruce, J.P., and Clark, R.H., 1966: Introduction to Hydrometeorology, Pergamon.

Buishand, T.A., 1977: Stochastic modelling of daily rainfall sequences. Medelingen Landbouwhogeschool, Wageningen.

Clarke, R.T., and Edwards, K.A., 1972: The applicaton of the analysis of variance to mean areal rainfall estimation. J. Hydrology, 15, 97-112.

Coe, R., and Stern, R.D., 1982: Fitting models to daily rainfall data. J.Appl.Meteor.,21, 1024-1031.

Coote, G.G., and Cornish, E.A., 1958: The correlation of monthly rainfall with position and altitude of observing stations in South Australia. Division of Mathematical Statistics Technical Paper No.4, CSIRO, Australia.

Cornia, R.L., and Pochop, L.O., 1975: Principal components and crop weather analysis. Transactions of ASAE, 18(2) 335-339.

Coxon, A.P.M., 1982: The users guide to multidimensional scaling. Heinemann.

Craddock, J.M., 1979: Methods of comparing annual rainfall records for climatic purposes. Weather, 34, 332-346.

Creutin, J.D., and Obled, C., 1982: Objective analyses and mapping techniques for rainfall fields: An objective comparison. Water Resources Research, 18, 413-431.

Davies, P.M., and Coxon, A.P.M., Eds., 1982: Key texts in multidimensional scaling. Heinemann.

Davis,J.C., 1973: Statistics and data analysis in Geology, Chapter 6: Map analysis. Wiley, New York, 298-411.

Fienberg, S.E., 1980: The analysis of cross-classified categorical data. The MIT Press.

Fisher, R.A., 1924: The influence of rainfall on the yield of wheat at Rothamsted. Phil. Trans. Roy. Soc. London Ser. B, 213, 89-142.

Gabriel, K.R., and Neumann, J., 1962: A Markov chain model for daily rainfall occurrence at Tel Aviv. Quart. J. Royal Meteor.Soc. 88,. 90-95.

Gadgil, S., and Joshi, N.V., 1980: Use of principal component analysis in ratical classification of climates. In Climatic Classification. ICRISAT, 17-26.

Gates, P., and Tong, H., 1976: On Markov chain modelling to some weather data. J.Appl.Meteor., 15, 1145-1151.

Gilchrist, R.,(Ed.) 1982: GLIM 82: Proceedings of the International Conference on Generalised Linear Models.

Glahn, H.R., 1968: Canonical correlation and its relationship to discriminant analysis and multiple regression. J. Atmospheric Sciences, 25, 23-31.

Gregory, S., 1965: Rainfall over Sierra Leone. Univ. of Liverpool, Dept. of Geography Research Paper No.2.

Haan, C.T., Allen, D.M., and Street, J.O., 1976: A Markov chain model of daily rainfall. Water Resources Research, 12, 443-449.

Huda, A.K.S., Ghildyal, B.P., and Tomar, V.S., 1976:Contribution of climatic variables in predicting maize yield under monsoon conditions. Agric.Meteor., 17, 33-47.

ICRISAT, 1980: Climatic classification: A Consultants' Meeting. 14th-16th April, 1980, ICRISAT Centre, Hyderabad, India.

Jackson, I.J., 1981: Dependence of wet and dry days in the tropics. Arch. Met. Geoph.Biokl.,Series B, 29, 167-179.

Johnson, N.L., and Kotz, S., 1969: Distributions in Statistics: Discrete distributions. Houghton Mifflin.

Kowal, J.M., and Knabe, D.T., 1972: An agroclimatological atlas of the Northern States of Nigeria. Ahmadu Bello University Press, Zaria, Nigeria.

McBoyle, G.R., 1971: Climatic classification of Australia by computer. Aust.Geogr. Studies, 9, 1-14.

McBoyle, G.R., 1972: Factor analytic approach to a climatic classification of Europe. Climatological Bull., 12, 1-11.

Mooley, D.A., 1971: Independence of monthly and bimonthly rainfall over Southeast Asia during the summer monsoon season. Mon. Wea. Rev., 99, 532-536.

Nelder, J.A., and Wedderburn, R.W.M., 1972: Generalised linear models. Proc. Royal Statist. Soc., Series A, 135, 370-384.

Pittock, A.B., 1978: A critical look at long term sun-weather relationships. Reviews of Geophysics and Space Physics, 16(3), 400-420

Rainbird, A.F., 1967: Methods of estimating areal average precipitation. Report No.3, WMO/IHD.

Ripley, B.D., 1981: Spatial Statistics. Wiley. London.

Roger, J.H., and Peacock, S., 1982: Fitting the scale as a GLIM parameter for Weibull, extreme value, logistic and log-logistic regression models with censored data. GLIM Newsletter, 1982 May.

Runge, E.C.A., 1968: Effects of rainfall and temperature interactions during the growing season on corn yield. Agric. J., 60, 503-507.

Stern, R.D., and Patel, A., 1980: Comments on "Independence of monthly rainfall over Southeast Asia during the summer monsoon season". Mon. Wea. Rev., Vol. 108, No.11, November 1980.

Waymire, E., and Gupta, V.K., 1981: The mathematical structure of rainfall representation. 1. A review of the stochastic rainfall models. Water Resources Research, 17, 1261-1272.

Welch, C.F., 1964: Variable rainfall temperature and yearly tobacco yields. Tobacco Science, 158, 17-20.

Woolhiser, D.A., and Pegram, G.G.S., 1979: Maximum likelihood estimation of Fourier coefficients to describe seasonal variations of parameters in stochastic daily precipitation models. J. Appl. Meteor., 18, 34-42.

WMO, 1970: The planning of meteorological station networks. Tech. Note No.111 (WMO No.265).

USEFUL MODELS OF TEMPERATURE DATA

Richard Coe

Department of Applied Statistics
University of Reading

1. INTRODUCTION

A considerable amount of work has been done on modelling most climatic variables and daily mean temperatures are no exception. The aim in this paper is to indicate some of the problems in modelling daily temperatures and show that very simple models are adequate for some aspects of interest to agriculture.

The data come from Busan, Korea, which was chosen for the unusually long record of daily data available - 73 years (1905-1977). Daily mean temperatures at Busan vary from around 2.0°C in January to 25.5°C in August. Throughout this paper the observed temperatures are denoted y_{it}, i=1...73, t=1...365, which are considered as observations on random variables Y_t, t=1...365.

In some applications daily maximum or minimum temperatures will be of more interest than means but the approach to analysis used in this paper will be appropriate.

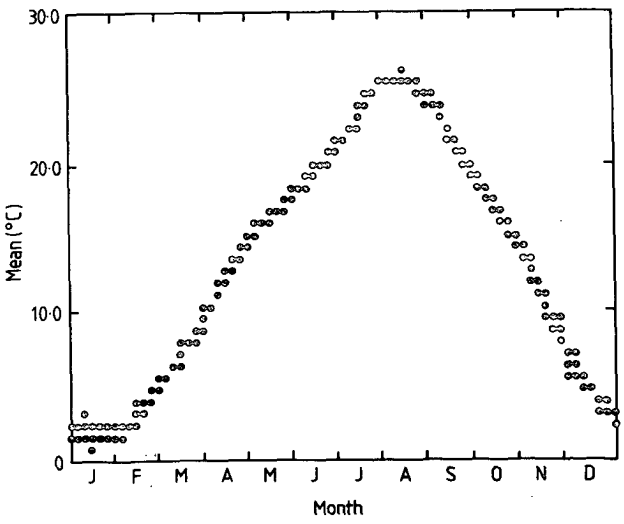


Fig. 1 Daily mean temperatures, $\bar{y}_t$, for Busan, Korea.

2. THE DATA

Before discussing any models it is worth looking at features of the data that comprehensive models might attempt to describe.

The variation in daily mean temperatures through the year, $\bar{y}_t$, t=1...365, is plotted in Fig.1.

The variance is also non-stationary. The within- day variances,

$$s_t^2 = \sum_{i=1}^{73}(y_{it}-\bar{y}_t)^2/72$$

show a considerable range (Fig. 2) with the variance in January being seven times that in June. The variation through the year is less smooth than for the mean, with the curious hump in July.

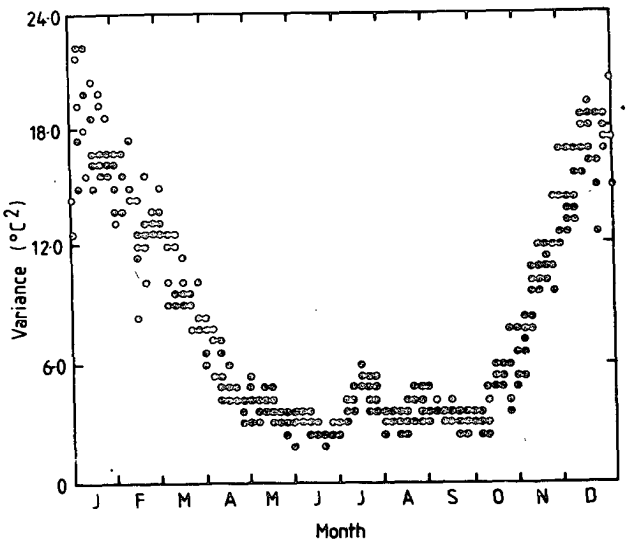


Fig. 2 Sample variances of temperatures, for each day of the year.

The correlation structure is often the hardest

component of temperature variation to model well. In many areas of application the autocorrelations in time series have to be assumed stationary as only a single realisation of the process is available. In the present example, however, assuming the autocorrelations are constant from year to year gives sufficient replication to estimate values for each day of the year. Fig.3 shows the first order autocorrelations,

$$r_{1t}=[\sum_{i=1}^{73}(y_{it}-\bar{y}_t)(y_{it-1}-\bar{y}_{t-1})]/(72s_t s_{t-1}), \quad t=2..365$$

These are clearly non-zero and non-stationary. Higher order autocorrelations are progressively smaller and show the same non-stationarity.

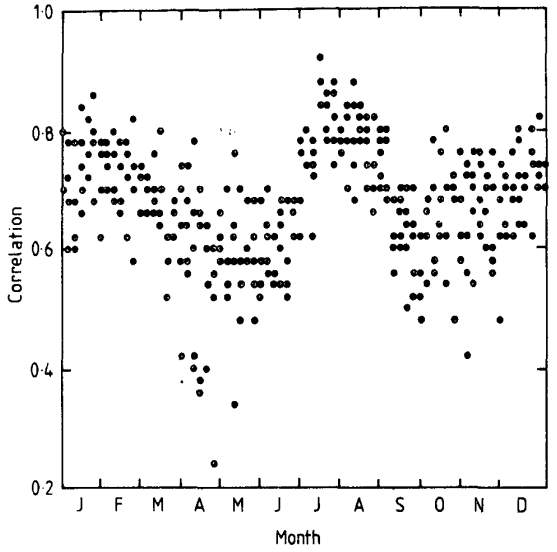


Fig. 3 Sample first order autocorrelations of daily temperatures, r_{1t}, t=1...365.

A model commonly used to represent the autocorrelation structure is the autoregressive process of order m,

$$y_{it} = \mu_t + \epsilon_{it}$$

where

$$\epsilon_{it} = \sum_{j=1}^{m} a_{jt}\epsilon_{it-j} + \nu_{it}$$

and the successive ν_{it} are independent. If m=1 then a_{1t} can be estimated by r_{1t} (Fig.3). If m=2, a_{1t} and a_{2t} are estimated from r_{1t} and r_{2t}. These estimates are plotted in Fig.4 and Fig.5 which show that for most of the year a_{2t} is non-zero and a second order autoregresive model will be more appropriate than a first.

Another aspect of temperature data that might be modelled is the dependence on other climatic variables, rainfall being the most commonly used (e.g., Dumont and Boyce, 1974, Nicks and Harp, 1980). At Busan both the mean temperatures and the variances depend on whether a day is wet or dry, with wet days warmer in Winter and cooler in Summer than dry days. Table 1 gives the values, averaged over months for clarity. The wet-day variance is lower than that for dry days in Winter and about equal in Summer (Table 1).

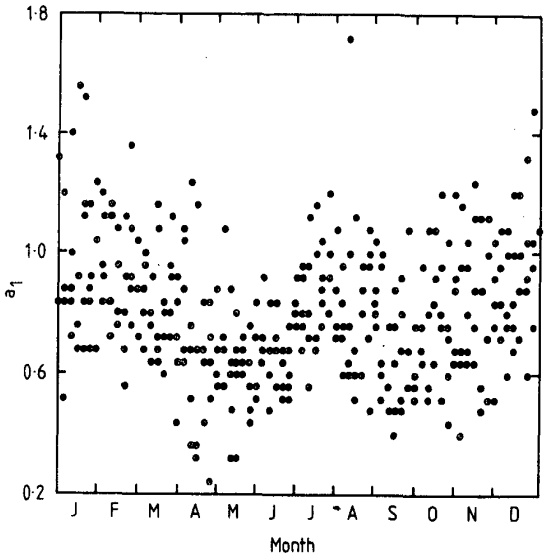


Fig. 4 Estimates of parameters a_{1t}, t=3...365 in a second order autoregressive model.

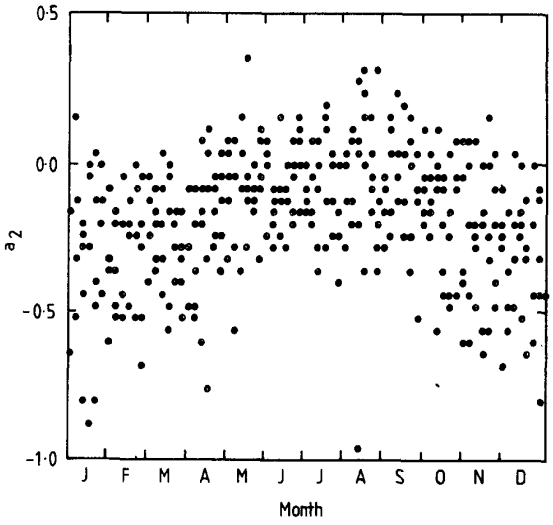


Fig. 5 Estimates of parameters a_{2t}, t=3...365 in a second order autoregressive model.

Table 1 Mean and variance of daily temperatures on wet and dry days, averaged over months.

	Mean		Variance	
Month	Dry	Wet	Dry	Wet
January	1.22	5.11	15.36	10.39
February	2.65	5.53	12.94	9.01
March	6.72	8.29	9.72	7.82
April	12.32	12.32	5.10	5.22
May	16.80	16.01	3.59	2.78
June	19.87	19.54	2.65	2.38
July	24.49	22.84	3.84	3.32
August	25.88	24.81	2.64	3.45
September	21.61	21.81	2.86	3.35
October	16.38	17.34	5.01	3.69
November	10.27	12.46	10.87	8.05
December	4.14	7.40	15.41	10.86

The form of the distribution of temperature on a particular day must be specified to complete a model. Normal distributions are commonly assumed, but a range of other distributions have been used to describe the skewness. (Dumont and Boyce, 1974, Hansen and Driscoll 1977, Furman 1978). The values of Fisher's g_1 statistic (Bliss, 1967) are given in Table 2. It should be noted that the skewness is also non-stationary.

Table 2 Skewness coefficients, with estimated standard errors = 0.05, of daily temperature distributions, pooled over months, for Busan, Korea.

Month	Skewness
January	-0.053
February	-0.156
March	-0.043
April	0.005
May	0.098
June	-0.184
July	-0.312
August	-0.583
September	0.004
October	-0.529
November	-0.499
December	-0.044

3. MODELLING

Most aspects of temperature of interest to applied workers may be investigated by simple distribution-free methods, without using any model. The mean or median of monthly mean temperatures for July, for example, is simply obtained by calculating the mean for each of the 73 years of the record, and then summarising the 73 values. The probability of a day with temperature less than 15°C in March is estimated as the proportion of the 73 years that have a temperature less than 15°C in March. It is these simple summary statistics that are usually used in agricultural studies, (Bichambaut et al. 1968, Kowal and Kassam, 1978), the many models described not being applied. What, then, are the reasons for fitting models? Three are considered.

First, the model may help an understanding of the processes causing variation. Furman (1978), for example, concludes that variations about the mean are determined by persistent weather patterns affecting a large area, rather than by local effects, these conclusions being reached after fitting models to data from eight sites. The form of model considered may be dictated by some knowledge of the system. Hansen and Driscoll (1977), for example, modelled hourly temperature values, including terms to represent annual and diurnal variation.

Second, a well-fitting model provides a summary of the data. Typically, a record of daily values for many years is summarised in a few parameter values. This data summary may now be used directly to compare places or times. Furman (1978) demonstrates the effect of altitude and aspect on daily maximum temperatures by comparing the parameter estimates of the models fitted to the eight sites. However, parameters of many models are difficult to interpret individually and may have little relevance to the aspects of temperature data of interest. A model may still be a useful summary if the items of interest can be calculated easily from it. This can result in a considerable increase in computational efficiency as the model parameters, rather than the original data, have to be accessed for each new item considered. This use of a fitted model is illustrated in Section 4.

The third reason a model-based analysis may be an advantage is the increased efficiency in the use made of the data. The simple non-parametric analysis generally requires a long record of data to achieve estimates of acceptable precision. For many sites data are only available for a few years, between 5 and 10 years being typical. The increase in efficiency in using a model depends, of course, on the model describing that aspect of the data well. The gain in efficiency is often difficult to quantify but a simple example is given in Section 6.

In this paper the aim of modelling is to estimate some aspects of temperature data of interest in agriculture. The simplest of these must be the mean over some period. Monthly means are used to determine agro-ecological regions and assess the suitability of an area for a given crop (Kowal and Kassam, 1978). Degree-day sums are a second aspect of interest. They have been considered by several authors in relation to the rate of development of crops (Smith et al., 1982). These are defined as the sums of temperatures above a lower limit over a period of days. The distribution of these series may then be related to the distribution of plant development rates. A third aspect, when considering the timing of planting or other agricultural operations, is the occurrence of 'hot' or 'cold' days during a period, where these are defined as temperatures above or below upper and lower limits (Brichambaut et al., 1968).

4. A SIMPLE MODEL

Section 2 shows that a comprehensive model of daily temperature data, describing all the various sources of variation noted, is bound to be complex. In this section it is demonstrated that if the model is required to estimate some items of agricultural interest then it may be very much simplified.

The simplest model considered was one which modelled the time dependence of the mean and variance of the daily temperatures but ignored the serial correlations, i.e.,

$$y_{it} = f_1(t) + \epsilon_{it}$$

with $\epsilon_{it} \sim N(0, \sigma_t^2)$, $\sigma_t^2 = f_2(t)$ where $f_k(t)$, $k=1,2$, are Fourier series with n_k harmonics.

The fitting of the model is made very easy with the package GLIM. The sample variances, s_t^2, have distributions which are multiples of a χ^2 distribution. Maximum likelihood estimates of the coefficients of $f_2(t)$ are given by specifying a gamma error distribution and scale parameter of 2 (Baker and Nelder, 1978). If the fitted values are $\hat{\sigma}_t^2$ then coefficients of $f_1(t)$ are estimated

by a weighted least squares fit to $\bar{y}_t$, with weights $\hat{\sigma}_t^{-2}$. Care must be taken in the criteria used to assess the number of harmonics, n_k, required as some of the assumptions used to derive properties of standard tests, such as independence of observations, are known to be violated. In this case n_1 and n_2 were both taken as 5. Thus the fitted model has 22 parameters estimated.

Is this simple model any use? Estimates of monthly mean temperature are calculated from the fitted values, as

$$M = (\Sigma\hat{\mu}_t)/m \quad ,$$

where $\hat{\mu}_t$ is the fitted value of μ_t, the summation being over the m days of the month. The values in Table 3 show that the model estimates monthly means extremely well. Means for periods of any other length can obviously be obtained equally easily.

Table 3 Monthly mean temperatures estimated from the simple model and directly from the data.

Month	Model	Direct
January	2.05	1.99
February	3.29	3.24
March	7.26	7.22
April	12.25	12.36
May	16.59	16.56
June	19.65	19.76
July	23.69	23.74
August	25.53	25.46
September	21.64	21.64
October	16.63	16.56
November	10.76	10.72
December	4.72	4.74

Another item of interest is the probability of a hot day during a given period, where 'hot' is defined as a temperature above an upper limit υ. These probabilities are again easily calculated from the fitted model and are compared with those calculated directly from the data in Table 4. The model predicts the distribution well towards the tails but is not so good around the centre. This discrepancy is probably due to ignoring the correlations rather than inadequacy of the normal distribution.

Table 4 Probability of at least one day with temperature above υ in the periods specified estimated from the simple model and directly from the data.

Period (days)	υ=15°C		υ=20°C	
	Model	Direct	Model	Direct
61 - 75	0.06	0	0	0
76 - 90	0.19	0.16	0	0
91 - 105	0.61	0.47	0	0
106 - 120	0.99	0.96	0.02	0.03
121 - 135	1	1	0.19	0.18
136 - 150	1	1	0.70	0.41
151 - 165	1	1	0.98	0.93

Degree-day sum means are easily estimated when the probability of temperatures below the lower limit is very small, in which case simple sums of the fitted values are used. Clearly these must be estimated as well as mean temperatures. If the probability of temperatures below the lower limit, ℓ, is not negligible then the contribution to the degree-day sum on day t, S_t, is

$$S_t = 0 \quad \text{if } Y_t < \ell$$

$$= Y_t - \ell \text{ if } Y_t > \ell.$$

S_t therefore has a mixed distribution, with the continuous part being a truncated normal. The mean of this (Johnson and Kotz, 1970) is used to give the expected value of S_t,

$$E(S_t) = (1 - p_t)\left\{\mu_t - \ell + \frac{z_t}{1 - p_t}\,\sigma_t\right\}$$

where $p_t = Pr(Y_t < \ell) = \Phi[(\ell - \mu_t)/\sigma_t]$

and $z_t = (1/\sqrt{2\pi})\exp(-(\ell - \mu_t)^2/2\sigma_t^2)$

The mean of a degree-day sum is therefore found by summing the values of $E(S_t)$ predicted by the model. The values in Table 5 are calculated for periods when p_t is not negligible and show that the model tends to slightly over estimate the values compared to the direct summary of the data. However, this bias is not large compared to the standard errors. This suggests that, for this application at least, the normal distribution is adequate despite the evidence in Section 2 that it is not correct.

Table 5 Degree day sum means calculated for lower limit ℓ from the simple model and directly from the data.

Period (days)	ℓ (°C)	Model	Direct	
			mean	s.e
1 - 30	5	17.4	17.3	1.68
31 - 60	5	24.7	23.6	1.99
61 - 90	5	85.6	80.0	3.16
1 - 60	5	42.0	41.5	2.53
121 - 150	10	196.2	182.5	3.19
91 - 150	10	251.4	255.0	4.13
61 - 150	10	285.9	263.8	4.67
121 - 180	10	482.3	467.7	4.52

Table 6 Variance of monthly mean temperatures estimated from the simple model and directly from the data.

Month	Model	Direct
January	0.55	3.09
February	0.50	2.10
March	0.32	1.30
April	0.18	0.47
May	0.11	0.63
June	0.10	0.49
July	0.12	1.42
August	0.12	1.04
September	0.11	0.61
October	0.16	0.79
November	0.37	1.72
December	0.54	3.18

The mean temperatures used in many agricultural studies are seen to be predicted surprisingly well by this very simple model. It is, however, important to remember that this model is a gross simplification, ignoring

correlations between successive days. One consequence of this is that the variances predicted by the model for periods of more than one day will be too small. Table 6 illustrates this with monthly values.

5. A MORE REALISTIC MODEL

In an attempt to provide better estimates of variances a model was fitted that included autocorrelations. Fitting a model with constant first order autoregressive parameter improved the estimates of monthly variances but these were still too low for most of the year. A model with non-stationary first order autoregressive parameter was then tried,

$$y_{it} = f_1(t) + \epsilon_{it}$$

$$\epsilon_{it} = \rho_t \, \epsilon_{it-1} + \gamma_{it}$$

with $\gamma_{it} \sim N(0, \sigma_t^2)$, independent, $\sigma_t^2 = f_2(t)$

and $\rho_t = f_3(t)$.

The functions $f_1(t)$ and $f_2(t)$ were again taken as Fourier series while $f_3(t)$ was a transformation of a Fourier series.

The fitting of this model has to be done in three stages

(i) Estimate parameters of $f_3(t)$, by fitting a Fourier series to $0.5\log[(1 - r_{it})/(1 + r_{it})]$ by least squares, giving smoothed values $\hat{\rho}_t$.

(ii) Calculate generalised differences

$$y'_{it} = y_{it} - \hat{\rho}_t \, y_{it-1}$$

together with their variances

$$s_t'^2 = \Sigma(y'_{it} - \bar{y}'_t)^2/72.$$

The $s_t'^2$ again have $\sigma_t^2 . \chi^2$ distributions so $f_2(t)$ is fitted as in Section 4.

(iii) The parameters of $f_1(t)$ are then estimated by a weighted least squares fit to $\bar{y}'_t$ after having taken generalised differences of the Fourier coefficients.

Calculating monthly means from this fitted model is again straightforward. However calculation of variances is more complex. Approximate variances are easily found by assuming $f_2(t)$ constant within each month. The monthly values in Table 7 show the means are still predicted well. There is an improvement in variances over the simple model, but the model now over estimates variances during the winter months. It should be noted that this is just the part of the year when the second order autoregressive parameter (Fig.5) is clearly non-zero.

Table 7 Monthly means and variances estimated from the non-stationary first order autoregressive model and directly from the data.

	Mean		Variance	
Month	Model	Direct	Model	Direct
January	1.85	1.99	4.22	3.09
February	3.39	3.24	3.84	2.10
March	7.51	7.22	1.74	1.30
April	12.97	12.36	0.74	0.47
May	16.22	16.56	0.44	0.63
June	20.09	19.76	0.48	0.49
July	22.59	23.74	1.25	1.42
August	25.87	25.46	1.15	1.04
September	21.61	21.64	0.52	0.61
October	16.42	16.56	0.93	0.79
November	10.55	10.72	2.18	1.72
December	4.70	4.74	3.86	3.18

6. CONCLUDING COMMENTS

Statistical modelling of any climatic variable should be with an end in mind other than producing a 'good fit', since this may determine the type of model fitted. Any model will be a simplified representation of the physical process but the example in this paper shows that this simplification can be extreme and yet the model still be useful. It is however important to understand the features of the data that are being modelled well and those that are being ignored.

If the aim of modelling is to provide good estimates of some well specified aspects of the data, the adequacy of the model may be assessed by determining how well it represents these particular aspects. Models are also developed as part of larger climatic models to be used in, for example, crop simulation studies (Dumont and Boyce,1974, Jones et al.,1970). In this case it is not obvious exactly which aspects of temperature are going to be important so the models have to describe as much of the variation noted in Section 2 as possible. The resulting models may not be the most useful for the first type of analysis. An example is the paper by Taylor (1972) describing the daily temperatures in Britain. This uses a very ingeneous device to model the annual variability in skewness and variance but results in a model in which the parameters are very difficult to interpret and which cannot be used analytically.

The distribution free methods of analysis will always give estimates of high precision if there are enough years of data available. The 73 years used in these examples will be sufficient for the simple summary statistics considered. However, it is important to note the increased precision of estimates obtained via the model. This can be quantified for monthly means by calculating standard errors of the estimates derived from the simple model (Section 4) and directly from the data (Table 8). In most months the distribution free method gives an estimate with standard error between two and three times that from the model. This means that, to obtain estimates of similar precision, between four and nine times the length of record is required if the distribution free method is to be used rather than the model. If

this is generally true for model based analyses, it demonstrates very clearly the importance of finding useful models so that maximum use may be made of the short records of daily temperature data available in the majority of countries.

Table 8 Estimated standard errors of monthly means estimated from the simple model and directly from the data.

Month	Model	Direct
January	0.0728	0.206
February	0.0676	0.170
March	0.0559	0.133
April	0.0428	0.080
May	0.0342	0.093
June	0.0324	0.082
July	0.0354	0.140
August	0.0350	0.119
September	0.0336	0.091
October	0.0414	0.104
November	0.0592	0.153
December	0.0762	0.209

The computational efficiency of a model based analysis must also be stressed. The example in this paper shows that a simple model involving just 22 parameters can be useful for estimating degree-day sum means. If such sums are to be estimated for a range of periods and lower limits, the distribution free analysis requires repeatedly accessing and recoding of the raw data. The model based analysis needs the model to be fitted just once and then only the 22 parameters used to calculate the mean for any period or lower limit. This computational efficiency is particularly important with the increased use of micro-computers in, for example, smaller agricultural research stations.

Future work in this area must be to find a class of models that can be used to estimate a wider range of aspects of temperature data of interest in agriculture in a range of climates. The same approach to model fitting should also be applied to a bivariate analysis of rainfall and temperature as it is in relation to rainfall that temperature is often of most importance in agriculture.

7. REFERENCES

Baker, R.J. and Nelder J.A., 1978: The GLIM System, Release 3. Numerical Algoriths Group, Oxford.

Bliss, C.I., 1967: Statistics in Biology, Vol. 1. McGraw-Hill, New York.

Brichambaut, G.P. and Wallen C.C., 1968: A Study of Agroclimatology in Semi-Arid and Arid Zones of the Near East. WMO Tech. Note No.56, WMO, Geneva.

Dumont, A.G. and Boyce, D.S., 1974: The Probabalistic Simulation of Weather Variables. J. Agric. Eng. Res., 19, 131-145.

Furman, R.W., 1978: Time Series Analysis of Daily Maximum Temperatures on a Mountain Slope. J. Appl. Meteor., 17, 110-115.

Hansen, J.E. and Driscoll, D.M., 1977: A Mathematical Model for the Generation of Hourly Temperatures. J. Appl. Meteor., 16, 935-948.

Johnson, N.L. and Kotz, S., 1970: Continuous Univariate Distributions - 1. Houghton Mifflin, Boston.

Jones, J.W., Colwick, R.F. and Threadgill, E.D., 1970: A Simulated Environment Model of Temperature, Evaporation, Rainfall and Soil Moisture. ASAE Papers, 70-404.

Kowal, J.M. and Kassam A.H., 1978: Agricultural Ecology of Savanna. Clarendon, Oxford.

Nicks, A.D. and Harp, J.F., 1980: Stochastic Generation of Temperature and Radiation Data. J. Hydrology, 48, 1-17.

Smith, P.J., Bootsma, A. and Gates, A.D., 1982: Heat Units in Relation to Corn Maturing in the Atlantic Region of Canada. Agric. Meteor., 26, 201-213.

Taylor, C.J., 1972: A Stochastic Model of Temperature Variations at Weather Staions in Britain. Applied Statistics, 21, 248-260.

CALIBRATION IN PALEOCLIMATOLOGY WHEN THE BIOLOGICAL RESPONSE IS BINARY

S.R. Esterby and L.D. Delorme

Aquatic Ecology Division
National Water Research Institute
Canada Centre for Inland Waters
Burlington, Ontario, Canada L7R 4A6

1. INTRODUCTION

The features common to attempts to construct empirical calibration models in paleolimnology and paleoclimatology are 1) development of a calibration model from modern data consisting of a biological response and corresponding observations on environmental variables, and 2) use of the calibration model with fossil data to predict unobservable environmental conditions. Webb and Clark (1977) cite examples where the data include tree rings, pollen, foraminifera, diatoms, coccoliths and Radiolaria. The calibration models are usually based on quantitative biological data.

The present paper reports preliminary results on the construction of an empirical calibration model relating ostracode occurrence to environmental variables. Delorme has sampled 6,720 stations across Canada, identifying the ostracode species present and recording physical and chemical characteristics of the station. In addition, meteorological data has been obtained from the Canadian Atmospheric Environment Service. The present study differs from many others in that the organisms present in modern samples were not counted but only identified, and thus the biological response at each station is binary, that is, a species was either present or absent. In this case, the procedures of normal-theory linear models are not directly applicable. An alternative model, for which practical procedures are available, is the linear logistic model. If sampling sites can be grouped, the frequency of occurrence of a species for the stations in a group replaces the binary observations at the stations. This permits the use of the empirical logistic transform and the procedures of normal-theory linear models are available. The crucial assumption is that a suitable grouping of the stations can be found.

In the subsequent sections, the model is specified for our calibration problem; the methods used to reduce the number of variables and group the stations, for the aquatic habitat variables, are described; and some of the regression results are given. In the last section, alternative methods are suggested, and the use of the regression model with fossil data and the extension of the methods are discussed.

2. CALIBRATION MODEL

Let Y_i be the binary biological response variable, then $Y_i=1$ if the species is present at station i and $Y_i=0$ otherwise. Using the linear logistic model (Cox, 1970), the relationship between the probability that a species is present at station i and the values of the environmental variables at the station are given by

$$\lambda_i = \log\left(\frac{\theta_i}{1-\theta_i}\right) = \sum_{\ell=1}^{p} x_{i\ell}\,\beta_\ell \qquad \text{for } i=1,2,\ldots,n \quad (1)$$

where $\theta_i = P(Y_i=1) = E(Y_i)$, x_{i_1}, $x_{i_2},\ldots,x_{ip}$ are the values of the p environmental variables at station i, and n is the number of sample stations. If the stations can be grouped so that in any group the probability of a species being present can be assumed to be constant, then for k groups, the model becomes

$$\lambda_j = \log\left(\frac{\theta_j}{1-\theta_j}\right) = \sum_{\ell=1}^{p} x_{j\ell}\,\beta_\ell \qquad \text{for } j=1,2,\ldots,k \quad (2)$$

where now x_{j_1}, $x_{j_2},\ldots,$ x_{jp} are values of the environmental variables for the groups. Let n_j be the number of stations in the j^{th} group and R_j be the number of stations in the j^{th} group at which the species is present. Use of the empirical logistic transform, in a form suitable for weighted regression analysis, leads to fitting the model

$$E\left(\frac{Z_j}{\sqrt{V_j}}\right) = \sum_{\ell=1}^{p} \frac{x_{j\ell}}{\sqrt{V_j}}\,\beta_\ell \qquad \text{for } j=1,2,\ldots,k \quad (3)$$

by normal theory regression methods, where

$$Z_j = \log\left(\frac{R_j-\frac{1}{2}}{n_j-R_j-\frac{1}{2}}\right) \text{ and } V_j = \frac{(n_j-1)}{R_j(n_j-R_j)} \qquad (4)$$

V_j is the estimate of the variance of the transform Z_j and is used as the weighting factor.

3. VARIABLE REDUCTION AND GROUPING

For the problem as described above, only the formation of groups satisfying the assumption that $P(Y_{i=1})$ for all stations in the group is required. However, for the present set of data it was necessary to reduce the number of variables as well because, after grouping, the number of variables is relatively large and because strong relationships exist between some of the environmental variables. Principal components analysis using the correlation matrix of the 19 quantitative habitat variables allowed a considerable reduction in dimension. The 10 largest eigenvalues accounted for 91.4 percent of the total sum of squares, and, as can be seen from Table 1, as few as five components account for many of the important features and 69.6 percent of the variation.

Table 1. Principal component analysis using the correlation matrix of the 19 quantitative habitat variables for 932 Saskatchewan stations.

Component	Important Variables	Percentage of Variation	
		Component	Cumulative
1	Major ions	29.7	29.7
2	Temperature	16.2	45.9
3	pH, alkalinity, O_2, Fe, PO_4	12.5	58.4
4	Water depth, distance from shore	6.8	65.2
7	Distance from shore, O_2	4.4	69.6
5		6.6	76.2
6		5.0	81.2
8		3.8	85.0
9		3.5	88.5
10		2.9	91.4

As well as the 19 quantitative habitat variables, five unordered categorical variables describing the aquatic habitat and land vegetation were recorded. These unordered categorical variables were used as a five-way classification to define groups of stations. The present analysis was restricted to 1574 stations in the province of Saskatchewan, and, the setting of a minimum number of stations per group lead finally to 16 groups, hereafter called regions, with distinct combinations of the five categorical variables and a final total of 932 stations, after removal for missing data (examples of regions are given in Table 2).

To further increase the homogeneity of the groups of stations, cluster analysis within region was performed using a k-means, non hierarchical, Euclidean-distance clustering algorithm (Anderberg, 1973) on the first 10 principal component score vectors. Most of the regression results reported in the next section are for region eight. A multivariate analysis of variance (Cooley and Lohnes, 1971, MANOVA program) and plots of cluster means separately for each variable showed that the seven groups within region eight have different characteristics with respect to the quantitative aquatic habitat variables.

Table 2. Examples of the regions defined by the five unordered categorical variables.

Region	Source of Water, State of Permanence	Land Vegetation	Soft or Hard Substrate, Bottom Vegetation
1	Littoral, permanent	Aspen parkland	Hard, no vegetation
2	Pond, permanent	Low grassland	Soft, vegetation
3	Pond, permanent	Aspen parkland	Soft, vegetation
4	Pond, temporary	Low grassland	Soft, vegetation
6	Pond, temporary	Medium grassland	Soft, vegetation
8	Pond, temporary	Aspen parkland	Soft, vegetation

4. REGRESSION ANALYSIS

The regression model, given by equation (3), if written for all groups involves terms for the quantitative variables and terms for the regions which were derived from the five unordered categorical variables. The latter are included by the use of "dummy variables" (Draper and Smith, 1981).

Results from a small subset of the data are shown here. Regions six and eight were analyzed for two species: 1) Cypridopsis vidua which is abundant, being present in 60 percent of the stations over all 16 regions; and 2) Candona renoensis which is present in only 17 percent of the stations. For both species, the regression model was fitted using a term accounting for the region to which the group belongs and terms for the vectors of scores on the principal components. It was found, in both cases, that the proportion of stations at which the species was present did not differ significantly for regions six and eight. It can be seen from Table 3 that vectors of scores on the first three principal components provide a reasonable fit for Cypridoposis vidua, and explain 76 percent of the variation, while they explain only 65 percent of the variation for Candona renoensis.

Table 3. Regression analysis for two species versus the score vectors for the first three principal components and regions six and eight.

Species	Region	Group	Observed N_j	Observed R_j	Estimated $\hat{\theta}_j$	Estimated $\hat{R}_j$
Cypridopsis vidua	8	1	12	10	0.79	9.5
		2	58	36	0.65	37.9
		3	12	8	0.62	7.5
		4	34	23	0.58	19.6
		5	52	25	0.49	25.7
		6	18	15	0.85	15.3
		7	28	14	0.43	11.9
	6	1	25	19	0.74	18.5
		2	29	18	0.68	19.6
		3	53	25	0.55	29.0
Candona renoensis	8	1	12	2	0.13	1.5
		2	58	10	0.20	11.5
		3	12	4	0.40	4.8
		4	34	14	0.27	9.3
		5	52	11	0.20	10.6
		6	18	5	0.25	4.5
		7	28	6	0.24	6.7
	6	1	25	4	0.14	3.6
		2	29	10	0.37	10.8
		3	53	11	0.22	11.6

5. DISCUSSION

The environmental variables analyzed to date do not inlcude the climatological variables which were obtained from a separate source, but the methods and the logic of the analysis apply to climatological variables alone or together with the aquatic-habitat and land-vegetation variables. Variables such as mean annual temperature, annual precipitation and potential evapotranspiration are to be included.

The use of the empirical logistic transform, and consequently normal theory regression methods, means that the method of inverse estimation (Draper and Smith, 1981) is available. That is, given a value of Z in equation (3), derived from an observed fossil proportion, values of the environmental variables can be estimated. This becomes a complicated task if many environmental variables are included. An example of inverse estimation in multiple regression, for a different type of problem, is given by El-Shaarawi and Esterby (1981).

The use of least squares to fit the model means that groups, for which $R_j=0$ or $R_j=n_j$, are excluded. For very sparse or very abundant species maximum likelihood estimation may be preferable. Extension to a multivariate analysis is also possible.

The purpose has been to outline the general methods being used to analyze a complicated set of data with the features of 1) a binary response variable and 2) many independent variables, which include both quantitative and unordered categorical variables. The approach should be applicable to other calibration data sets, with modifications likely at the variable reduction and grouping steps. The importance of many sources of variation needs to be determined with the present set of data before a calibration model can be proposed.

6. REFERENCES

Anderberg, M.R., 1973: Cluster Analysis for Applications, Academic Press, New York.

Cooley, W.W. and P.R. Lohnes, 1971: Multivariate Data Analysis, Wiley, New York.

Cox, D.R., 1970: The Analysis of Binary Data, Methuen, London.

Draper, N.R. and H. Smith, 1981: Applied Regression Analysis, Wiley, New York.

El-Shaarawi, A.H. and S.R. Esterby, 1981: Analyse de régression de la variation spatiale d'une charactéristique limnologie. Eau d'u Québec, 14, 222-228.

Webb, T. and D.R. Clark, 1977: Calibrating micropaleontological data in climatic terms: A critical review. Annals of New York Academy of Sciences, 288, 93-118.

AN ANALYSIS OF THE SPACE AND TIME VARIATION OF GROWING SEASON RAINFALL IN THE CENTRAL UNITED STATES

Peter J. Lamb[1] and Michael B. Richman

Illinois State Water Survey
Champaign, Illinois 61820

1. INTRODUCTION

It is well known that summer rainfall is of great importance for crop production in the rich agricultural region of the central United States, and that this rainfall is overwhelmingly convective and hence rather variable in space and time. An understanding of its intraseasonal and interannual fluctuations in terms of the relative importance of several possible contributing climate system processes — for example, the local presence and external supply of water vapor, and the troposphere's predisposition to provide the rain-inducing vertical motion — is therefore desirable, particularly since this knowledge is essential for the development of physically (as opposed to statistically) based climate prediction schemes. We are in the early stages of a research endeavor that seeks such understanding. It seems mandatory that the overall project commences with and is based upon an objective identification of the characteristic spatial patterns (or modes) of rainfall variation in the study region on time-scales ranging from a few days to the entire season (here defined as May-August), and also on knowledge of the time (intraseasonal and interannual) fluctuations of those patterns. Such work should permit the choice of a truly representative range of case study years for the proposed central investigation of the climate system processes that underlie the major modes of rainfall variation. The present contribution provides a preliminary report on the results of the rainfall pattern analysis for the following time-scales — 3, 7, 15, 30, and 61 (∿ half-season) days, and the entire season (123 days). Much more complete treatments of this and related material will be presented elsewhere in due course (Richman and Lamb, 1983; Lamb and Richman, 1983).

2. DATA

Because our interest includes several time-scales of less than the calendar month, for which there existed no readily available historical data sets with the spatial resolution necessary for the treatment of convective rainfall, we have created a new data set for use in the overall project outlined above. It contains the individual daily rainfall amounts for 402 stations in the central United States for the May-August periods of 1949-80, which total more than 1.58 million observations. The station network is displayed in Fig. 1; it is bounded to the west and east by the approximate edges of the Rocky and Appalachian mountain systems, respectively. We plan to extend the network northward into Canada (to 55°N) in due course. Full details of the construction and characteristics of this data set will be presented elsewhere, with a brief review sufficing here. The first step in its development was the acquisition of all daily cooperative substation rainfall totals for the study region and period held in the National Climatic Center's "Summary of Day Elements" file, a data deck that was recently quality controlled. Many stations commenced or ceased operation during the period and/or had gaps in their records. The finest network resolution that was found to be

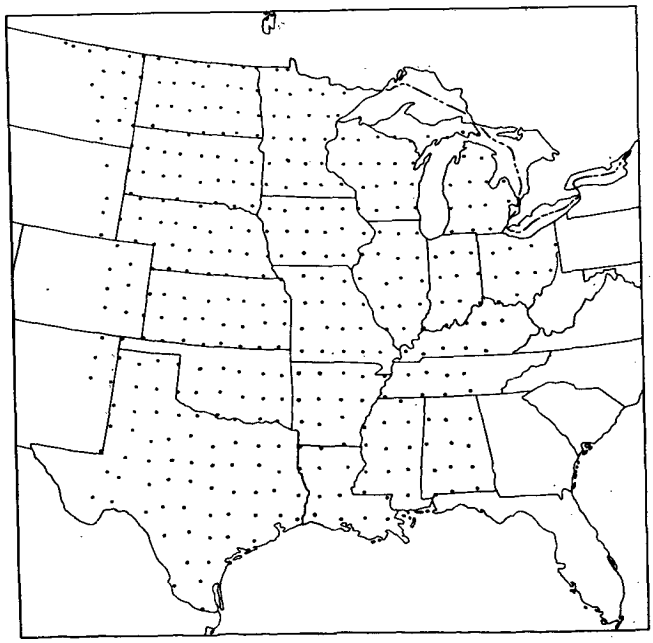


Fig. 1. Station network for central United States daily rainfall data set for May-August periods of 1949-80.

[1]Also affiliated with Department of Atmospheric Sciences, The University of Illinois.

practicable from the standpoints of spatial and temporal uniformity and data set completeness was ∿ 80-100 km. A 1° latitude/longitude "target" grid was therefore established, and the station nearest to each gridpoint with the most complete and seemingly most reliable record was identified. These stations were usually located within 10 km of their gridpoint, and were chosen to form the present data set (Fig. 1). For the days they did not possess observations (generally 2-3% of the total), the value from one of two adjacent "alternate" stations was substituted without adjustment. These alternate stations were usually situated within 15 km of the "primary" site, and were selected because their records, although often incomplete, covered the gaps in those of the primary station.

The present paper utilizes square root transformations of 3-day (i.e., 1-3 May, ..., 29-31 August), 7-day (1-7 May, ..., 21-27 August), 15-day (1-15 May, ..., 14-28 August), and 30-day (1-30 May, ..., 30-July - 28 August) totals computed from the aforementioned daily data, and also untransformed 61-day (1 May - 30 June, 1 July - 30 August), and 123-day (1 May - 31 August) totals. The square-root transformation of the shorter time-scale totals was employed to reduce the well known skewness of their distributions. It was chosen after considerable experimentation with alternative transformations, and had the desirable effect of increasing the variance explained by the Principal Component patterns presented below without altering their morphology. These results were also found to be unaffected by either the slight northward decrease of the west-east station spacing (Fig. 1) or the fact that cooperative substation observation times vary.

3. METHODS

Since our ultimate need included both an objective identification of the characteristic spatial patterns of rainfall variation and information on their time fluctuations, the foregoing data sets for different time-scales were analyzed statistically using eigenvectorial methods. Only the spatial patterns are presented here, however. Recent years have seen meteorological and oceanographic research utilize an increasing range of eigenvectorial techniques, a development that has been exploited in this investigation. In addition to computing the well known and standard unrotated Principal Components (PCs) of the station-versus-station correlation matrix, we have also obtained from this input orthogonally (VARIMAX criterion) and obliquely (Harris-Kaiser Case II criterion) rotated PCs. Experimentation showed that the covariance input matrix yielded near identical results. A secondary objective has been the comparison of the results yielded by the different analytical methods. Since unrotated PCs are basically data compression devices that merely seek to maximize the variance explained over the entire domain, they may not adequately capture the shape of the dominant modes of variability of the input data, particularly if the latter include distinct clusters such as characterize our rainfall data (not shown). In the present context, this translates into a potentially improper delineation of the regions of strongest spatial rainfall coherence; the quest for maximum variance extraction for the entire domain could merge distinctive subareas. The rotation of PCs seeks to avoid this problem by fitting them more closely to the data clusters. It accordingly produces spatial loading patterns that typically contain many near-zero values and a smaller number of larger ones, and hence generally feature somewhat stronger anomalies covering smaller portions of the domain than occurs without rotation (cf. Figs. 2 and 3). Although the VARIMAX rotation used here is an orthogonal one, the PC loadings (i.e., the spatial patterns) are allowed to be spatially correlated. Their time dependent scores, however, remain uncorrelated. Oblique rotations relax the orthogonality constraint entirely.

4. RESULTS

The unrotated and VARIMAX orthogonally rotated PC results for all time-scales considered are summarized in Table 1 and Figs. 2-6. The VARIMAX results were essentially duplicated by those yielded by the Harris-Kaiser oblique rotation. This enhanced the credibility of the former and pointed to the dominant modes of variation of the study phenomenon being nearly orthogonal in time.

Table 1 shows that the first 10 of the 402 total unrotated PCs account for 44% of the variance of 3-day square-root transformed summer rainfall in the central United States. These first 10 unrotated PCs, unlike all subsequent ones, were suggested to contain nonrandom signals by the application of the "scree" test (Cattell, 1966) to their corresponding eigenvalue series. The use of a Monte Carlo testing procedure (e.g., Preisendorfer and Barnett, 1977) would have retained an undesirably large number (∿ 43) of unrotated PCs, as is typical for substantial sample sizes (here N = number of observations = 1312 3-day periods). Table 1 also shows that the eigenvalues for each of the first five unrotated PCs are significantly separated from the immediately following eigenvalue according to the North et al. (1982) criterion.

The spatial loading patterns of the first eight unrotated 3-day PCs are depicted in Fig. 2. Their statistical (but not necessarily physical) reality is strongly supported by the fact that extremely similar patterns were obtained from three subsets of the data. This was also true of unrotated PCs 9 and 10. The data subsets were: (a) all observations for a lower (∿ half) density network of 199 stations, (b) one-half of the observations (those for 1949, 1951, ..., 1979) for that 199 station network, and (c) the remaining half of those observations. The 3-day unrotated PC 1 (11.13% of total domain variance) is completely dominated by an anomaly of one sign that covers virtually the entire study area from an Iowa center. Even within this center, however, the maximum station variance it explains (given by the square of the loading) never exceeds 31%. The next two members of this PC set, in contrast, feature anomalies of opposite sign in the northern/southern (No. 2, 9.33%) and western/eastern (No. 3, 5.55%) halves of the study region. While unrotated PC 4 (4.54%) contains a west-east oriented anomaly through the center of the domain that has the opposite sign to similarly aligned anomalies in the extreme north and south, and number 5 (3.51%) has a col-like form, numbers 6-8 exhibit little structure and would seem to be especially difficult

Table 1. Eigenanalysis results derived from station-versus-station correlation matrices of central United States summer rainfall for several time scales. The station network included 402 stations for all time-scales except 61-days (199 stations, see text). The numbering of the components/regions in Figs. 2-6 is consistent with the root numbers given here. An asterisk after an eigenvalue indicates that it was not significantly separated from the immediately following one according to the sampling error criterion of North et al. (1982).

	Time-Scale											
	3-day			7-day		15-day		30-day		61-day		123-day	
	1312 Observations			544 Observations		256 Observations		128 Observations		64 Observations		32 Observations	
Root No.	Unrotated Eigenvalue	Unrotated Percent Total Variance Explained	VARIMAX Percent Total Variance Explained	Unrotated Eigenvalue	VARIMAX Percent Total Variance Explained	Unrotated Eigenvalue	VARIMAX Percent Total Variance Explained	Unrotated Eigenvalue	VARIMAX Percent Total Variance Explained	Unrotated Eigenvalue	VARIMAX Percent Total Variance Explained	Unrotated Eigenvalue	VARIMAX Percent Total Variance Explained
1	44.76	11.13	5.11	49.58	5.64	51.42	6.75	54.73	7.24	30.99	8.93	57.57*	11.79
2	37.49	9.33	4.79	41.84	5.55	43.98	5.65	46.37	6.94	17.73	7.79	51.35	8.09
3	22.30	5.55	4.70	23.65	5.20	21.94	5.64	21.20*	6.00	12.60*	7.02	28.58*	7.55
4	18.25	4.54	4.62	18.24	5.14	18.21	5.57	19.19	5.93	11.88	6.70	23.91	7.24
5	14.09	3.51	4.56	14.56	4.93	12.96	5.27	15.25	5.92	8.00*	4.90	17.94*	5.48
6	9.84*	2.45	4.55	10.05	4.85	10.81	4.61	11.98	4.68	7.13*	4.86	16.25*	5.26
7	9.73	2.42	4.26	8.91	4.56	9.09*	4.18	10.43	4.32	6.07*	3.85	15.06*	5.12
8	7.43	1.85	4.08	7.72	4.05	8.70	4.13	8.94*	4.16	5.39*	3.75	14.16*	4.02
9	6.94*	1.73	3.75	7.20	3.65	7.10*	3.72	8.25*	2.91	4.82*	3.73	12.77*	3.88
10	6.60	1.64	3.72	6.62	3.29	6.65	1.96	7.56	2.62	4.60	3.34	11.46	3.74
Sum 1-10		44.14	44.14		46.86		47.48		50.72		54.87		62.17

to interpret physically. Furthermore, the maximum station variance explained by unrotated PCs 3-8 seldom exceeds 15%. It is also interesting to note that the spatial loading patterns of the first five unrotated PCs show a remarkable similarity to the common set Buell (1975, Figs. 5 and 9) extracted from different correlation functions when the domain was rectangular as in the present case (Fig. 1). The Buell patterns appear in the lower left insets of the panels of Fig. 2.

The 3-day VARIMAX rotated PC results appear in Table 1 and Fig. 3. Table 1 shows that the rotated versions of the first 10 PCs account for the same entire domain cumulative variance fraction as the unrotated originals. Since rotation is applied to only those unrotated PCs that are considered to contain nonrandom signals, it seeks to redistribute their cumulative explained variance as evenly as possible among the same number of rotated PCs (as Table 1 illustrates), as was partially implied above. The aforementioned North et al. (1982) significance criterion is therefore not applicable once PCs have been rotated. More importantly, however, the statistical reality of our rotated 3-day PC results is strongly supported by the fact that their spatial loading patterns (Fig. 3) were all completely replicated by the analysis of each of the three data subsets listed previously. Fig. 3 gives these patterns for the first 8 rotated PCs. They are consistent with the earlier statement that rotated PCs feature relatively small but strong anomaly regions. In this case such regions tend to have a spatial scale equivalent to several medium-sized states, within which the explained station variance generally exceeds 15% and always rises to central values of 40-50%. Furthermore, the patterns do not degenerate and lose their structure with increasing root number; number 8 is just as distinctive as number 1. Interestingly, no rotated PC is centered on Iowa. These results are thus in considerable contrast to the unrotated patterns in Fig. 2.

Fig. 4a subdivides the study area into regions within which 3-day summer rainfall tends to be spatially coherent, on the basis of the VARIMAX rotated PC results in Fig. 3. This regionalization is considered to be more complete, realistic, and interpretable, in a physical sense, than any subdivision that could be derived from Fig. 2's unrotated results. Some of its apparently more meaningful aspects include Region 2's approximation to much of the "Corn Belt", the linking of northern Illinois with Wisconsin rather than the rest of Illinois, the inclusion of northeastern Iowa, southeastern Minnesota, and evidently southern Ontario into what is a "Great Lakes" region (number 4), the separation of northwest Texas and western Oklahoma/Kansas from the rest of those states, and the suggestion that North Dakota and eastern Montana are more linked with the Canadian prairies than South Dakota.

Figs. 4b-d provide counterparts to Fig. 4a for the 7-, 15-, and 30-day time-scales. They were obtained in exactly the same manner as the

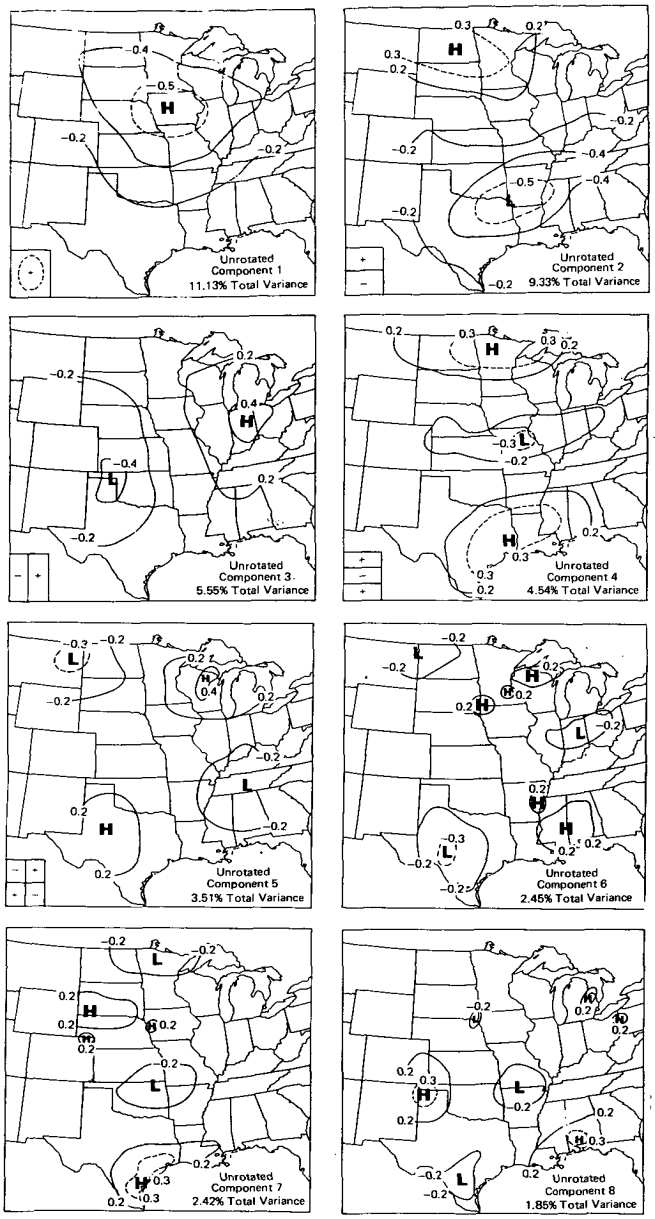


Fig. 2. Unrotated Principal Components 1-8 of a correlation matrix derived from square-root transformed 3-day summer rainfall. Buell's patterns (see text) are displayed in the lower left insets of Numbers 1-5.

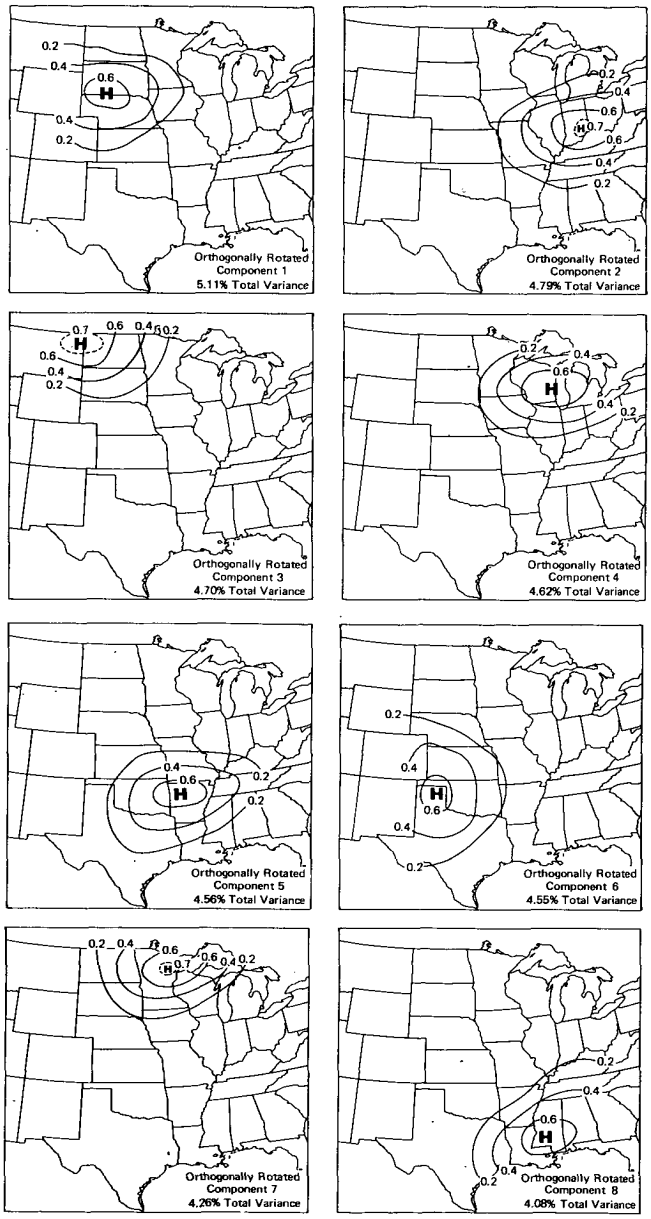


Fig. 3. VARIMAX orthogonally rotated Principal Components 1-8 of a correlation matrix derived from square-root transformed 3-day summer rainfall.

3-day regionalization. It is readily apparent that the areas within which 7-day summer rainfall tends to be spatially coherent are essentially identical to those discussed above for 3-day periods (Figs. 4a, b). This similarity is also very characteristic of the 15-day pattern (Fig. 4c), with the only real difference there being the virtual elimination of the central 3/7-day region that included southeastern Nebraska, northeastern Kansas, southern Iowa, and northern Missouri (Figs. 4a, b). The latter primarily resulted from the southward and eastward expansion of the region that closely coincided with Nebraska and Kansas in the 3- and 7-day analyses. Despite the relative irregularity of its regional boundaries, the 30-day pattern is essentially the same as the 15-day one (cf. Figs. 4c, d).

The regions within which summer rainfall in the central United States tends to be spatially coherent thus exhibit surprisingly little change as the time-scale is increased from 3 to 30 days. Furthermore, Table 1 shows that the cumulative variance fraction explained by the 10 VARIMAX rotated PCs on which these patterns are based (and of course their unrotated forerunners) increases only slightly (from 44% to 51%) with this same time-scale increase. The spatial loading patterns of the first 10 unrotated PCs for the 7-, 15-, and 30-day rainfall data (not shown) are

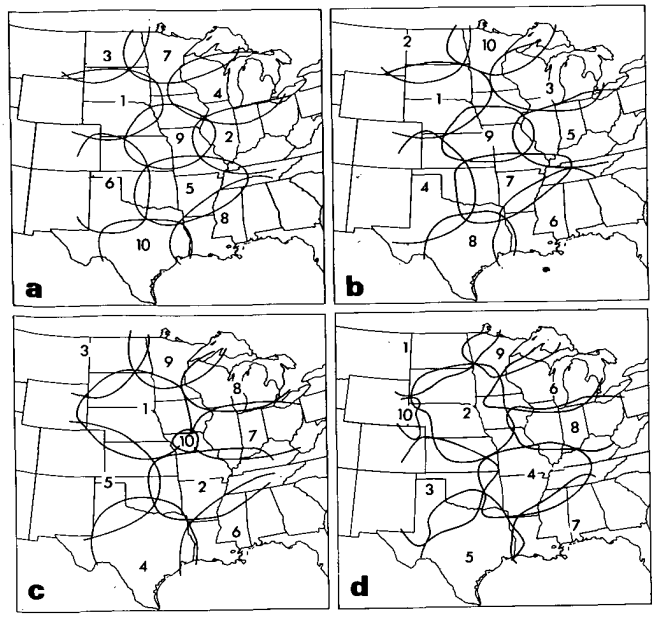


Fig. 4. Regionalization of central United States for short-period (square-root transformed) summer rainfall on the basis of the patterns for the first 10 VARIMAX orthogonally rotated Principal Components: (a) 3-day, (b) 7-day, (c) 15-day, (d) 30-day. The regional boundaries are the +0.4 loading isopleths for each Principal Component.

extremely similar to those presented in Fig. 2 for 3 days. It should also be noted that the eigenvalues for the unrotated 7-, 15-, and 30-day PCs are almost all significantly separated from the immediately following eigenvalue for the same time-scale according to the North et al. (1982) criterion (Table 1). In particular, all 7-day eigenvalues pass this test.

The VARIMAX rotated PC patterns extracted from the half-season (61 days) and full-season (123 days) rainfall totals (Figs. 5 and 6) are in considerable contrast to those presented in Figs. 3 and 4 for the shorter time-scales. While the 61-day results do yield a regionalization of the type appearing in Fig. 4 (not directly shown but implicit in Fig. 5), it differs somewhat from that obtained for 30 days. The greatest similarity occurs in the Upper Midwest, for which the shorter time-scale regions centered on (a) North Dakota, (b) northern Minnesota, and (c) Wisconsin-Michigan remain well defined at 61 days. Further south , however, the regional boundaries are frequently different from those in Fig. 4. An exception is the moderate preservation of the eastern Gulf Coast region. The 123-day VARIMAX results (Fig. 6), in pronounced contrast to those already discussed (Figs. 3-5), do not yield an adequate regionalization of the study domain. The patterns for roots 7-10 are particularly ill-defined. This occurs despite the cumulative variance fraction accounted for by the first 10 PCs being as high as 62%. Whereas the "success" of pattern analysis is often considered to be directly proportional to the explained variance fraction,

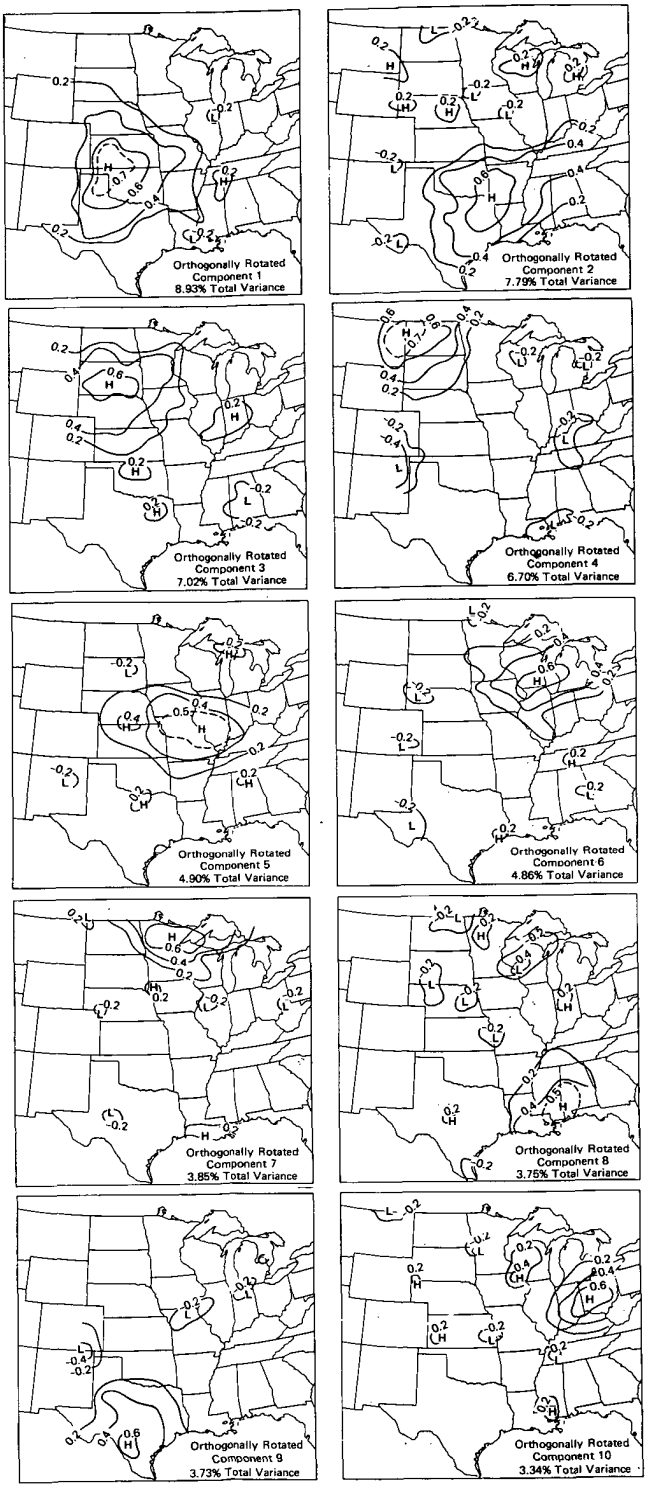


Fig. 5. Same as Fig. 3 for first 10 VARIMAX Principal Components of untransformed half-season (61-day) rainfall totals.

in this case an increase in that fraction seems to have been accompanied by a degradation in the meteorological meaning and utility of the results. It should also be added that, because of the very small sample (32 observations) of 123-day rainfall totals available, the eigenvalues for the unrotated

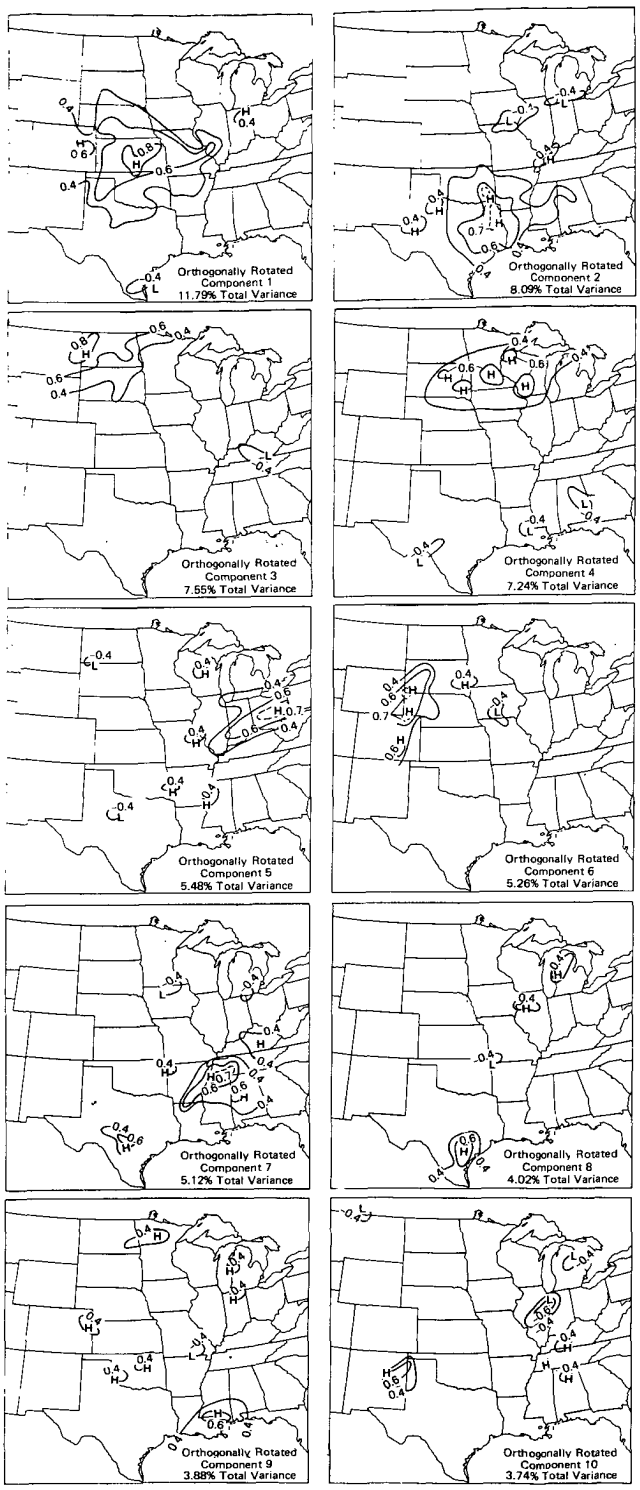


Fig. 6. Same as Fig. 5 for untransformed seasonal (123-day) rainfall totals.

PCs for this time-scale are not significantly separated from each other (Table 1). Finally, even at this long end of the time-scale continuum, the unrotated PC spatial loading patterns (not shown) still possess considerable resemblance to the Buell-like ones shown in Fig. 2 for 3-day periods.

5. CONCLUDING REMARKS

The foregoing results have thus suggested the shorter time-scale VARIMAX rotated PC analyses to have particular utility for our overall project along the lines of the need stated at the outset. The 7-day regionalization is now being particularly exploited in this regard because (a) it is derived from a larger population and possesses finer resolution score time-series than those for 15 and 30 days, and (b) its score time-series are more compact and stable than the 3-day ones. This approach is in pronounced contrast to most previous empirical research into large-scale climatic (including rainfall) variation. The latter has tended to utilize the monthly and seasonal data that is relatively easy to access and compact to process and analyze, and where eigenvectorial analysis has been involved it has generally been of the orthodox unrotated variety.

6. ACKNOWLEDGMENTS

This study was supported by NSF Grant ATM 81-00158. We thank Diane Portis and Keith Hendrie for their computational assistance, and Don Gilman, Bob Livezey, John Walsh, and Brian Weare for helpful comments at an early stage of the work.

7. REFERENCES

Buell, C. E., 1975: The topography of empirical orthogonal functions. Preprints Fourth Conf. on Probability and Statistics in Atmos. Sci., Tallahassee, Florida, Amer. Meteor. Soc., 188-193.

Cattell, R. B., 1966: The scree test for the number of factors. Multivar. Behav. Res., 1, 245-276.

Lamb, P. J., and M. B. Richman, 1983: On the modes of variation of growing season rainfall in the central United States. In preparation.

North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng, 1978: Sampling errors in the estimation of empirical orthogonal functions. Mon. Wea. Rev., 110, 699-706.

Preisendorfer, R. W., and T. P. Barnett, 1977: Significance tests for empirical orthogonal functions. Preprints Fifth Conf. on Probability and Statistics in Atmos. Sci., Las Vegas, Nevada, Amer. Meteor. Soc., 169-172.

Richman, M. B., and P. J. Lamb, 1983: On the climatic pattern analysis of short-period (3-7 day) summer rainfall in the central United States: Some methodological considerations and a regionalization. In preparation.

FUZZY MULTIFACTORIAL EVALUATION IN CLIMATOLOGICAL ASSESSMENT

Chen Guofan

Academy of Meteorological Science
Beijing, China

1. INTRODUCTION

In many cases information is found to be naturally fuzzy as when representing personality traits, physical features of individuals, and subjective opinions and judgments in situations such as policy preferences, medical decision making, economic forecasting, and personnel evaluation. Obviously, most of the meteorological phenomena are also of fuzzy state. For example, "hot summer", "lovely weather", "rainy season", and so on . All these no precise formulation and no well defined boundaries. If these problems can be done in a manner that allow the exact representation to be a special case of the fuzzy representation, then, clearly, the fuzzy representation is the more powerful. I

In 1965, L. A. Zadeh suggested the fuzzy set theory to study and deal with the fuzziness of things. He defined fuzzy subsets of a universe of discourse as generalized traditional subsets such that the grades of membership of an element of the universe in a specified fuzzy subset may have any value in 0,1 instead of 0 or 1 only. The membership value $\mu_A(x)$ can be interpreded as the degree of compatibility of the predicate associated with the fuzzy subset A and the object x. The closer the value of $\mu_A(x)$ is to 1, the more x belongs to A; conversely, the closer the value of $\mu_A(x)$ is to 0, the more x not belongs to A. All the classical settheoretic operation can be extended to for the fuzzy sets. This approach provides an effective tool for describing and studying the objects which are in vague.

Since 1979, the problem as how to represent the membership function of meteorological elements and the application of fuzzy recognition to daily circulation patterns have been attacked. Thereon satisfactory objective results were obtained by using fuzzy clustering dealing with partition of circulation processes and classification of synoptic situation. Now in vast numbers of weather service fuzzy set theory is applied to the forecasting of long-range precipitation, typhoon track, and cold wave. In this paper, the basic principle of fuzzy multifactorial evaluation is introduced and an example of climatological assessment is given.

2. THE BASIC PRINCIPLE OF FUZZY MULTIFACTORIAL EVALUATION

A fuzzy set is actually a generalized subset of a classicalset. The classical union($\cup$) and intersection($\cap$) of ordinary subsets of X can be extended by the following formulas,

$$\forall x \in X \quad \mu_{A\cup B}(x) = \max(\mu_A(x), \mu_B(x)),$$

$$\forall x \in X \quad \mu_{A\cap B}(x) = \min(\mu_A(x), \mu_B(x)),$$

where μ_A, μ_B, $\mu_{A\cup B}$, $\mu_{A\cap B}$ are respectively the membership functions of A, B, A∪B, and A∩B.

Let U and V be limited sets,

$$U = \{u_1, u_2, \cdots, u_n\},$$
$$V = \{v_1, v_2, \cdots, v_m\}.$$

A and B are two fuzzy sets on U and V respectively. They could be expressed as fuzzy vectors,

$$A = (a_1, a_2, \cdots, a_n), \quad 0 \le a_i \le 1,$$
$$B = (b_1, b_2, \cdots, b_m), \quad 0 \le b_i \le 1.$$

Let R be a fuzzy relation on U×V,

$$R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{pmatrix}, \quad 0 \le r_{ij} \le 1$$

A fuzzy multifactorial evaluation of a fuzzy relation R and a fuzzy set A is defined by

$$A \circ R = B,$$

where the symbol ∘ represents the composition of fuzzy relations. We say that B is a fuzzy set induced from A through R. The transformation can be viewed as a multiplication of matrix, but where plus is the operation "max" and product is the operation "min". The element of B is

$$B_k = \bigvee_i \left[a_i \wedge r_{ik} \right], \quad \begin{matrix} i=1,2,\ldots,n, \\ k=1,2,\ldots,m. \end{matrix}$$

When we are going to give an evaluation to an objective reality, it is easier to draw the conclusion depend on single factor. But usually the target of evaluation is not affected by only one factor. That is why multifactorial evaluation is necessary.

3. AN EXAMPLE OF CLIMATOLOGICAL ASSESSMENT

According to medical research, the optimum conditions for human is that temperature is 17°C, wind speed 2m/sec, and relative humidity 60%. Based on these three factors, we give the climatological assessment to Qingdao, Dalian, Chengdu, and Erlianhowte in summer half year(Mar. to Aug.) by using fuzzy multifactorial evaluation.

Let U denote the fuzzy set of factors,

$$U = \{\text{temperature, wind speed, humidity}\}.$$

Probably we attach different importance to different factor. Let A be the weight matrix, which shows the importance of corresponding factor,

$$A = (a_1, a_2, a_3),$$

generally $\sum_i a_i = 1$.

The evaluation are supposed to be given out in 5 grade, denoted as

$$V = \{\text{best, better, good, no good, bad}\}.$$

For every single factor we can get a evaluation in either grade. All the evaluation together a matrix R would be formed. This matrix gives the relation on U×V. The composition of R and A is the multifactorial evaluation we expect.

Take Dalian as example, on the basis of historical data, we obtain R as follow:

$$R = \begin{pmatrix} 0.33 & 0.02 & 0.33 & 0.12 & 0.17 \\ 0.00 & 0.10 & 0.58 & 0.32 & 0.00 \\ 0.28 & 0.15 & 0.12 & 0.23 & 0.22 \end{pmatrix}.$$

Assume A=(0.20 0.50 0.30), then

$$B = A \circ R = (0.28 \quad 0.15 \quad 0.50 \quad 0.32 \quad 0.22).$$

It shows that by multifactorial evaluation, Dalian just a good place for people in summer. Similarly, a result shows that Qingdao is the best place in this season.

4. REMARKS

In spite of the simplicity of the algorithm, fuzzy set theory is quite efficient for meteorology. We anticipate it will get more fruitful results in this field.

5. REFERENCE

Cao, H. and G. Chen, 1983: Some applications of fuzzy sets to meteorological forecasting. Fuzzy Sets & Systems, 9, 1-12.

THE ROLE OF STATISTICS IN CLIMATE IMPACT ANALYSIS

T.M.L. Wigley

Climatic Research Unit,
University of East Anglia,
Norwich NR4 7TJ, U.K.

1. INTRODUCTION

There is little doubt that both weather and climate can have important influences on many aspects of society; agriculture, energy demand, water resources, recreation, etc. In general these impacts cannot be assessed using well-defined deterministic principles because of the complexities of both climate and of the facets of human livelihood which are susceptible to climatic impact. Society is particularly uncooperative in this regard since it constantly strives to reduce the impact of climate and climatic variability. In the absence of deterministic models, statistics and statistical models must be the main tools for identification, definition and quantification of climatic impact; but the application of these tools is not always straightforward.

The purpose of this paper is to discuss the role that statistics plays in climate impact analysis. No attempt is made to give an exhaustive review; indeed, I have been rather selective and have deliberately concentrated on relatively simple approaches. I have, however, attempted to place these statistical techniques in the broader descriptive framework of climate impact analysis as a whole.

Different types of impact can be distinguished according to time scale. Short-term impacts occur on time scales of minutes to days, often as the result of sudden weather events such as a tornado, tropical storm or flash flood. While more properly classified as weather impact (as opposed to climate impact) these effects can still be considered as resulting from prevailing climatic conditions with the particular weather event embedded in and possibly related to the medium time scale run of climatic variability. Medium-term impacts are those occurring on the weekly to annual time scale. Examples are the overall economic impact of a severe winter, agricultural losses due to anomalous climate over the growing season, reductions in hydroelectric generating capacity due to drought, etc. Long-term impacts occur on time scales of decades or longer and may be the cumulative result of shorter time scale effects associated with longer time scale changes in climate.

Identification of climate as the cause of impact becomes more difficult as the time scale becomes longer. For short-term impact, cause and effect are generally obvious or easy to isolate. For medium-term impact linkages are usually more difficult to establish, both a priori, because of the complexities in both natural and societal systems, and a posteriori, because of the statistical difficulty of isolating cause-effect relationships in a web of interconnected climatic, social, economic and technological variables. Long-term impacts are even more difficult to establish due to the dynamic qualities of societal systems which constantly alter their resilience to climate. Such impact analyses tend to be based on qualitative (and often oversimplified) reasoning.

In the following sections of this paper I will concentrate on medium- and long-term impacts. The general framework of climate impact will be introduced using some simple conceptual models which illustrate the complexities which may be involved. Statistical approaches will be classified into stratification techniques and correlation/regression techniques. General statistical problems will be described and both methods and problems will be illustrated with examples. Finally, a probabilistic model for climate impact will be discussed.

2. GENERAL CONSIDERATIONS

Conceptual models

Statistical methods can be used to establish the nature and magnitude of a relationship whose existence is not open to question, or to establish the reality of a link which is suspected but still subject to doubt. In symbolic terms, suppose I is some impact variable (the predictand) thought to be susceptible to the vagaries of climatic variability, and suppose that C is some set of possible causal climate variables identified a priori (candidate predictors). The primary aims of climate impact analysis are to answer the following questions:

1. *Identification*: is I related to C?
2. *Definition*: which subset of C is most important in determining I?
3. *Quantification*: what fraction of the I

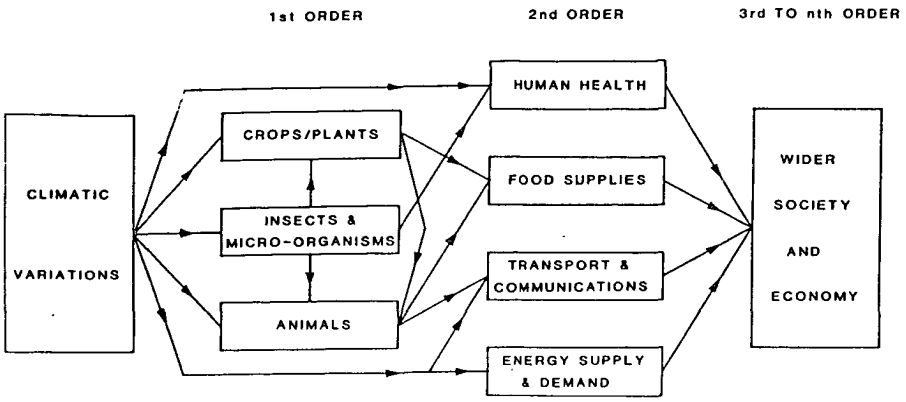


Fig. 1. Conceptual model of climate impact illustrating the interlinking chain of cause and effect relationships involved (from Ingram et al, 1981).

variance is explained by this subset and what is the quantitative relationship between I and C?

Unless a precise deterministic model is available, statistical techniques must still be employed to describe and quantify the I-C links. Generally the mechanisms involved in linking climatic data to observed impact are complex. Figs. 1 and 2 show conceptual models of climate impact illustrating this complexity.

Figure 1 depicts climatic impacts as an interlinking chain of cause-and-effect relationships with an assigned (somewhat arbitrarily) order of propogation, first, second,...to nth order impact. This sort of model underlies the methodological approach adapted by a number of large-scale climate impact studies, such as the U.S. Department of Transportation's Climatic Impact Assessment Program (CIAP) which was concerned with the impact of high-flying supersonic transports on climate and the subsequent climatic impacts on Man (Dept. of Transportation, 1974).

The distinction used in Fig. 1 is that first-order impacts are those affecting the biosphere. Second to nth order impacts involve the human element at a level where perception of the impact may lead to positive action on the part of the person or group affected to ameliorate the perceived impact. Indeed, in the interaction of climate and society "impact" is inextricably bound to human response. Impact is partly a function of the accumulation of individual and collective actions - technological, behavioural, socio-structural - taken before the event in response to past experience or in anticipation of future occurrences, and of the actions taken directly to minimise disruption and share the burden of losses during adverse climatic events (Burton et al. 1978; SCOPE, 1978). These actions, in addition to cultural changes which take place irrespective of weather and climate, constantly alter both the vulnerability and resilience of society. Thus, the relationships between climatic occurrences and societal impacts are dynamic and changeable over time, aspects which can create considerable difficulties for statistical analysis. Feedbacks like this are illustrated further in Fig. 2. In general the higher the order of the impact and/or the longer the time scale of the impact the more difficult it is to establish a quantitative relationship between an impact variable and climate.

Statistical strategies

Which statistical technique is appropriate depends on the amount and quality of available data, the particular questions asked (i.e. identification, definition or quantification of impact), and the order and time scale of the impact. Three different approaches can be distinguished:

(a) subdivision or stratification of the data on the basis of the values of a climate variable and examination of the groups for statistically significant differences,

(b) subdivision of the data into a discrete number of groups according to both the

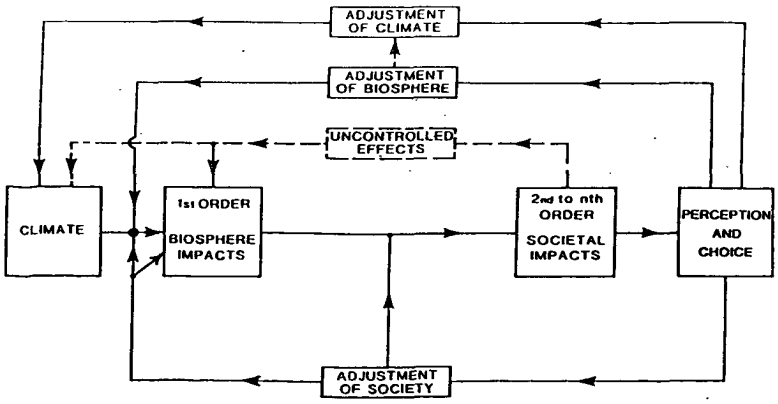


Fig. 2. Conceptual model of climate impact showing, in a simplified way, some of the complexities that can arise because of Man's power to modify environmental and societal responses to climatic variability (from Ingram et al, 1981).

magnitude of impact and the magnitude of the suspected climate forcing factor(s), and then testing for a significant relationship,

(c) direct comparison of time series using correlation or regression techniques.

Methods (a) and (b) are more appropriate for identification or confirmation of suspected impact relationships and can be applied to both qualitative and quantitative data. Method (c) has greater potential for establishing precise quantitative relationships between variables.

Statistical difficulties

These are all rather routine statistical procedures. Even so, in applying them to impact assessment problems, one is liable to encounter one or more of a number of problems which can affect the significance of the results. These problems have been discussed frequently in one important area of climate-impact assessment, crop-yield/climate modelling (e.g. Katz, 1977), and the literature in this area abounds with spurious results arising from elementary statistical errors. The main difficulties are -

Nature of the data base. In many studies data are available for only short time periods, data series may be inhomogeneous (socio-economic data are often given relative to a constantly changing base year, or may be produced in different ways in different years), and data may be of doubtful quality. Statistical estimates may, therefore, have unsatisfactorily wide confidence limits and statistical uncertainties may be compounded by uncertainties in the original data.

Non-Gaussian probability distributions. Some climate variables have highly non-Gaussian distributions and this may make standard statistical techniques inappropriate. Often a normalizing transformation is possible, but for some variables (such as precipitation) this may still be inadequate.

Non-stationary time series. Impact analyses invariably involve historical data from uncontrolled experiments. Trends in these data tend to be the rule rather than the exception. It is usually important to remove these trends before subjecting the data to statistical analysis; although in some cases the trend may be the factor under consideration. In the former case, simple curve-fitting or filtering techniques may be applied to define the trend, but there is an element of arbitrariness here which can pose further problems (see below). In the latter case, spurious associations may be detected because impact and climate variables may have similar trends by chance.

Autocorrelation. Both climate data and impact data often show persistence in time manifest as a positive lag-1 autocorrelation coefficient. Individual data points may, therefore, not be independent and the effective sample size may be substantially less than the actual number of data points. Unless this smaller effective sample size is taken into account, spurious significance levels may be obtained. There are standard techniques for allowing for the effects of autocorrelation (Quenouille, 1952; Mitchell et al, 1966), but these must be applied with caution because of the difficulty in distinguishing between trend (non-stationarity in the mean) and autocorrelation. The distinction between trend and serial correlation is an important one. Although both will be manifest by a similarity in consecutive observations, the former corresponds to a general change in the mean level, while the latter represents a tendency for fluctuations about the mean to perpetuate themselves. Some of the methods of allowing for autocorrelation are strictly applicable only to stationary data.

Spatial autocorrelation may also be a problem. Given a dearth of data at a single site, one may be tempted to aggregate data from nearby areas to boost the sample size (examples here come mainly from crop-yield/climate modelling; see, e.g., Starr and Tripathi, 1975). If the data are spatially autocorrelated, as they frequently are, spurious results may be obtained.

Intercorrelated predictors. In multiple regression analysis multicollinearity is frequently a problem, leading to unstable or totally spurious results. Climate variables are notorious for their intercorrelations, so considerable care needs to be taken in choosing predictor variables (see, for example, Palutikof, this volume). Once again there are standard methods for dealing with this problem (careful predictor selection, techniques like ridge regression and principal components regression, etc.). In some cases, intercorrelations may be used to advantage to reduce the number of candidate predictor variables.

Multiplicity. Models of climate impact are rarely well-defined and may involve a number of possible candidate predictor variables. In selecting predictor variables from a large candidate set, if enough independent candidates are correlated with a given impact variable a statistically significant result would eventually arise by chance even if the data were random.

Time dependence of relationships. This is perhaps the most invidious problem. Because humans are not passive receptors and may change the way they respond to environmental stress, either spontaneously or in response to perceived impacts, statistical relationships between impact predictand and climate predictors may change with time both in magnitude and form. Consequently, there may never be enough data to establish the impact-climate link; or, even if there are sufficient data, the derived relationship may not be applicable to the future.

Non-linearities. It is well known that relationships between, say, crop yield and temperature and precipitation may be non-linear. Non-linear relationships may also occur in other climate-impact situations. If a non-linear form is anticipated, then this presents a problem only in that an additional predictor variable may be required, reducing the number of degrees of freedom. If a strong linear relationship is indicated from a statistical analysis of the data this may simply be the linear part of a more complex link. Extrapolation beyond the calibration data set may, in such cases, lead to invalid conclusions. The underlying problem here is one of structure - unless a statistically-based model has the correct structure its value must be strictly limited.

3. STATISTICAL TECHNIQUES AND EXAMPLES

Stratification according to predictor values

This is one of the simplest means of statistically rigorous identification of climatic impact, valuable as an exploratory tool and/or when there are relatively few predictor variable data. I will illustrate the approach with modern data on wheat yields (per planted acre) from North Dakota. In this particular case, more sophisticated techniques can be used to describe the impact of climate on yield since there is an abundance of meteorological data available. For climate-impact studies in the historical past, however, when meteorological data are sparse, the approach has wider applicability.

The choice of using modern wheat yield data, incidentally, exposes a particularly vexing problem in climate impact analysis, a problem which deserves some elaboration before continuing. Most impact variables show long term trends (most noticeably, in the agricultural sector, an upward trend in crop yield over the past 3 to 5 decades) which are obviously not related to climate. For crop yields these trends are loosely referred to as technological trends although they include the aggregate influence of management practices (use of fertilizers and pesticides, irrigation, etc.), land use changes, government policies, prices and so on. In order to properly analyse climatic impact they must be accounted for independently and removed. However, in general, even less is known about the quantitative influence of technology than about the influence of climate, so, in order to establish the climate-yield link, technology effects are often factored out in an *ad hoc* way. The usual method is to assume the technology effect to be a piecewise linear or quadratic function of time and to consider only the residuals after fitting a simple curve of this type to the raw data. Essentially this is imposing a model of the following form:

$$Y = T_L + T_S + C$$

where Y is the impact variable, T_L and T_S are long and short-to-medium "technology" (more precisely, "non-climatic") factors, and C is the climate factor. T_L is chosen *a priori* to have some simple form, and T_S appears in the unpredictable error part of any statistical relationship between $(Y-T_L)$ and C. Instead of using a prescribed form for T_L one can filter the data with a low-frequency filter and consider only residuals from the filtered values. This amounts to admitting defeat in trying to pin down the long-term impact of climate on Y (since these effects cannot be distinguished from non-climatic effects on these time scales), and concentrating on identifying the short-to-medium time scale effects. The impact variable used for further analysis is therefore

$$y = Y - \tilde{Y}$$

where $\tilde{Y}$ is Y-filtered.

Of course, there is no reason why climatic and non-climatic effects should combine additively, so an alternative is to consider

$$y^* = Y/\tilde{Y}$$

(see, for example, Starr and Tripathi, 1975). We therefore have a considerable range of choices in deciding on the particular form of the impact variable to use in further statistical analysis: we can assume either an additive or a multiplicative model; and we have some freedom in choosing the "technology" function's form, or the length of the filter to use to isolate the shorter time scale fluctuations in the impact variable.

To return to North Dakota wheat yields, I have chosen to remove the technology effects using differences from 30-year Gaussian filtered values. As a climate predictor I have used a measure of drought extent and intensity, namely the number of division-months of Palmer drought index values of -2.00 or less (i.e. at least moderate drought) summed over April to July (these data were kindly provided by Dr R.A. Warrick). On this basis, each year can be categorized as a year of little or no drought, moderately extensive drought or widespread drought conditions. Fig. 3 shows the residual yields for 1892 to 1980 and indicates, for each year, the drought severity. The importance of drought (as defined here) in determining yield can be judged by examining the average yields in drought and non-drought years (Table 1).

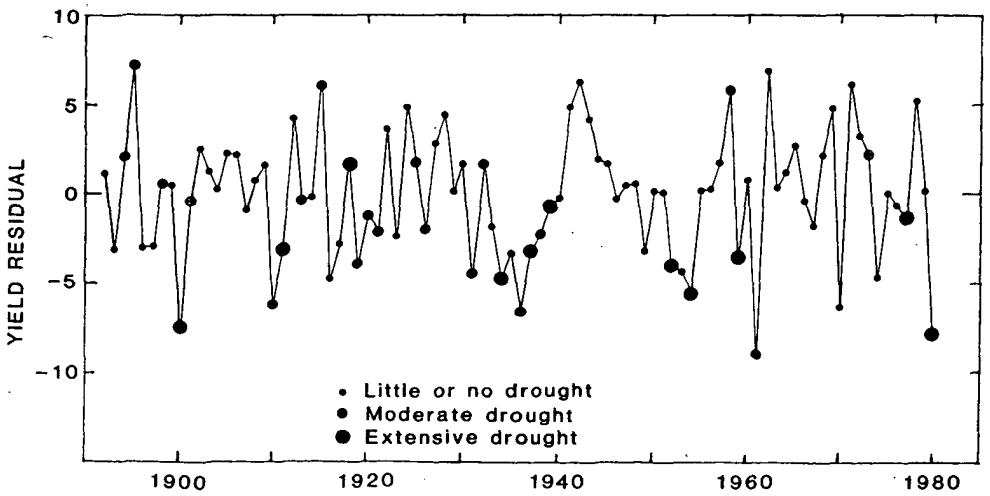


Fig. 3. Spring wheat yields for North Dakota, 1892-1980, as residuals from 30-year Gaussian filtered values (bushels per planted acre).

	No. of years	Average residual yield	Standard deviation
Years with extensive drought	11	-3.21	2.78
Remainder	78	0.40	3.41
Years with moderate or extensive drought	29	-1.51	4.09
Remainder	60	0.66	3.03

Table 1: Means and standard deviations of North Dakota wheat yields (residuals from a 30-year filter in bushels per planted acre).

Comparing years of extreme drought with those with less severe conditions, the t-value for the difference in means is 3.35 (87 d.f.). For years of moderate-to-extensive drought compared with those with little or no drought the corresponding t-value is 2.82 (87 d.f.). Both differences are highly significant. Nevertheless, it is apparent from Fig. 3 that the drought-yield relationship is far from being a perfect correspondence, so that other factors, both climatic and non-climatic, probably play important roles. For example, in some of the years of low yield which were relatively drought-free rust contributed to harvest losses (viz. in 1916, 1917, 1923 and 1949; rust occurred also in 1950 and 1954 - R.A. Warrick, personal communication).

An interesting application of this technique in a historical context has been given by de Vries (1980) in trying to assess the impact of the severe climatic conditions which prevailed in Europe during the Little Ice Age. De Vries compared average butter prices at Leiden for years with frosty Marches to those for other years over the period 1658-1757. He found a statistically significance difference, with prices in the "cold" years about 12% above those in other years. De Vries, in fact, did not use a t-test to test for differences in the means, but used the Normal distribution approximation. This made no difference to his result. The most interesting facet of de Vries's work is that it represents one of the few studies of historical climate impact in which a rigorous statistical approach has been used.

An aside on variability

Before considering a slightly more complex statistical technique for assessing impact, it is of interest to consider the effect of technology in attempting to assess whether or not crop yields are becoming more or less sensitive to interannual climate fluctuations. This is a question of considerable importance since it has immediate bearing on Man's susceptibility to future variations in climate. A superficial examination of many crop-yield time series shows two characteristics; a general upward trend in yields already alluded to, and an apparent change in interannual variability in recent decades. Whether or not these changes in variability are associated with changes in climatic variability or with changes in the sensitivity of modern crops and agricultural practices to climatic fluctuations is not known; different investigators have reached different conclusions (compare, for example, the findings of NOAA (1973), Perrin and Heady (1975), Haigh (1977) and Newman (1978)).

Before attempting to analyse data to find the cause of a change in variability, however, we might first ask whether there has been a change in variability at all. This question has been addressed by Waggoner (1979) and Dennett (1980), but their results are inconclusive because the answer depends on whether one assumes that technology has an additive or multiplicative effect. As an example, Fig. 4 shows non-overlapping decadal values of the standard deviations of residual yields for North Dakota wheat using both models. For the additive model there is an increase in yield variability, while for the multiplicative model, yield variability has decreased in recent decades. The results also depend (to a lesser degree) on the type of filter used in defining the background technological trend. This is a salutory indicator of the uncertainties which can arise in impact assessment because of our inadequate understanding of the non-climatic factors involved.

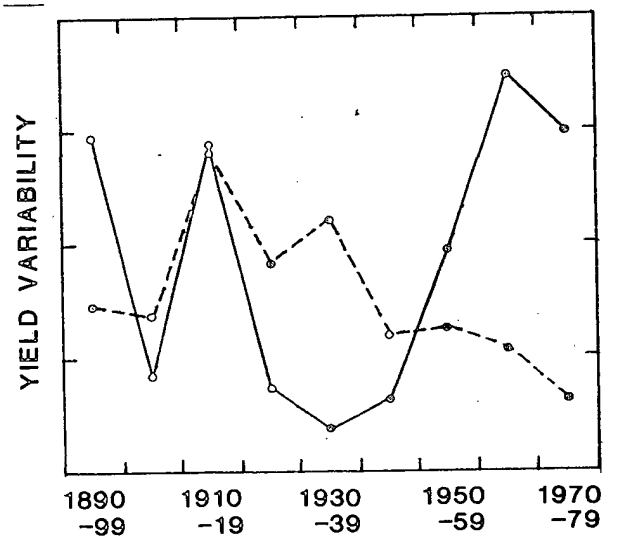


Fig. 4. The effect of the method of removing the technology effect on the interannual variability of residual yields of North Dakota wheat. Dashed line is for a multiplicative model, full line is for an additive model. Arbitrary scale.

Stratification of predictand and predictor variables

Here impact may be assessed using contingency tables. The method is particularly useful in dealing with data which are qualitative or quantitatively uncertain, or when data are available only as incomplete time series. These data deficiences do not, however, preclude the use of rigorous statistical techniques for testing the reality of climatic impact or for identifying important predictor variables. The appropriate technique can be illustrated with an example of historical climate impact, Ogilvie's (1981) analysis of the impact of climate on grass growth

(an important precursor of hay yield) in eighteenth century Iceland. The reality of a link in the twentieth century has been established using agronomic studies and regression analysis. An examination of contemporary eighteenth century descriptive accounts suggests that grass growth and hay yields were adversely affected by cold weather in spring. (This in turn is related to the incidence of sea ice around the coasts of Iceland which was rather more frequent in the eighteenth century than today.) Although only qualitative data are available, these can be grouped into cold, average and mild for spring weather, and poor, average and good for grass growth. Data covering the period 1681-1780 from a number of regions are grouped together in Table 2 which shows the observed number of occurrences in each joint category and the expected number of occurrences (in brackets) based on the null hypothesis of no relationship between growth and spring temperature.

Grass growth \ Spring temp.	Cold	Average	Mild	Totals
Poor	89 (51.8)	10 (32.1)	6 (21.1)	105 (35.2%)
Average	46 (67.1)	67 (41.5)	23 (27.4)	136 (45.6%)
Good	12 (28.1)	14 (17.4)	31 (11.5)	57 (19.1%)
Totals	147 (49.3%)	91 (30.5%)	60 (20.1%)	298 (100%)

Table 2. Climate and grass growth in Iceland, 1681-1780. Numbers denote number of occurrences, bracketted numbers are estimates based on the assumption of no crop-climate relationship. $\chi^2 = 118.8$; significance level better than 0.001.

A significant difference between observed and expected frequencies, based on the value of the χ^2 test statistic indicates a relationship between variables. The value of χ^2 (118.8) in this case is highly significant and indicates a strong relationship; the observed incidence of poor grass growth following cold springs is much larger than would be expected by chance.

Correlation methods

The most obvious way in which climatic impact can be demonstrated when time series of data are available is simply to correlate the impact variable with the suspected climate predictor variable(s). In such analyses one must try to avoid problems of multiplicity. This aside, there are still pitfalls which, while obvious to the statistician, are not always noticed by those studying climate impact.

As an example, consider the comparison made by Libby (1977) between the price of wheat (normalised to account for changes in the value of money) and late-winter/early-spring temperatures in England. Fig. 5 shows these two time series over the period since about 1660 when temperature data first became available. Libby subjectively judged the relationship to be significant, although when the data are correlated over the whole period 1660 to 1900, the correlation coefficient is only about 0.22 which is not significant at the 5% level. The eye is, however, immediately drawn to the period 1725-1900 over which the two curves are more in line. The correlation

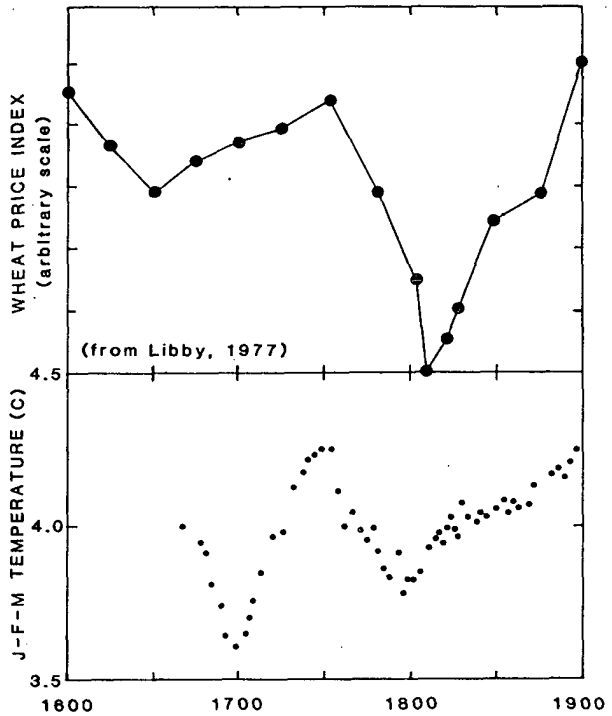


Fig. 5. Variations in the price of European wheat corrected for the changing value of currency using the price of sterling silver (the index is the ratio of the price of sterling to the price of wheat), compared with variations in Jan.-Feb.-Mar. average temperature (from Manley, 1974). Diagram adapted from Libby (1977).

coefficient here is 0.63 which is apparently significant at the 1% level (n=19). This is, however, a spurious result induced by the strong autocorrelation in both time series which severely limits the number of degrees of freedom. A simple (but approximate) method for accounting for autocorrelation is that due to Bartlett (see, for example, Quenouille, 1952, p. 168). The effective number of degrees of freedom is determined using

$$\nu_{eff} = (n - 2)(1 + 2r_{1,a}\, r_{1,b} + 2r_{2,a}\, r_{2,b} + \ldots)^{-1}$$

where $r_{i,a}$ and $r_{i,b}$ are the lag-i autocorrelation coefficients for the two time series being considered and n is the sample size. The difficulty in using such a formula as this is that the autocorrelation structure of both climate and impact variables tends to be rather unstable (Fig. 6 shows a rather extreme example). In the light of such instability, a simpler formula (the AR1 equivalent of the above) is probably all that is warranted:

$$\nu_{eff} = (n - 2)\,\frac{1 - r_{1,a}\, r_{1,b}}{1 + r_{1,a}\, r_{1,b}}$$

In the Libby example, both series have lag-1 autocorrelations exceeding 0.75 and the effective sample size ($n_{eff} = \nu_{eff} + 2$) is only about 7. The correlation coefficient, 0.63, is not statistically significant.

Regression analysis

One of the most powerful statistical tools

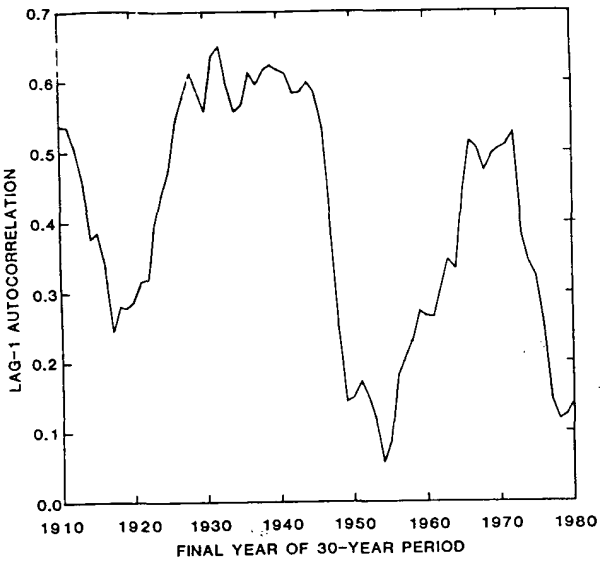


Fig. 6. Running 30-year lag-1 autocorrelations of Northern Hemisphere annual mean surface air temperature (data from Jones et al, 1982).

for climate impact assessment is regression analysis. Provided some form for the functional relationship between impact predictand(s) and climate predictor(s) can be defined then unknown parameters can be optimized using regression techniques to produce a complete quantitative equation. Depending on how well the impact mechanism, the structure of the relationship, is known the resulting regression equation may represent, at one extreme, an expression of the actual processes involved or, at the other extreme, a "black box" with little or no mechanistic information or insight.

All of the difficulties outlined earlier apply to regression analysis. If suitable data are available one is first likely to have to remove obfuscating non-climatic factors such as the effects of technology. The sensitivity of subsequent results to the method of removal (additive or multiplicative model, choice of filter, etc.) has rarely been examined (but see, e.g., Mostek and Walsh, 1981). In choosing candidate climate predictors, one may be faced with potential multicollinearity and/or multiplicity problems. When the regression has been performed, a subset of predictors may have to be chosen from the original candidate predictor set. The decision here may be based on the significance levels of individual candidate predictors (e.g. as in Briffa et al, 1983) or on minimizing some overall test statistic (Shibata, 1980, reviews some of the possibilities). Finally, the statistical significance of the final regression equation must be estimated. This is a particularly difficult problem. The standard F-test can easily give a false indication of the reliability of a regression result; particularly when there has been a posteriori selection of predictors from a larger candidate predictor set (Barnett and Hasselmann, 1979). "No one knows how to appraise a (regression result) safely except by using different bodies of data from those that determined it" (Mosteller and Tukey, 1977, p. 37) so some form of cross-validation or verification is essential. In spite of this need for independent testing of regression results, adequate cross-validation is rare in climate impact analyses.

These are all well-known problems that apply to regression analysis in general. In climate impact applications there is the added possibility that the structure and/or magnitude of relationships may change with time. A striking example of this comes from the energy field. In analysing the sensitivity of summer electric power demand in the U.S. midwest to temperature and other climatic characteristics, Johnson et al. (1969) and McQuigg et al. (1972) found steadily increasing sensitivity with time over the period 1962-70. Their results, relating peak daily load in summer to daily mean temperature, are shown in Fig. 7. Both absolute load and sensitivity to temperature increase steadily over the study period, a change which they attribute to economic and life style changes. The upward trend in sensitivity has continued at least up to 1974 (McQuigg, 1975, Fig. 13). Such changes in climate impact relationships may be the rule rather than the exception. In agriculture, in particular, there are many reasons to expect this to be so. Apart from direct effects such as the increasing use of irrigation to reduce drought sensitivity, more indirect changes may result from the effects of large-scale atmospheric pollution through acid rain and increasing carbon dioxide levels.

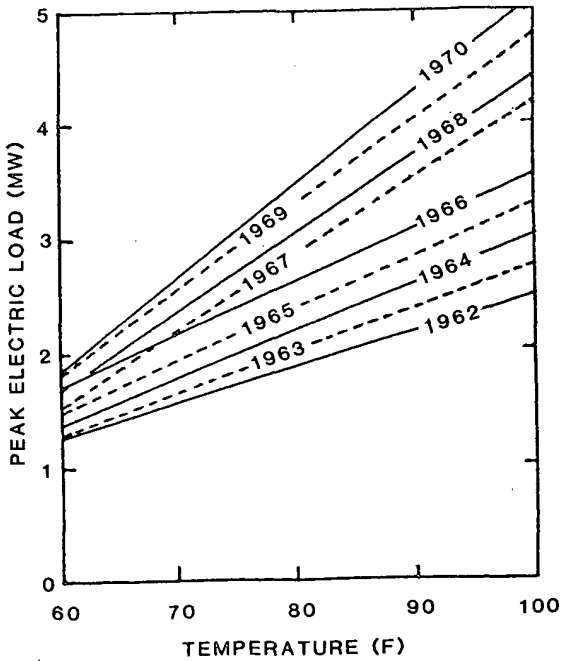


Fig. 7. Best fit regression lines relating daily peak electric load to daily mean summer temperature in the US midwest (from McQuigg et al, 1972).

4. PROBABILISTIC MODELS OF CLIMATE IMPACT

Parry (1981) has proposed a simple probabilistic model which can be used to show how short- to medium-term impacts of climate on society may have a cumulative longer-term effect in the presence of a changing climate. Although devised in a historical context the ideas have contemporary relevance given the possibility of a major climatic change due to the increasing levels of

atmospheric carbon dioxide and the carbon dioxide greenhouse effect. Parry argues that, while year-to-year fluctuations in the weather or climate of a region may have identifiable impacts, the resilience of societies is such that only the impacts due to rare extreme events are liable to have effects that carry over to the subsequent year. Cumulative effects of significance can, therefore, result when extremes follow each other in consecutive years. There are a number of examples of this in the historical climate literature (see, e.g., Pfister, 1981, and Flohn, 1981), and the recent occurrence of severe winters in the eastern U.S.A. in 1976/7 and 1977/8 is a modern example of two consecutive severe seasons.

If we suppose that extreme events occur at random with probability p then the probability of two extremes in consecutive years will be p^2. (The randomness of extremes is something which can easily be tested, although there are relatively few examples in the climatological literature. Interesting cases are the studies of the frequency of Indian monsoon failure by Mooley and Pant (1981) and the probability model for severe cyclonic storms in the Bay of Bengal developed by Mooley (1980). When periodicities exist, however, extremes may not be random; a particular case in point being the occurrence of drought in the U.S. Great Plains (Mitchell et al, 1979; Stockton et al, 1983).) Following Parry's hypothesis, long-term climatic impact is therefore related not to p, but to p^2. The importance of this proposition is that p^2 is far more sensitive to climatic change than is p. To illustrate, consider a numerical example and suppose that severe winters or drought summers occur with a probability p = 0.1. The return period for the occurrence of a single severe season is, therefore, 10 years, while the return period for the occurrence of two consecutive severe seasons is 100 years. Any change in climate will lead to a change in p, either through changing variability which will change p directly, and/or through a change in mean conditions which must also change p if severity is judged relative to an absolute threshold. Alternatively, p may change through changes in some critical impact threshold as a result of land use changes, new crops or crop mixes, increasing population pressure, etc. If p becomes 0.2 then the return period for a single severe season is halved to 5 years. The return period for consecutive severe seasons, however, is reduced by a factor of four to only 25 years.

Although highly idealized, this model is a thought-provoking one. It highlights a number of factors which need to be considered in assessing the impacts of future climatic change on society. Firstly, the model highlights the importance of changes in variability; but note that equivalent effects, through changes in the frequency of extremes, might also arise simply through a change in mean climatic conditions, or through changes in non-climatic factors. Secondly, the model introduces the concept of a threshold of sensitivity to climate impact and to the possibility that the impacts of long-term climatic change may be non-linear. And thirdly, the resilience of society to such changes may depend on how rapidly recovery can be made to ensure minimal carry-over effects to subsequent years. These are not new ideas, but they are factors which need to be explored at greater depth both through detailed case studies and through more sophisticated approaches to climate-impact modelling.

5. CONCLUDING REMARKS

Because of the lack of deterministic models of climate impact, statistics must play a central role in the field, in identification, description and quantification of climate-impact relationships. Such studies are plagued by difficulties. In order to isolate the climatic component, other non-climatic factors must first be eliminated. Since these are also poorly understood the resulting derived climate effects may be partly dependent on how the non-climatic effects are accounted for. At the very least, ones success at determining the non-climatic elements may set an upper bound on how well the climate aspect of impact can be defined. Having hurdled this difficulty, numerous statistical problems can arise, which, if not handled carefully and honestly may lead to false estimates of the strength of the derived climate-impact link.

It is no coincidence that the majority of examples considered here have involved either first order impacts, or impacts in which the link between climate and the impact variable is direct. Very few publications in the field have considered those higher order impacts which are influenced by the dynamic nature of the response of society to environmental stress. Likewise, few studies have been able to address the problem of trends or changes in the nature or strength of climatic impact which are a necessary consequence of our ability to adapt to perceived external pressures. Changes in climate-impact relationships resulting from uncontrolled modifications of the environment (acid rain and increasing levels of atmospheric carbon dioxide are two obvious examples) have yet to be investigated in any rigorous statistical way. There are opportunities and challenges here for both statisticians and climatologists.

Other challenges lie in the assessment of climate impact in the lesser developed countries of the world. Here, the lack of data is a major problem so there must be a synthesis where results from other regions are modified and transposed to new conditions and where statistical techniques which are less sensitive to data deficiencies are employed.

There is considerable scope for making quantitative assessments of second and higher order impacts of climate on society by using rigorously derived primary impact results in complex econometric, ecological or systems dynamics models. Most work on climate impact analysis espouses a reductionist approach in which a complex problem is broken down into simple constituent parts. However, the whole is, in this case, not always the sum of the parts, but also includes the myriad interactions between these parts and the buffering (or amplifying) capacities of those parts which experience no direct impact. Simulation of these interactions using empirical socio-economic models may be one way to estimate some of the impacts of future climatic changes, but this type of analysis must be based on statistically rigorous input data

and must involve models which themselves have been properly tested and evaluated.

Finally, the examples given in this paper have all used quite simple statistical procedures. In many cases this is all that the available data warrant, and even these procedures can be misused. Nevertheless, there is considerable scope for the application of more sophisticated procedures, and techniques like jack-knifing, multivariate regression, factor analysis, ARIMA modelling, intervention analysis, etc., are largely unexplored. As climate impact questions become more sharply honed and data requirements more clearly defined, we can expect that such statistical tools will make an increasingly significant contribution to studies of climate and society.

Acknowledgements: Discussions with Dr R.A. Warrick were most valuable and (if a suitable test could be devised) most probably had a statistically significant impact on this paper.

6. REFERENCES

Barnett, T.P., and K. Hasselmann, 1979: Techniques in linear prediction with application to oceanic and atmospheric fields in the Tropical Pacific. Rev. Geophys. Space Phys., 17, 949-968.

Briffa, K.R., P.D. Jones, T.M.L. Wigley, J.R. Pilcher and M.G.L. Baillie, 1983: Climate reconstruction from tree rings: Part 1, Basic methodology and preliminary results for England. J. Climatology (in press).

Burton, I., R.W. Kates and G.F. White, 1978: The Environment as Hazard, Oxford University Press, New York.

Dennett, M.D., 1980: Variability of annual wheat yields in England and Wales. Agric. Met., 22, 109-111.

Dept. of Transportation, 1974: Climatic Impact Assessment Program Monographs 1-6, Washington, D.C.

Flohn, H., 1981: Short-term climatic fluctuations and their economic role. In, Climate and History (T.M.L. Wigley, M.J. Ingram and G. Farmer, Eds.), Cambridge University Press, Cambridge, 310-318.

Haigh, P.A., 1977: Separating the effects of weather and management on crop production. Report to the Charles F. Kettering Foundation.

Ingram, M.J., G. Farmer and T.M.L. Wigley, 1981: Past climates and their impact on Man: a review. In, Climate and History (T.M.L. Wigley, M.J. Ingram and G. Farmer, Eds.), Cambridge University Press, Cambridge, 3-50.

Johnson, S.R., J.D. McQuigg and S.R. Rothrock, 1969: Temperature modification and costs of electric power generation. J. Appl. Meteor., 8, 919-926.

Jones, P.D., T.M.L. Wigley and P.M. Kelly, 1982: Variations in surface air temperatures: Part 1. Northern Hemisphere, 1881-1980. Mon. Wea. Rev., 110, 59-70.

Katz, R.W., 1977: Assessing the impact of climatic change on food production. Climatic Change, 1, 85-96.

Libby, L.M., 1977: Correlation of historic climate with historic prices and wages. Indian J. Met. Hydrol. Geophys., 28, 247-250.

McQuigg, J.D., 1975: Economic Impacts of Weather Variability. Report to N.S.F. under Grant No. GI37218X, ESR73-07752-A01, Dept. of Atmospheric Science, Univ. of Missouri, Columbia.

McQuigg, J.D., S.R. Johnson and J.R. Tudor, 1972: Meteorological diversity - load diversity, a fresh look at an old problem. J. Appl. Meteor., 11, 561-566.

Manley, G., 1974: Central England temperatures: monthly means 1659 to 1973. Quart. J. Roy. Met. Soc., 100, 389-405.

Mitchell, J.M., Jr., B. Dzerdzeevskii, H. Flohn, W.L. Hofmeyr, H.H. Lamb, K.N. Rao and C.C. Wallén, 1966: Climatic Change. W.M.O. Tech. Note No. 79 (W.M.O. No. 195, T.P. 100), Geneva, Switzerland.

Mitchell, J.M., Jr., C.W. Stockton and D.M. Meko, 1979: Evidence of a 22-yr rhythm of drought in the western United States related to the Hale solar cycle since the 17th century. In, Solar Terrestrial Influences on Weather and Climate (B.M. McCormac and T.A. Seliga, Eds.), Reidel, Dordrecht, 125-143.

Mooley, D.A., 1980: Suitable probability model for severe cyclonic storms striking the coast around the Bay of Bengal. In, Statistical Climatology (S. Ikeda, E. Suzuki, E. Uchida and M.M. Yoshino, Eds.), Elsevier, Amsterdam, 349-357.

Mooley, D.A., and G.B. Pant, 1981: Droughts in India over the last 200 years, their socio-economic impacts and remedial measures for them. In, Climate and History (T.M.L. Wigley, M.J. Ingram and G. Farmer, Eds.), Cambridge University Press, Cambridge, 465-478.

Mostek, A., and J.E. Walsh, 1981: Corn yield variability and weather patterns in the U.S.A. Agric. Met., 25, 111-124.

Mosteller, F., and J.W. Tukey, 1977: Data Analysis and Regression. A second course in statistics. Addison-Wesley, Reading, Mass.

Newman, J.E., 1978: Drought impacts on American agricultural productivity. In, North American Droughts (N.E. Rosenberg, Ed.), Westview Press, Boulder, Colorado.

NOAA, 1973: The Influence of Weather and Climate on United States Grain Yields: Bumper Crops or Droughts. Report to the Administrator NOAA (J.D. McQuigg et al.).

Ogilvie, A.E.J., 1981: Climate and Society in Iceland from the Medieval Period to the Late Eighteenth Century. Unpublished Ph.D. dissertation, Univ. of East Anglia, Norwich, U.K.

Palutikof, J.P., 1983: Climate-transport relationships in Great Britain. Preprints, Second International Meeting on Statistical Climatology, Lisbon (this volume).

Parry, M.L., 1981: Climatic change and the agricultural frontier: a research strategy. In, Climate and History (T.M.L. Wigley, M.J. Ingram and G. Farmer, Eds.), Cambridge University Press, Cambridge, 319-336.

Perrin, R.K., and E.O. Heady, 1975: Relative contributions of major technological factors and moisture stress to increased grain yields in the midwest, 1930-71. CARD Report 55, Center for Agricultural and Rural Development, Iowa State University.

Pfister, C., 1981: An analysis of the Little Ice Age in Switzerland and its consequences for agricultural production. In, Climate and History (T.M.L. Wigley, M.J. Ingram and G. Farmer, Eds.), Cambridge Univ. Press, Cambridge, 214-248.

Quenouille, M.H., 1952: Associated Measurements, Butterworth Scientific Publications.

SCOPE (Scientific Committee on Problems of the Environment), 1978: Workshop on Climate/Society Interface, SCOPE secretariat, Paris.

Shibata, R., 1980: Selection of the number of regression parameters in small sample cases. In, Statistical Climatology (S. Ikeda, E. Suzuki, E. Uchida and M.M. Yoshino, Eds.), Elsevier, Amsterdam, 137-148.

Starr, T.B., and R. Tripathi, 1975: The response of United States spring wheat yields to variation in climate. In, A Detailed Model of of the Production and Consumption of Spring Wheat in the United States, Institute for Environmental Studies, Univ. of Wisconsin-Madison.

Stockton, C.W., J.M. Mitchell, Jr., and D.M. Meko, 1983: A reappraisal of the 22-year drought cycle. In, Weather and Climate Responses to Solar Variations (B.M. McCormac, Ed.), Colorado Associated University Press (in press).

De Vries, J., 1980: Measuring the impact of climate on history: the search for appropriate methodologies. J. Interdisciplinary History, 10, 599-630.

Waggoner, P.E., 1979: Variability of annual wheat yields since 1909 and among nations. Agric. Met., 20, 41-45.

ASSESSING THE IMPACT OF CLIMATE CHANGE ON THE DEVELOPMENT OF SURFACE WATER RESOURCES

V. Klemeš

National Hydrology Research Institute
Department of the Environment
Ottawa, Ontario, K1A 0E7, Canada

J. Němec

Hydrology and Water Resources Department
World Meteorological Organization
Geneva 20, Switzerland

1. INTRODUCTION

Large storage reservoirs with controllable rates of release are, at the present time, the most effective means for the development of surface water resources. They are the only means capable of redistributing the irregular natural streamflow in accordance with water needs of a region, in particular of increasing streamflow during dry periods and reducing it during wet periods.

There are several reasons why the planning of these reservoirs calls for an accurate knowledge of future climate. The first and most basic is that the size of a reservoir required for a given degree of water resource development depends on streamflow properties during its period of operation which in turn are a result of climate during that period. In the current practice, the design of these reservoirs is based on the assumption that the climate is approximately stationary - that during the next several decades it will not be significantly different from its conditions in the past 50 to 100 years. If this should not be the case then our present estimates of reservoir sizes, feasible levels of water resource development, and the overall water resource planning would be in error. The second reason is that these errors might easily involve billions of dollars since, as a rule, the reservoirs involve the construction of large dams, new roads, bridges, etc., which all are expensive undertakings. And the third reason is that all these structures, once finished, are rather difficult to change so that any errors made are literally cast in concrete.

It seems obvious that an accurate forecast of climatic conditions (not to mention an accurate forecast of streamflow in specific rivers) over the period of, say, the next 50 years is an unrealistic proposition, notwithstanding the repeated attempts to discover deterministic trends in climate associated with a wide variety of phenomena ranging from the wobbling motion of the Earth to the enigmatic sun spots (the hypothesis that the 11 year sun spot cycle is the key factor enjoys a popularity which itself seems to exhibit an 11 year period - now again it may be passing through a maximum because of the periodicity of the presently fashionable layers in the Pichi-Richi Pass; Williams, 1981; Pittock, 1983).

Thus the fact remains that contemporary water resource development planning must take the uncertainty regarding future climate into account. In general, taking uncertainty into account is not new to water management. For example, the uncertainty resulting from the stochastic nature of streamflow series was first considered in reservoir design about 70 years ago by the famous American hydrologist and engineer Allan Hazen (the inventor of "probability paper").

The first problem that water resource management is facing vis-à-vis a possible climate change in the near future is to estimate the range of the consequent change in the runoff and compare it with the range of random fluctuations of runoff which can be expected even under the assumption of a stationary climate. Only if it can be conclusively shown that runoff changes during the next two or three decades (beyond which horizon no serious planning activity is now possible) are likely to be much higher than those due to the stochastic nature and the attendant sampling variability of a stationary runoff process, only then can the water resources planner face a practical problem of what can now be done about them. In other words, the first task is to find out whether runoff changes due to a climate change are statistically significant within the context of 1) our knowledge of the stochastic structure of the runoff process, 2) a planning horizon of about 20 years, and 3) the hydrological data base used for the design of large storage reservoirs.

The aim of this paper is to examine these problems and to indicate ways that may lead to their better understanding.

2. QUANTIFYING THE WATER RESOURCE DEVELOPMENT POTENTIAL

A convenient means for quantifying the potential for development of surface water

resources at a given point along a river is the so called reservoir regime function which shows the reliability with which a given constant rate of flow can be sustained in the river by a reservoir of a given storage capacity. This rate of flow is usually referred to as the target release (from the hypothetical reservoir) or safe yield, and its reliability (or degree of safety) is expressed either as a percentage of time during which the release does not drop below the target value, or as a probability that it does not drop below the target during a period of one year, or by some similar characteristic. As an example, a reservoir regime function for the Leaf River (at Collins, Mississippi) is shown in Fig. 1.

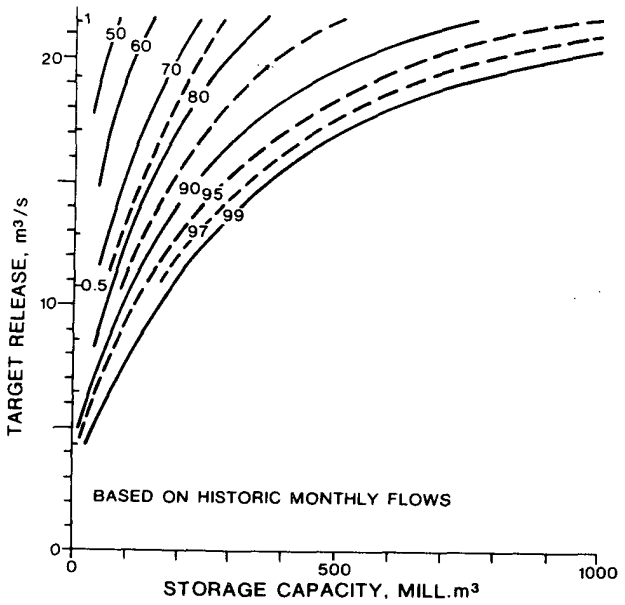


Fig. 1. Reservoir regime function for Leaf River (isolines are drawn for time-based reliability, in percent).

Quantitative relations between the three variables of the regime function (the target, storage, and reliability) are uniquely specified by the properties of the stochastic process representing the streamflow. Thus, in theory, the effect of climate change on the potential for surface water development can be determined from the difference between regime function based on the stochastic process representing the historic streamflow and regime function based on the stochastic streamflow process corresponding to the conditions after a climatic change. The effect of climatic change could thus be expressed either as a difference in the reliability with which a given reservoir could maintain a given target release, or as a difference in the value of target release corresponding to given values of reliability and storage capacity, or finally as a difference in the storage capacity of a reservoir needed for the maintenance of a given release with a specified reliability.

Since both of the stochastic streamflow processes (pre- and post-climate change) would be estimated from finite data samples, and the regime functions constructed for some finite period of reservoir operation, each of the three variables being compared would have a probability distribution. In general, the extent of the overlap of the pre- and post-climate change distributions would indicate the degree of statistical significance of the climate change impact on the surface water resource development potential.

3. PROBLEMS INVOLVED

While the program outlined above is theoretically straightforward, its practical execution poses several serious problems. One is a quantitative estimate of the changes in the primary climatic variables such as air temperature, radiation, precipitation, and evapotranspiration; moreover, it is not only the changes in the long-term normals that are of interest here but also the changes in seasonal variability and, in general, in the properties of the associated stochastic processes. The second problem, which in theory should be less difficult to overcome but in practice does not seem to be so, is the modelling of the mechanism by which the primary climatic variables are transformed into the streamflow process. To be sure, a large number and a great variety of such models (commonly called hydrological) exist, ranging from purely black-box input-output (transfer function) models to "grey-box" conceptual models of various shades of grey. However, none of them qualifies as a climate-transferable model, or at least none has so far been conclusively proven to qualify as such (Klemeš, 1982b). A third problem is our lack of understanding of the stochastic structure of the streamflow process even under stationary conditions (Klemeš, 1974) - a sine qua non for establishing a firm basis for meaningful comparisons; this problem is of a fundamental nature since our streamflow records are generally short and do not contain enough information for solving it (Moran, 1957; Wallis and O'Connell, 1973).

4. THE OPTIONS

Under the circumstances outlined in the preceding section there are basically two options: one is to refrain from any attempts to assess the impact of climate change on the development of water resources; the other is to see how far one can get using the current state of the art. The first approach is rigorous and safe; the second is likely to be just the opposite but, we believe, more productive. It may, and often will, lead to wrong answers and dead ends. But by forcing the currently available concepts, models, etc. to work and interact in a clearly defined framework designed with a specific aim in mind, it should be possible to identify more readily the weakest links, to see where, and what kind of, research is needed, and whether one direction is likely to be more promising than another. As a matter of fact, the more suspect and contradictory the results, the greater the incentive for

improvement and rethinking of the concepts employed. It may be pertinent in this context to quote the Nobel prize winner in physics, Richard Feynman: "The thing that doesn't fit is the thing that's the most interesting, the part that doesn't go according to what you expected...[our laws] sometimes look positive, they keep on working and all of a sudden, some little gimmick shows that they're wrong. And then we have to investigate the conditions under which [they are wrong] and so forth, and gradually learn the new rule that explains it more deeply" (NOVA, 1983).

The only real danger posed by this second approach is not that its results may be wrong, but that they might be accepted uncritically as correct, either because of the prestige of their authors, or of the institutions with which the authors are associated, etc., and then used as sacred dogma to suppress whatever might contradict them. To put it more positively, the second approach can be productive only if perceived as a catalyst in a learning process.

5. SOME RESULTS

An example of the second approach is a recent attempt to assess the impact of climate change on the development of surface water resources in two specific locations of the United States made by Němec and Schaake (1982) and followed up by Klemeš (1982a, 1983).

The first two authors considered 12 different scenarios of climate change covering the approximate range of possibilities presently considered compatible with the buildup of atmospheric CO_2. These scenarios were characterized by all combinations of three different changes in air temperature converted into three corresponding values of evapotranspiration (E) changes of 12%, 4% and -4%, and the following four precipitation (P) changes, 25%, 10%, -10% and -25%. These changes were then introduced as perturbations of the E and P variables in a Sacramento conceptual hydrological model (which had been calibrated on historical records of climate variables and streamflow) for two U.S. basins described in Table 1, and 12 hypothetical post-climate-change

Table 1. Characteristics of River basins under study

	Leaf R. Collins, Miss.	Pease R. Vernon, Tex.
Basin area	1949 km^2	9034 km^2
Annual precip.	1314 mm	540 mm
Annual runoff	409 mm	11 mm
Record length	18 y	11 y

streamflow series were generated for both basins. The third author then constructed reservoir regime functions of the type shown in Fig. 1 for the historic streamflow series and for the 12 simulated series.

An example of the result is presented in Fig. 2a, b, c, d. It shows how the reliability of four different levels of development (for the Leaf River), for all of which the reliability under a stationary climate is 95% (black circle), might change for any of the 12 climate change scenarios (open circles). We do not claim that the reliability values shown are accurate; most likely they are not because of the presently inevitable imperfections both in the representation of the post-change climatic forcing functions and in the hydrological model employed. However, they give an idea of a possible order of magnitude of the impact and of some of its likely features that may be of interest to the water manager, for instance: 1) the drop in reliability might be - percentagewise - much faster than the drop in precipitation or the rise in evapotranspiration; 2) the impact of a drier climate would be more severe where the present level of development is high (Fig. 2a, b) than where it is low (Fig. 2c, d); 3) the relative effect of the precipitation change would probably be greater than that of the evapotranspiration change.

6. STATISTICAL SIGNIFICANCE AND ITS PROBLEMS

While the changes such as those shown in Fig. 2 are interesting, they lack a proper perspective unless their statistical significance is assessed, i.e. unless they are compared with conditions that might occur simply because of sampling fluctuations of a stationary climate or, in the present context, streamflow. Such a perspective can be provided by confidence limits corresponding to a given reliability value for stationary climate, e.g. to the value of 95% in our case. As mentioned before, these limits can be derived if the stochastic process of the streamflow is known. Unfortunately, in our case the length of the historic flow record is very short so that the stochastic structure of the process cannot be satisfactorily identified. Respecting Occam's razor, one has to settle for a random (uncorrelated) series model for the annual flows. On this basis, and taking into account the variability inherent in a sample of size 18 (see Table 1), 90% confidence regions were constructed for a 20-year planning horizon (Klemeš, 1982a); they are represented by the shaded areas in Fig. 2. In all likelihood, these confidence regions are too narrow since their underlying stochastic model probably is grossly oversimplified. It is important to see that this cannot be corrected by any amount of mathematical-modelling legerdemain involving the given historic data since, because of the shortness of the record, there is no more information in them.

In this context, one problem should be emphasized which is not always appreciated. It is common knowledge that the statistical significance of a difference between two hypotheses, parameters, etc., drops as the underlying data base shrinks and the sampling uncertainty increases. This reduction of significance follows from the flattening of the sampling distribution and is graphically

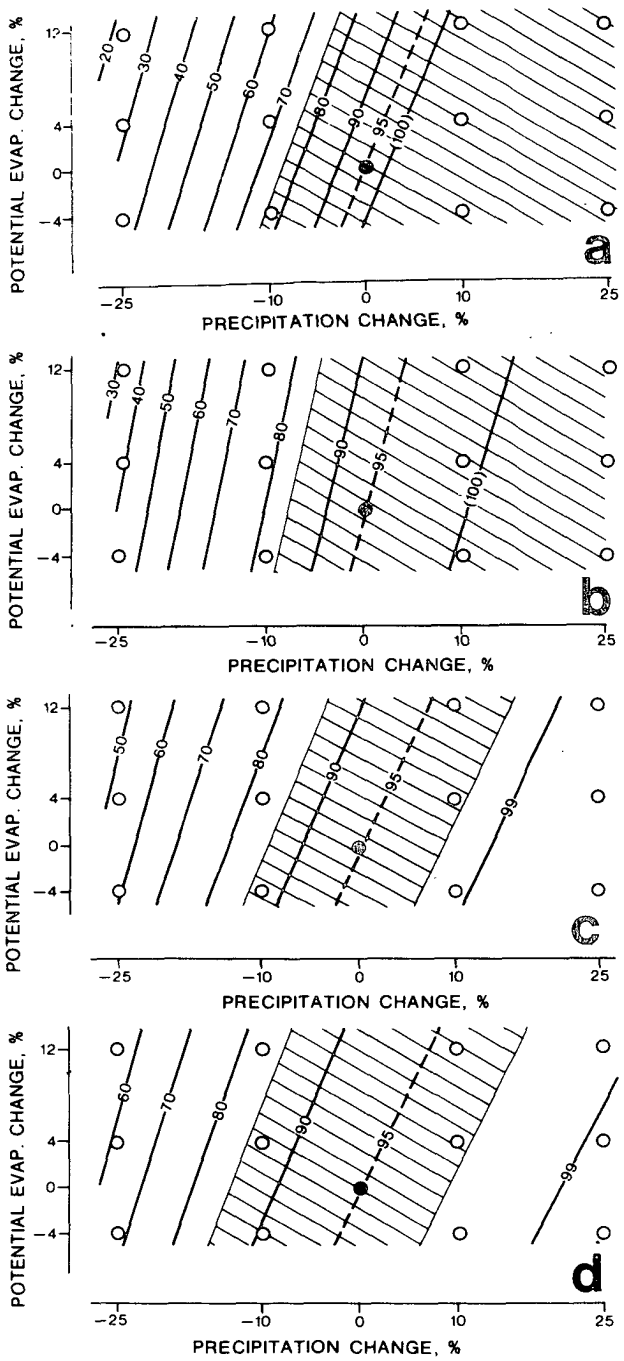


Fig. 2. Reliability (for Leaf River data) as a function of precipitation and potential evaporation changes, for the following 4 levels of development (all with an original 95% reliability):

Level	a	b	c	d
Target (m^3/s)	19.3	15.0	10.8	6.5
Storage (mill m^3)	600	280	125	40

Shaded areas represent 90% confidence regions (for the 95% reliability contour) computed on the basis of a stochastic model fitted to data, for sample size N=18 and a 20-year planning horizon.

reflected in the widening of the confidence interval. It thus seems to follow that, because of the shortness of the flow series underlying the results shown in Fig. 2, the shaded confidence regions are, if anything, too wide rather than too narrow as claimed in the preceding paragraph. The point is however that the above claim does not refer to the confidence region width as a function of sample size but rather to the width as a function of process structure for a given sample size. In other words, in the claim, the shortness of the record was invoked as a reason for our inability to see a greater complexity which is probably inherent in the process and which would lead to wider confidence regions for N=18 than those shown in Fig. 2.

Thus, ironically, if we are concerned with assessing statistical significance of climate change impact on the basis of a short record, the paucity of data may lead to an inflation of its statistical significance because of an underestimation of the complexity of the natural climate process. While it is true, as Němec and Schaake (1982) point out in quoting Katz (1980), that "...without reliable statistical inference, any climatic changes reportedly discovered could just as well be attributed to the chance variation of essentially unpredictable natural fluctuations", it must also be remembered that reliable statistical inference is not possible without a knowledge of the dynamic structure of the climate process which a short record does not reveal. As Bartlet put it as early as 1954, "unless the statistician has a well defined and realistic model of the actual process he is studying, his analysis is likely to be abortive" (Klemeš, 1978).

It follows that statistical and dynamic (physical) analyses are inseparable if they are to be meaningful and that neither can benefit from ignoring the other.

7. INSIGHTS THROUGH PROXY DATA

In cases where historic samples of data against which the effect of climate change is to be evaluated are too small and inadequate for reliable statistical inference, proxy data may provide some insights. In general, best results are to be expected from proxy data which are most closely related to those on which information is being sought (some dangers inherent in the use of remotely related proxy data will be discussed later). In our case, the first choice obviously is streamflow data from basins as similar as possible to the two basins under study (Table 1), but with longer records. The similarity criteria may vary and in practice will often by dictated by the information available. Our data source (Yevdjevich, 1963) called for the adoption of mean annual runoff depth and basin area. The basins selected on this basis are listed in Table 2.

In order to simplify the analysis, we refrained from assessing the reliabilities (as

Table 2. Characteristics of proxy river basins

	Annual runoff mm	Basin area km^2	Record length years
A. Leaf River Proxy Data			
Current R., Van Buren Missouri, U.S.A.	390	4320	45
Petit Jean Creek, Danville Arkansas, U.S.A.	394	1910	41
Kaweah, near Three Rivers, California, U.S.A.	380	1345	53
Feather, Boidwell Bar, California, U.S.A.	476	3500	45
Boise, Twin Springs, Idaho, U.S.A.	486	2150	45
Saugeen, Walkerton, Ontario, Canada	428	2200	40
Elbe, Děčín Czechoslovakia*)	186	51100	100
B. Pease River Proxy Data			
North Llano, near Junction Texas, U.S.A.	26	2380	42
Pecos, near Anton Chico, New Mexico, U.S.A.	48	2720	46
Verde, below Barlett Dam, Arizona, U.S.A.	43	15900	50
Humbolt, Palisade, Nevada, U.S.A.	25	13000	46
Avoca, Coonoorer Bridge Victoria, Australia	27	2600	59

*) Novotný (1963)

in Fig. 2) and concentrated on the annual runoff instead, in particular on its mean and standard deviation which represent the two most important parameters on which the reservoir regime function (and hence the reliability) depends. Within this framework, the objective is to compare these two streamflow parameters of the 12 climate-change scenarios with those of the historic flows. This comparison is shown in Figs. 3 and 4 where the four connected triads of full circles show the parameters of the 12 scenarios and the double circle those of the historical flow series (base). Both parameters are made dimensionless with respect to those of the base series which are set equal to unity. To be able to relate, at least approximately, Fig. 2 with Figs. 3 and 4, the latter figures contain ellipses representing 90% confidence regions for the true population parameters, on the assumption (see section 6) that the annual flows form random series.

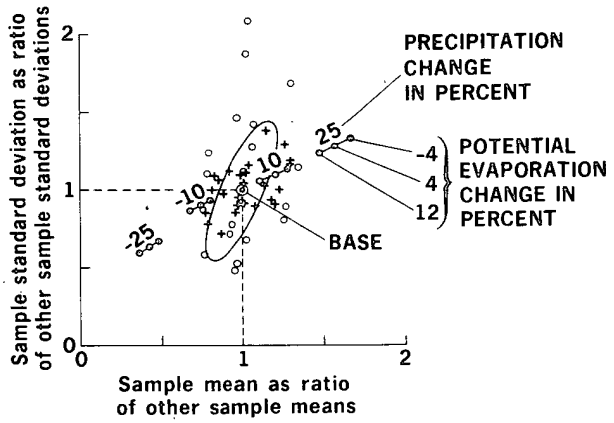


Fig. 3. Climate-change induced changes in sample parameters of annual flows for Leaf River (connected triads) compared to natural parameter variability in samples of equal size (N=18) from proxy basins listed in Table 2 (crosses; open circles correspond to Elbe River data).

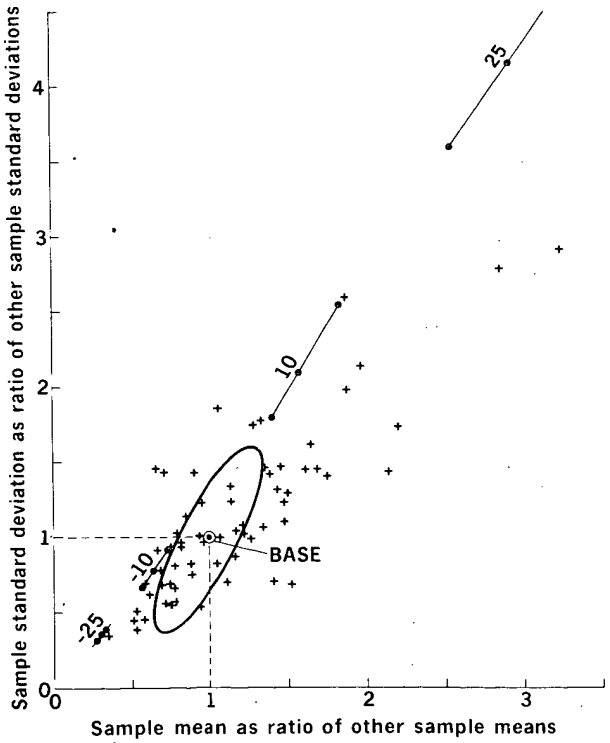


Fig. 4. Climate-change induced changes in sample parameters of annual flows for Pease River (connected triads - see legend in Fig. 3) compared to natural parameter variability in samples of equal size (N=11) from proxy basins listed in Table 2 (crosses).

If statistical significance of the differences between the "post-change" parameters and the "base" parameters were to be judged solely on the basis of primary data (i.e. without any reference to proxy data) it would have to be done as follows. Any point within the ellipse could be regarded as defining the population parameters and would serve as a centre of a "sample confidence" ellipse for parameters of samples of size equal to the planning period (if the planning period were of the same length as the historic period, this sample-confidence ellipse would be similar to the one shown). Those scenarios whose parameters would fall inside the sample confidence ellipses could then be considered statistically indistinguishable (at a given level of significance) from the historic period since their sample parameters would be compatible with the same population. In our two cases, this analysis leads to a conclusion that annual flows corresponding to a precipitation change of about ±10% would not be significantly different from those that might be expected under a stationary climate.

However, taking into account the proxy data shown in Table 2, the above conclusion seems rather conservative as indicated by the spread of the crosses in Figs. 3 and 4. These crosses were constructed in such a way that nonoverlapping segments of the same length as the base series (N=18 for the Leaf River, N=11 for the Pease River) were formed from the proxy series, the parameters of each segment were in turn set equal to one and those of the other segments expressed as their ratios. This representation shows the range of relative differences between compatible samples. The variability indicated by the crosses, while not associated with any specific confidence level, is much higher than that indicated by the procedure described in the preceding paragraph. The variability of the Pease River proxy data seems to be markedly higher than that for the Leaf River. This is suspected to be an artefact of the analysis and caused by the fact that the Pease River sample has size N=11 as compared to N=18 for the Leaf River. Because of this, more independent samples can be formed from a proxy series of a given length in the former case than in the latter so that the sampling variability of the latter case will be lower. To get a spread at the same level of significance for both basins, the Leaf River would require proxy series about 80-100 years long to produce the necessary number of independent samples. The most similar flow series of such a length that we could find was one for the Elbe River at Děčín in Czechoslovakia (Novotný, 1963). Its variability is believed to be of the same order as that of the Leaf River despite the differences both in mean annual runoff and in basin area (Table 2) since their effects would tend to compensate each other. Parameters of the Elbe River samples of size N=18 are shown by open circles in Fig. 3.

In general, the proxy data indicate that the inherent long-term variability of a stationary annual runoff series is higher than that of a random series and, consequently, that a blind statistical analysis of climate-change impact, based on short historic records, can be grossly in error. This conclusion is in agreement with results of some related earlier studies (Klemeš, 1979; Klemeš and Bulu, 1979).

8. RELEVANCE OF PROXY DATA AND OF THEIR ANALYSES

The main reason for the use of proxy data is to extend the length of record of a process under study in the hope that a longer proxy record will reveal important features of the process that are not apparent from a short primary record. While this objective is legitimate, the methods for its achievement are often inadequate and the inferences wrong. The most common inadequacy is an exclusive reliance on formal statistical methods both in the assessment of the closeness of the relation between the proxy and the primary data, and in the analysis of the proxy time series. Such an approach can lead to correct results only in exceptional circumstances, to quote Norbert Wiener, only if "the main elements of the dynamics of the situation are either explicitly known or implicitly felt" (Klemeš, 1978) as they are in our case. Here the main elements of the dynamics are relatively clear since both the primary and the proxy data are streamflows and most of the proxy data come from basins with similar climatic and hydrologic conditions. Moreover, and this is important, the information being sought is concerned only with relative statistical variability of successive short samples and definitely not with extrapolation of deterministic features (trends, periodicities, etc.) depending on real-time connectedness between the proxy and primary data.

Instances where the main elements of the dynamics are not so straightforward involve all the cases where the proxy data relate to a different physical process than the primary data; for instance when properties of a runoff process are being inferred from records of lake levels, tree rings, varves, etc. In such cases, the physical (deterministic) transfer mechanism between the two processes must first be identified and included in the analysis since it can cause drastic differences in the statistical properties of the two processes.

To illustrate the point, the use of lake level records as proxy data for streamflow will be discussed in more detail. In general, both processes tend to follow a similar long-term pattern in the sense that during long dry periods both lake levels and streamflows are lower than in long wet periods. However, this similarity can be grossly distorted in shorter periods. This has been repeatedly pointed out by hydrologists but routinely ignored by the modern "correlation hunters" mass-produced by typical contemporary graduate programs in hydrology. Thus, for instance, Langbein (1961) warned that, in using the fluctuations of lake levels as indexes of climatic variations, it must be remembered that "The interpretation of lake levels depends on the value of the response times, k. The longer the response time, the

greater is the possibility that climate and lake levels may be out of phase". More recently, Kurdin (1975, 1977) showed by an analysis of levels of Lake Balkhash that the dynamics of its water balance (in particular, the conditions in the delta of its tributary, the Ila River) may reduce or amplify long-term fluctuations of lake levels caused by climate variability.

The main reason why lake levels and streamflow may exhibit markedly different fluctuation patterns is that, in principle, lake level fluctuations reflect a running integral of the net precipitation process while streamflow may reflect either this process itself, or its running integral, or a mixture of both. For example, if the precipitation process were a random (serially uncorrelated) process, a lake level record would be highly serially correlated, streamflows in some rivers of the basin could be only weakly serially correlated while in other rivers they could exhibit high serial correlation. Moreover, in some rivers the streamflows could belong in the category of short-memory processes while in others they could form long-memory processes. This is because 1) lake levels reflect a storage process (equivalent to running integral) of the (net) precipitation process; 2) streamflow formed on an impervious part of the basin would have a structure similar to the precipitation process, with only a slightly stronger persistence (short memory); and 3) rivers fed from large lakes may have streamflow structure similar to that of the lake levels, i.e. belong in the long-memory category (Klemeš, 1974, 1978). Hence to use statistical analysis (stochastic model) of a lake level record for inferences about statistical properties of streamflow in all rivers in the area (or of rainfall, etc.) could be very misleading.

A specific example of this situation is the case of Lake Victoria levels. Their record shows a sudden rise of about 2.5 m around 1964, then a gradual drop of about 1 m followed, in 1978 and 1979, by a rise almost back to the 1964 level. A classical statistical analysis of the record would point to a nonstationary process, perhaps indicative of a climate change in the early sixties. However, precipitation records do not indicate any such change. This apparent discrepancy can be explained by the above-mentioned fact that lake levels behave, in general, like an integral of (net) precipitation. Precipitation which was quite high (and probably still underestimated) in the early sixties was capable of shifting the lake level to a new quasi-equilibrium. This was demonstrated by Kite (1981) with the aid of a deterministic water balance model for the lake. In a later, purely statistical, analysis of Lake Victoria levels, Kite (1982) found that the lake level record could not be homogenized without removing the above mentioned step and rightly concluded: "...if a [stochastic] model must include the possibility of random jumps of such a magnitude its usefulness for planning purposes is doubtful". The latter study also demonstrates the validity of point 3 above. It shows that the streamflow record of Victoria Nile (which is the Lake Victoria outflow) closely follows the lake level record including the sudden rise in the early sixties which does not appear in streamflows of rivers in the area not originating in Lake Victoria or in other large equatorial lakes.

9. CONCLUSIONS

The importance of an accurate assessment of the impact of climate change on the development of water resources must not be underestimated. But neither should be its difficulty which stems from the following:

a) uncertainty in quantitative estimates of change of primary climate variables which is mainly due to our imperfect understanding of climate dynamics;
b) the limited capability of current hydrological models which is a consequence of our imperfect understanding of hydrological mechanisms;
c) shortness of historic records which hinders the identification of the correct structure (and parameters) of the streamflow stochastic process which constitutes the basis of any assessment of a statistical significance of climate change impact;
d) inadequacy of current statistical methods for making useful inferences on the basis of small samples;
e) difficulties with information transfer from proxy variables in the absence of solid understanding of dynamic relationships between them and the primary variables, namely streamflow.

In short, the quality of climate impact modelling as well as that of the assessment of statistical significance of the impact depend primarily on our understanding of the physical mechanisms involved. With the present state of knowledge, and with the aid of the current tools reflecting it, only a very crude picture can be obtained whose usefulness to the planning process is extremely limited.

10. REFERENCES

Katz, R. W., 1980: Statistical evaluation of climate experiments with general circulation models: inference about means. Rep. No. 15, Climatic Research Inst., Oregon State Univ.

Kite, G.W., 1981: Recent changes in level of Lake Victoria. Hydrological Sciences Bull., 26(3), 233-243.

Kite, G. W., 1982: Analysis of Lake Victoria levels. Hydrological Sciences J., 27(2), 99-110.

Klemeš, V., 1974: The Hurst phenomenon: a puzzle? Water Resour. Res., 10(4), 675-688.

Klemeš, V., 1978: Physically based stochastic hydrologic analysis. Adv. Hydroscience, 11, 285-355.

Klemeš, V., 1979: The unreliability of reliability estimates of storage reservoir performance based on short streamflow records. In: Reliability in Water Resources Management, pp. 193-202, Water Resour. Publ., Fort Collins, Colorado.

Klemeš, V., 1982a: Effect of hypothetical climatic change on reliability of water supply from reservoirs. Report of National Hydrol. Res. Inst. of Environment Canada to WMO.

Klemeš, V., 1982b: The desirable degree of rigour in the testing of rainfall-runoff models. Amer. Geoph. Union, Fall Meeting, San Francisco.

Klemeš, V., 1983: Climatic change and the planning of water resource systems. Proc. 6th Canad. Hydrotech. Conference, pp. 485-500, Canad. Soc. Civil Eng., Ottawa.

Klemeš, V. and A. Bulu, 1979: Limited confidence in confidence limits derived by operational stochastic hydrologic models. J. Hydrology, 42, 9-22.

Kurdin, R.D., 1975: On long-term fluctuations of levels of large natural lakes in Kazakhstan and in Central Asia (in Russian). Proceedings, 4th All-Union Hydrological Congress, Vol. 5, Gidrometeoizdat, Leningrad.

Kurdin, R.D., 1977: On the role of the dynamics of the delta of the Ila river and of climatic factors in the fluctuations of water levels in lake Balkhash (in Russian). Water Resources, No. 1, USSR Academy of Sciences, Moskow.

Langbein, W.B., 1961: Salinity and hydrology of closed lakes. Geol. Survey Professional Paper 412, U.S. Government Printing Office, Washington, D.C.

Moran, P.A.P., 1957: The statistical treatment of flood flows. Eos, 38, 519-523.

Němec, J. and J. Schaake, 1982: Sensitivity of water resource systems to climate variation. Hydrological Sciences J., 27(3), 327-343.

NOVA, 1983: The pleasure of finding things out. WGBH Educational Foundation, Boston, Massachusetts.

Novotný, J., 1963: Hundred-year streamflow series from two Czech rivers (in Czech). Transactions of the Hydrometeorological Institute (HMU), No. 2, Prague.

Pittock, A.B., 1983: Solar variability, weather and climate: an update. Quat. J. R. Met. Soc., 109, 23-55.

Wallis, J.R. and P.E. O'Connell, 1973: Firm reservoir yield - how reliable are historic hydrological records? Hydrological Sciences Bull., 18(3). 347-365.

Williams, G.E., 1981: Sunspot periods in the late Precambrian glacial climate and solar-planetary relations. Nature, 291, 624-628.

Yevdjevich, V.M., 1963: Fluctuations of wet and dry years, Part I. Hydrology Papers, No. 1, Colorado State Univ., Fort Collins, Colorado.

CLIMATIC VARIATION IN SOUTH-EAST ENGLAND AND THE EFFECT OF SHORT-TERM ASSOCIATIONS BETWEEN WEATHER PARAMETERS ON EXPERIMENTATION WITH PERENNIAL CROPS

Ian G. Cumming

East Malling Research Station
East Malling, Maidstone, Kent, England. ME19 6BJ

1. INTRODUCTION

East Malling Research Station is the principal centre in Britain devoted to study of the culture and storage of fruit. It was founded in 1913 and is now part of the Agricultural Research Service, largely funded by the British Government.

Regular daily meteorological observations were started on 1st August 1914 and, although there has been considerable subsequent increase in the scope of recording, certain basic measurements have been made every day since that date. As field experiments with perennial crops may last upwards of ten years, various studies have been made of the effects of month to month and year to year variations in weather on the cropping and growth of fruit trees. This paper is primarily concerned with long-term characteristics of the weather itself and the effects which short-term patterns may have on such horticultural studies.

2. ANALYSIS OF 60 YEARS MONTHLY METEOROLOGICAL DATA

For the first analyses the monthly and annual values of rainfall, sunshine hours and mean air temperature for the period January 1915-December 1974 were used. This choice enabled the data conveniently to be subdivided into six periods of ten years in order to assess the long-term consistency of characteristics and relationships.

For each month of the year and for whole calendar years the means and standard deviations of rainfall, sunshine hours and temperature were calculated. Various tests, described below, for trends and periodicity were then carried out.

(i) Mean values, monthly and annual

Tables 1-3 show the variation between the sixty years of the monthly and annual values of the three parameters.

Rainfall was highest and also, when standard deviation was used as a measure, most variable in the autumn and early winter. The distributions within months were positively skew, but the distribution of annual totals was nearly symmetrical, since occasional severe storms tend to average out over twelve-month periods. The coefficient of variation did not appear to vary systematically through the year: it had an average value of 57% within a calendar month and 17% between years. These are similar to values previously reported at Kew, England (Ratcliffe, Weller and Collison, 1978) and elsewhere (Tabony, 1977).

Table 1. Monthly mean rainfall (mm) at East Malling, 1915-74.

Month	Mean	S.D.	Min.	Max.	C.V.
Jan.	60.1	27.0	17.3	132.6	45
Feb.	47.3	31.2	2.3	149.9	66
March	40.7	27.6	0.3	126.7	68
April	47.7	23.3	3.8	121.1	49
May	45.5	20.9	11.2	102.4	46
June	43.6	27.9	3.3	125.7	64
July	55.0	27.2	8.9	121.2	49
Aug.	58.9	29.3	0.3	127.8	50
Sept.	58.5	36.8	2.8	183.9	63
Oct.	65.9	45.9	2.5	211.1	70
Nov.	77.0	41.7	17.3	191.0	54
Dec.	65.6	38.0	10.2	153.2	58
Total	665.7	113.7	329.7	948.4	17

S.D. = Standard Deviation

C.V. = Coefficient of Variation (%)

Sunshine hours were naturally highest in midsummer, but were also most variable at that season when the standard deviation was used to measure variability. The coefficient of variation, however, was least in summer. Within calendar months the distributions were approximately symmetrical.

Table 2. Monthly mean sunshine hours.

Month	Mean	S.D.	Min.	Max.	C.V.
Jan.	47	14	23	85	30
Feb.	67	21	23	121	31
March	119	32	50	197	27
April	145	37	65	226	26
May	202	33	121	257	16
June	210	41	83	280	20
July	197	45	108	305	23
Aug.	185	33	123	281	18
Sept.	149	33	89	220	22
Oct.	107	24	70	169	22
Nov.	59	17	30	106	28
Dec.	42	15	12	89	37
Total	1528	154	1257	1935	10

Table 3. Monthly mean temperature ^{o}F.

Month	Mean	S.D.	Min.	Max.
Jan.	38.8	3.5	28.2	46
Feb.	39.2	3.9	29	46.2
March	42.8	2.8	37.3	49.8
April	47.0	2.0	41.3	52
May	52.9	1.9	48	57
June	58.2	1.9	53.4	62.4
July	61.5	2.1	58	67.7
Aug.	61.1	1.8	57.2	65.2
Sept.	57.2	2.2	51	63
Oct.	50.6	2.5	45.9	58.5
Nov.	43.7	2.4	38	49.7
Dec.	40.1	3.2	32.1	48.3
Total	49.4	0.9	47.4	52.3

The highest mean air temperature was in July, but the winter months were much more variable. Again, the monthly distributions were approximately symmetrical. As the temperature scales on which meteorological observations are made are not absolute, the coefficient of variation has no particular significance. In this study Fahrenheit was used, partly because for most of the years and sites this scale had been used in recording until comparatively recent years and also because of the small differences which were of interest.

Sunshine hours and mean temperatures when plotted against successive months showed the well-established sinusoidal variation - see e.g. Trowell (1975) for illustrations. The seasonal variation of rainfall was less obvious, but also tended to a sinusoidal shape.

(ii) Correlation between calendar months

The correlation matrices for each parameter between months within calendar years were calculated and are shown in Tables 4.-6. It should be noted that the correlations between January and December are for months which are eleven months apart. The closer a result is to the diagonal of the matrix, the nearer together in real time are the two months to which it refers. The relevant statistically significant values of the correlation coefficient are 0.25 at $P < 0.05$, 0.32 at $P < 0.01$ and 0.41 at $P < 0.001$. The values which are significant at at least $P < 0.05$ are underlined.

Of the sixty-six coefficients in the rainfall matrix, only three were significant at $P < 0.05$, and in two of these cases the two months were a long way apart. Clearly there was very little evidence of correlation in rainfall between one month and another, at any time of the year. Holland (1962) obtained comparable results in a differently oriented study covering only the growing season for fruit trees.

With sunshine hours the position was quite different. Eight of the coefficients were significant at $P < 0.05$ and three of these were significant at $P < 0.01$. Two areas in the matrix were of particular interest. Between the months of July, August and September was a group of three significant coefficients, the highest being 0.47. This confirmed the popular impression that long, sunny summers and cold, dull summers are real, if occasional, phenomena in south-east England. For example, 1959 was consistently sunny over this period, whereas 1968 was consistently overcast. The other period showing high correlation was October/November - a situation which brings to mind the American expression an Indian Summer.

With mean temperatures seven coefficients were significant at $P < 0.05$ and four of these at $P < 0.01$. These were between adjacent pairs either in winter or in summer, namely - January/February; February/March; May/June and July/August.

Separately calculated for adjacent December/January months, the correlation coefficients were +0.30 for temperature and +0.12 for sunshine hours.

These calculations showed that in this area of England there were at certain times of the year tendencies for weather conditions to persist from one month to the next.

Table 4

Correlation Matrix of Monthly Mean Rainfall

Jan.	1.00											
Feb.	0.02	1.00										
March	0.06	0.13	1.00									
April	-0.16	0.13	-0.05	1.00								
May	-0.07	0.06	0.04	-0.04	1.00							
June	0.11	0.10	0.14	0.10	-0.04	1.00						
July	0.05	0.07	0.03	-0.07	-0.21	-0.03	1.00					
Aug.	0.01	0.22	-0.04	-0.01	0.04	0.05	0.12	1.00				
Sept.	-0.05	0.18	-0.15	-0.05	-0.07	0.00	0.18	0.25	1.00			
Oct.	-0.05	-0.03	-0.29	-0.04	0.23	-0.11	-0.14	-0.01	-0.11	1.00		
Nov.	-0.14	0.26	0.03	0.15	-0.22	-0.17	0.01	0.20	-0.12	0.04	1.00	
Dec.	0.23	0.11	0.11	-0.13	-0.13	0.02	0.09	0.03	0.01	-0.09	-0.08	1.00
	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.

Table 5

Correlation Matrix of Monthly Mean Sunshine Hours

Jan.	1.00											
Feb.	-0.03	1.00										
March	0.08	-0.01	1.00									
April	0.02	0.13	-0.12	1.00								
May	0.02	0.15	-0.07	0.17	1.00							
June	0.17	0.20	0.08	0.23	0.05	1.00						
July	0.15	0.20	0.06	0.25	0.05	0.24	1.00					
Aug.	0.16	-0.03	0.04	0.05	-0.02	0.26	0.29	1.00				
Sept.	0.06	0.06	-0.04	0.01	0.16	0.07	0.40	0.47	1.00			
Oct.	0.17	0.11	0.14	0.09	0.07	0.22	0.15	1.12	0.14	1.00		
Nov.	0.10	0.01	0.09	0.00	0.10	0.04	-0.06	0.00	0.16	0.38	1.00	
Dec.	0.09	0.07	0.30	0.12	0.01	0.31	-0.05	-0.01	-0.06	0.10	-0.04	1.00
	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.

Table 6

Correlation Matrix of Monthly Mean Temperature

Jan.	1.00											
Feb.	0.38	1.00										
March	0.11	0.34	1.00									
April	0.00	0.31	0.18	1.00								
May	0.00	-0.09	-0.01	-0.08	1.00							
June	-0.21	-0.04	0.08	-0.11	0.34	1.00						
July	0.06	0.09	-0.04	0.18	-0.13	-0.03	1.00					
Aug.	0.11	-0.08	-0.03	0.30	0.05	0.02	0.33	1.00				
Sept.	-0.03	-0.04	0.02	-0.03	-0.07	0.10	0.22	0.27	1.00			
Oct.	0.17	0.00	-0.01	0.13	-0.15	-0.08	0.23	0.17	0.23	1.00		
Nov.	-0.11	0.08	-0.11	-0.09	-0.20	0.03	-0.12	-0.17	0.21	-0.11	1.00	
Dec.	0.04	-0.09	-0.03	-0.14	0.20	-0.11	-0.12	-0.07	0.17	-0.02	-0.07	1.00
	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.

(iii) Long-term drifts

In a simple examination for long-term drifts or trends, the data for each variate were divided into six decades. Between and within decade analyses of variance were calculated, but among the thirty-six analyses of monthly data, only one produced a significant result. During the period 1945-54 the April sunshine hours were significantly higher than in preceding or following decades ($P < 0.001$). As a direct consequence of this the annual sunshine hours were also high during this decade.

The fact that in the twenty years 1945-64 there was a gradually decreasing amount of April sunshine will be referred to again later but the situation is well demonstrated in Fig. 1.

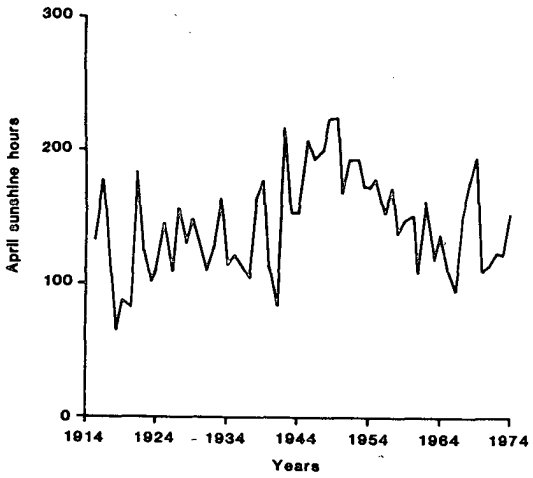


Fig. 1. April sunshine hours at East Malling 1915-74.

(iv) Periodicity in annual data

In order to assess periodicity in the annual data, the autocorrelogram was obtained for the three variables with lags of up to twenty years. Of the sixty correlation coefficients so obtained, only one was significant at $P < 0.05$, namely for rainfall with a lag of six years for which the coefficient was -0.30. This showed that within the rather limited scale of this data set there was little evidence of cyclic variation.

3. SHORT-TERM ASSOCIATIONS BETWEEN CLIMATIC VARIATES

Against this background of variation over sixty years it was then feasible to examine shorter-term effects which have or could have given rise to problems in the interpretation of relationships between weather and biological activity. In many experiments with perennial crop data is collected over ten years or more and for the present purpose this was the time scale of interest. Three illustrations are considered below.

(i) April sunshine hours

Referring again to Fig. 1. the unusual pattern of April sunshine hours was examined further. In the period 1945-64 April sunshine gradually decreased from a 5-year average of 210 hours to 130 hours. Expressed as a regression this situation was represented by

$$SH = 222 - 4.9 \text{ (Year A.D. - 1944)}$$

with a correlation coefficient of -0.88 ($P < 0.001$).

Persistence of this relationship would of course, have had disastrous effects on the climate culminating in negative sunshine in 1989! However, over the entire sixty years the coefficient had the statistically insignificant value +0.19. This particular example came to light during analysis of the weather data itself rather than incidentally in a biological experiment. However, as a real if perhaps extreme situation, it illustrated well how a spurious conclusion could have been drawn from a long-term experiment. If data had been collected from a perennial plant during the period 1945-64, or a sub-set of these years, of a variate which depended biologically on April sunshine, any significant effect might well have erroneously been ascribed to the age of the plant. That this is not sanctimonious pedantry should be more apparent in the succeeding examples.

(ii) May sunshine hours

In a previous investigation, Barlow and Cumming (1975) examined the effect of late spring weather on the yield of an established apple orchard. Yield was found to be positively and significantly correlated with sunshine recorded in the first two weeks of May when pollination is normally taking place but negatively in the later weeks of the month. It was observed that this rather surprising result may have indicated a genuine plant response, although a biological explanation was not obvious. It may merely have reflected a tendency for there to be a

negative correlation during the thirty-one years of this study between the sunshine hours early and late in May.

The correlation matrices of sunshine hours between the five weeks starting in May (up to 4th June) are shown in Table 7. for the thirty-one years concerned and also for the complete sixty years. Over the longer period the inconvenient cluster of negative correlations was no longer apparent. It is relevant, however, to remark that fruit trees are not normally grown for sixty years in modern orchard practice.

Table 7. Correlation matrices of sunshine hours in 5 weeks from 1st May.

(a) 31 years of Barlow and Cumming.

Weeks					
1	1.00				
2	-0.05	1.00			
3	0.07	0.20	1.00		
4	0.12	-0.34	-0.06	1.00	
5	0.01	-0.29	-0.26	0.06	1.00
	1	2	3	4	5

(r = 0.35 for P < 0.05, 29 D.F.)

(b) All years 1915-74

1	1.00				
2	0.09	1.00			
3	-0.19	0.08	1.00		
4	0.07	-0.01	0.03	1.00	
5	-0.05	-0.09	-0.03	0.18	1.00
	1	2	3	4	5

(iii) Mean temperature in April and June

In Table 6. those correlation coefficients between mean monthly temperatures which reach the P < 0.01 level of significance relate to adjacent months and are therefore close to the diagonal of the matrix. The further apart two months are in time, the less likely is the correlation of temperatures between them to reach a significant level.

However, in the course of studies of the effects of weather on national yields of apple orchards in England it has been observed that over the period 1951-65 the correlation between April and June mean temperatures reached a value of +0.63 (P < 0.01) (Jackson and Hamer, 1980). This compared with the non-significant value of -0.11 over the entire sixty years.

The situation is similar to that in the previous example. There is a risk that during this era for which recorded yields were available a biological effect due to warm (or cold) weather in April might have been erroneously ascribed to conditions in June, and vice versa.

It may be possible in studies of this type to separate the effects of correlated variables by multiple regression, but a series of consecutive years is not a random selection of values of an independent variable. A conclusion based on a series showing inter-year correlation may well have to be modified over a longer period of years. The examples are intended to illustrate a problem which may arise when year to year variations in an annually recorded biological parameter are assessed in relation to the characteristics of the weather.

4. ASSOCIATION BETWEEN ANNUAL VALUES OF CLIMATIC VARIATES

As annual values of rainfall, sunshine hours and mean air temperature were, at the time of writing, available for 67 consecutive years from this Station the correlation coefficients between these values were calculated with the following results:-

Rainfall x sunshine	-0.37	(P < 0.01)
Rainfall x temperature	-0.30	(P < 0.01)
Temperature x sunshine	+0.26	(P < 0.05)

These coefficients, although statistically significant, are not high. They do, however, show the obvious tendency for wet years to be cold and sunless and for sunny years also to be warm.

5. COMPARISON WITH OTHER AREAS IN BRITAIN

The geographical spreads of certain of the short-term associations and of annual correlations were then examined.

(i) April sunshine hours

The total April sunshine hours for the years 1945-64 for various stations covering the area of Britain were extracted from the Monthly Weather Report of the Meteorological Office. The sites chosen were a) Lerwick (Shetland), b) Stornoway (Western Isles), c) Duntuilm, Skye (Highland), d) Craibstone, Aberdeen (Grampian), e) Blackford, Edinburgh (Lothian), f) Auchincruive, Ayr (Strathclyde), g) Cockle Park, Morpeth (Northumberland), h) York (North Yorkshire), i) Sutton Bonington (Leicestershire), j) Harpenden (Hertfordshire), k) Long Ashton (Avon), l) Valley, Anglesey (Gwynedd), m) East Malling (Kent).

For each station the correlation coefficient of April sunshine hours with calendar years was calculated and these are shown in Fig. 2.

Using the negative correlation coefficient as a measure of the downward trend it is clear that this was most pronounced in the south-east, decreased with distance from London and disappeared completely in the furthest northern and western areas. Although such data could be analysed by more sophisticated methods this simple analysis was sufficient to demonstrate the geographical variation.

(ii) Mean temperatures in April and June

The mean temperatures in April and June were abstracted for the period 1951-65 at the same stations. The correlation coefficients between these two temperatures were calculated and are shown in Fig. 3. The situation is rather similar to that for April sunshine hours. In the extreme northern and western areas there was no evidence of a temporary association between April and June temperatures during 1951-65. On the other hand in the south and east, which incidentally includes most fruit growing areas in Britain, the correlation was significant and of a magnitude to be a nuisance in experimentation with perennial crops.

Over longer periods of years the association generally disappeared.

(iii) Annual values

A full analysis of the annual data from different stations is not presented here but a number of aspects are noted where they are relevant to the rest of the paper.

The correlation coefficients between annual values of rainfall, temperature and sunshine hours are shown in Fig. 4. The number of years of available data varied greatly from one station to another but was always more than fifty. The correlation coefficients were calculated for the maximum number of years in all cases.

Although often statistically significant, the numerical values of the correlation coefficients were generally small. The negative association between rainfall and mean temperature prevalent in the south and east was not evident in the north-west areas, which are subject to the Gulf Stream and relatively higher rainfall.

The association between rainfall and sunshine was everywhere negative, as was to be expected. Temperature and sunshine were positively correlated over most of Britain, again excepting the north-west, where the coefficients were insignificant, taking small negative values.

Although there was little indication of trends in temperature at East Malling, over much of Britain the mean temperatures in the 1920's were about 1°F lower than in the next thirty-five years. This phenomenon is, of course, well known - see e.g. Gribbin (1978).

The variability in annual data at some stations for which very long runs of data were available is shown in Figs. 5, 6 and 7. The patterns are fairly well defined and for rainfall, at least, agree with Tabony (loc. cit.).

Rainfall data were available from Stornoway for over one hundred years and the monthly values varied in a continuously sinusoidal manner. The coefficient of variation was somewhat higher in the spring months than in the rest of the year.

6. REFERENCES

Barlow, H.W.B., and I.G. Cumming, 1975: Effect of early summer weather on the yield of Lane's Prince Albert apple. Climate and the Orchard. ed. H.C. Pereira. Commonwealth Agricultural Bureau, Farnham Royal, U.K., 44-47.

Gribbin, P., 1978: Climatic change. Cambridge University Press.

Holland, D.A., 1962: The prediction of monthly rainfall as exemplified by data from south-east England. J. Agric. Science, 58, 327-331.

Jackson, J.E. and P.J.C. Hamer, 1980: The causes of year to year variation in the average yield of Cox's Orange Pippin apple in England. J. Hortic. Science, 55, 149-156.

Ratcliffe, R.A.S., J. Weller and P. Collison, 1978: Variability in the frequency of unusual weather over approximately the last century. Quart. J. of the Royal Meteorological Society, 104, 243-256.

Tabony, R.C. 1977: The variability of long duration rainfall over Britain. Scientific Paper No. 37, Meteorological Office.

Trowell, G.F., 1975: Forty years of weather at East Malling. Climate and the Orchard, 10-12.

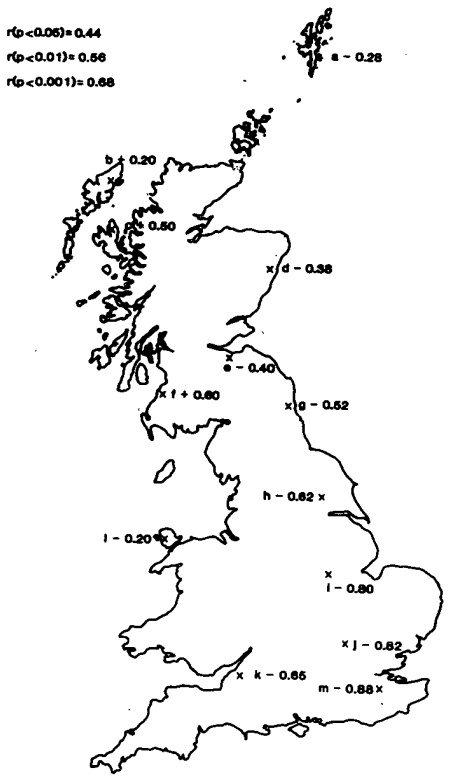


Fig. 2. Correlation coefficient April sunshine hours/calendar years 1945-64.

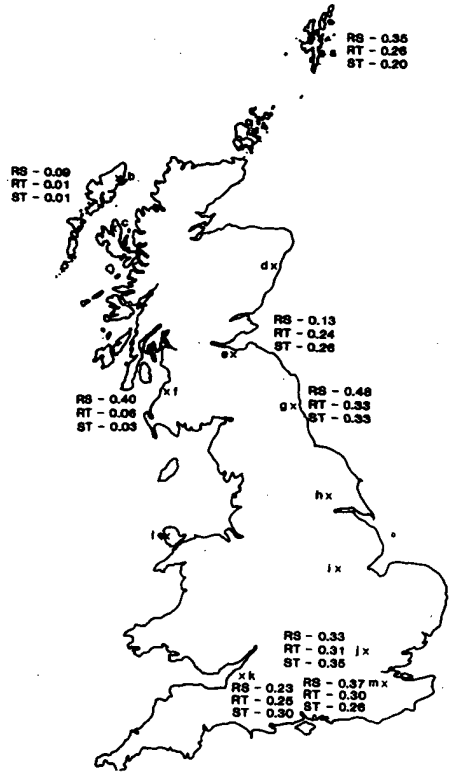


Fig. 4. Correlation coefficients between annual rainfall (R) sunshine hours (S) and mean air temperature (T).

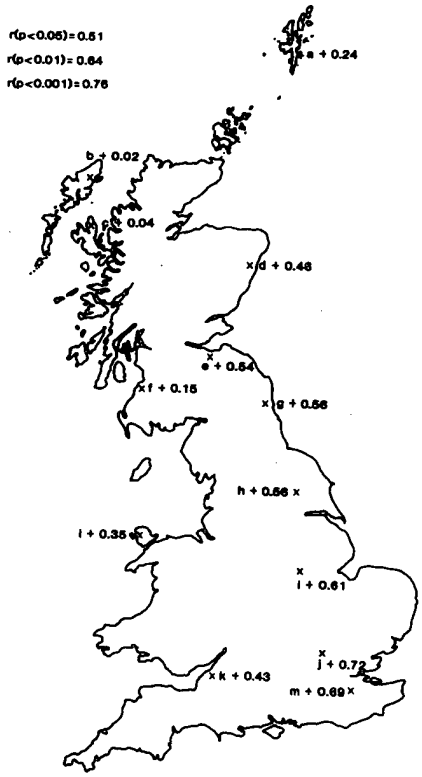


Fig. 3. Correlation coefficient April/ June mean temperatures 1951-65.

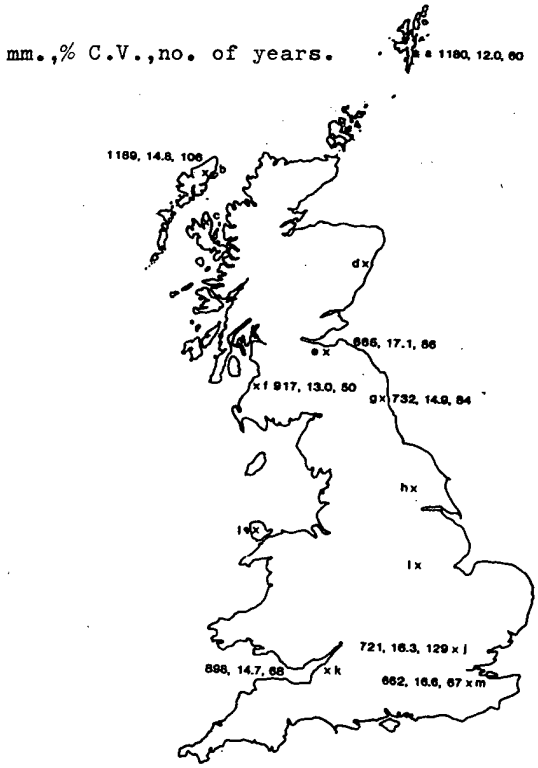


Fig. 5. Variability in annual rainfall.

hours,% C.V.,no. of years.

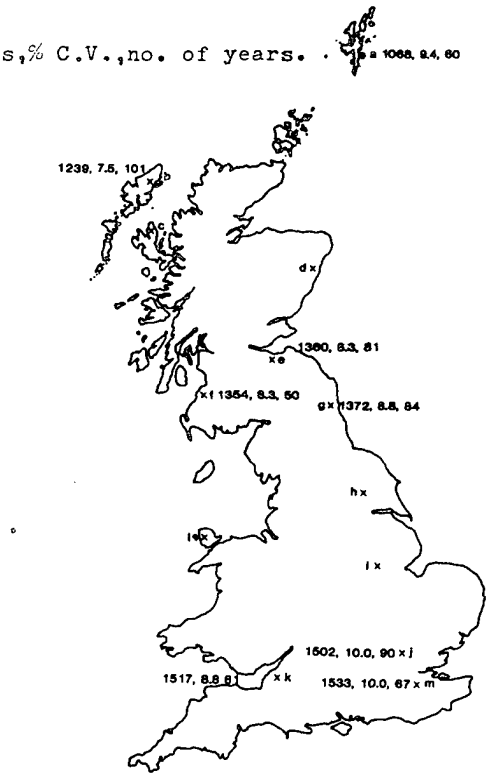


Fig. 6. Variability in annual sunshine hours.

°F, S.D.°F, no.of years.

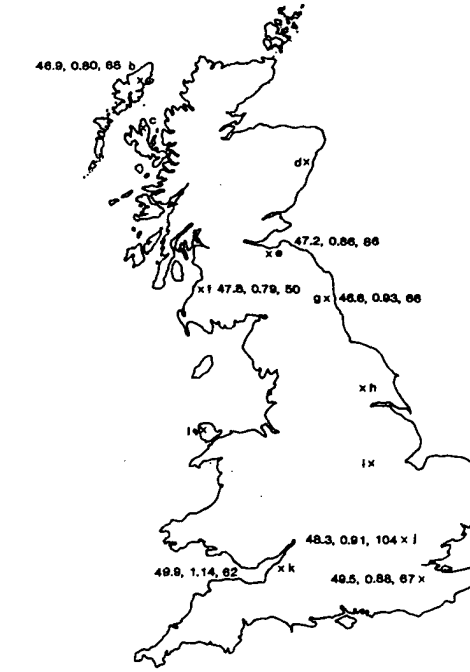


Fig. 7. Variability in annual mean air temperature.

CLIMATE - TRANSPORT RELATIONSHIPS IN GREAT BRITAIN

J.P. Palutikof

Climatic Research Unit,
University of East Anglia,
Norwich NR4 7TJ, U.K.

1. INTRODUCTION

It is widely recognized that the weather can cause extensive disruption of the economy, in terms of both costs and losses. Aspects of the economy where such vulnerability can be argued to exist include energy consumption (Tellings, 1978), agriculture (see, for example, Doornkamp et al, 1980), and transport. This paper examines possible causal relationships between winter weather (and climate) and a number of transport measures for Great Britain.

In an intuitive sense it is clear that weather, and thus more generally climate, and transport must be related. In this respect we are referring to transport indicators such as distances travelled and number of journeys made. Although climate-transport links may be argued to exist at all times of the year, their existence appears most likely in a temperate-latitude country such as Great Britain in the winter season. A particularly severe winter, such as that which occurred in Britain in 1962-3, causes disruption of traffic flows because of, for example, snowstorms, blocked roads and frozen points on the railway network (Codling, 1971; Hay and Young, 1972). The number of road accidents is also related to the severity of winter weather (Codling, 1974). In a more indirect sense, the known relationship between health and weather (Landsberg, 1972; Tromp, 1980) may cause fewer journeys to be made in a harsh winter.

In this paper an attempt is made to quantify the extent of the impact of winter weather on transport. Time series of three transport measures are compared with both extreme winter seasons and with more general departures from 'average' winter weather conditions. In fact, since extremes are a small sample drawn from general weather conditions, the two cannot be regarded as independent. To remove the undue influence of extremes in the comparison of general weather conditions with transport, the analysis was performed on the complete time series, and on the same time series with the extreme events removed.

2. THE TRANSPORT MEASURES

The three measures of transport used in this analysis are:

i) Index of road vehicle miles travelled, all vehicles (1958-79);
ii) Index of road vehicle miles travelled, commercial vehicles (1958-79); and
iii) Number of road casualties (1958-79).

Monthly values for these three measures are published in the Monthly Digest of Statistics (CSO, 1958-79).

Estimates of distances travelled are derived from roadside traffic counts. These counts take two forms: occasional short surveys at large numbers of sites to estimate the absolute level of traffic, and frequent counts at a small number of sites to estimate changes in the amount of traffic. The results are expressed as an index value in order to facilitate year-to-year comparisons. The base-year for the index is periodically changed: for example, in 1960 the index was to base 1958 = 100, whereas in 1970 it was to base 1963 = 100. However, since over the period 1958-79 the method of measurement was not changed, it is possible to use the overlap periods to arrive at a series with a common base. The index of road vehicle miles travelled for all vehicles is self-explanatory, in that it refers to all traffic. The index for commercial vehicles applies only to those over 30 cwt (1,530 kg) unladen weight, plus miscellaneous vehicles such as tractors, ambulances and road rollers.

Statistics on road casualties are compiled from police reports. Since there is no strict obligation to report an accident to the police a small proportion of minor injury accidents will be missing from the statistics. The figure used in this analysis is for the total number of casualties on roads in Great Britain irrespective of age of victim, nature of casualty or class of road user.

For further information on the derivation of these measures, reference should be made to Transport Statistics in Great Britain 1970-80 (Department of Transport et al, 1981). For all three, a 22-year series (to base 1958 twelve-month

mean = 100 for the index values) was derived.

3. THE IMPACT OF INDIVIDUAL EXTREME WINTERS

Within the period of the derived time series there are two winters which are usually judged to have been severe: 1962-3 and 1978-9. For many areas of England and Wales the 1962-3 winter was the coldest on record (Shellard, 1968). Table 1 shows the central England temperatures (Manley, 1974, updated by the U.K. Meteorological Office) for the winter months. All three months were the coldest of their respective time series within the 1958-79 period, although in the case of December, 1976 was equally cold. Heavy snowfalls occurred over most of the country between 12 December and 23 February. The average number of days with snow falling at ten stations in Great Britain was for this winter 44, a value only exceeded in the 1958-79 period by the winters of 1969-70 (50 days) and 1978-9 (63 days). However, because of the severity of the snowfalls and the low temperatures, the 1962-3 winter had the greatest recorded number of days with snow lying (U.K. Meteorological Office, 1979; Table 1). The 1978-9 winter was for most parts of the country the coldest since 1962-3 and the second coldest this century. Unlike the winter of 1962-3, severe conditions did not set in until January. December was reasonably mild: seven Decembers in the 1958-79 period had lower values of CET. Snowfall was about average in most places. Conditions deteriorated in early January and the weather continued cold with frequent snowfalls throughout the winter months.

Table 1. Characteristics of severe winters (for explanation and sources, see text).

	CET (°C)			No. of days snow falling	No. of days snow lying
	Dec.	Jan.	Feb.		
1962-3	1.8	-2.1	-0.7	44	69
1978-9	3.7	-0.4	1.0	63	57
Mean, 1958-79	4.4	3.6	3.8	31	20

In order to determine whether vehicle mileage was substantially reduced in these two severe winters, index values for each winter month were compared with the index mean for the preceding autumn (September, October, November) for each year in the time series. As an example, the index for all vehicles in January 1963 is 104, and the 1962 autumn mean is 127, giving a difference (winter month minus autumn mean) of -23. When this is done for every year, the results show whether the change between the preceding autumn and a severe winter is substantially different from that when winter conditions are closer to the average. The advantage of this method is that it minimizes the influence of long-term trends in the statistics. The all-vehicle index shows a pronounced upward trend between 1958 and 1979; for commercial vehicles there is a less-marked upward trend from 1958 to 1972 after which the curve flattens off.

Figure 1 shows the results of this analysis for the index of all vehicle miles travelled, for the three winter months. Bearing in mind the conditions described in the first paragraph of this section, the results fit well with what would be expected if we assume that the all-vehicle

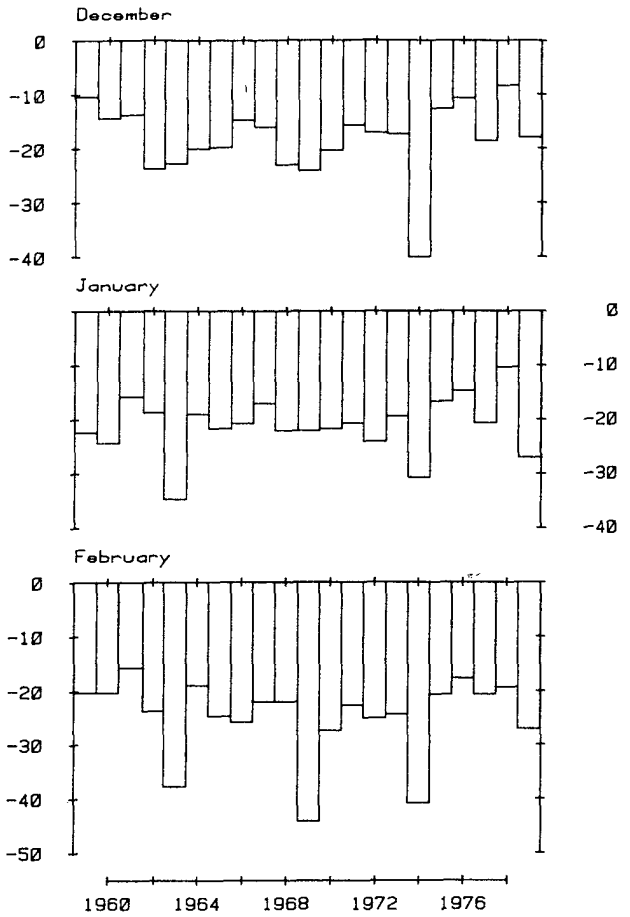


Fig. 1. Winter month Index of All-vehicle Mileage compared to preceding autumn mean, 1959-79.

index is influenced by severe winter weather. In the 1962-3 winter, when conditions were harsh throughout, index values in all three months were also substantially lower than the average. For the month of most severe weather, January, the index value for 1963 when compared with the preceding autumn is the lowest January value in the 1959-79 time series. February 1963 has the third lowest value for that month, and December 1962 the fifth. The situation for the 1978-9 winter, when severe conditions did not set in until January, is rather different. The December minus autumn value was near to the median, being the ninth lowest of all Decembers between 1959 and 1979. The January and February 1979 differences, however, are ranked third and fourth lowest respectively.

The evidence suggests that there is a substantial impact on all-vehicle mileage during a severe winter. In an attempt to discover what categories of vehicle are most affected, the same analysis was performed on the index of vehicle miles travelled for commercial vehicles. The results are illustrated in Figure 2. The winter of 1962-3 apparently had much less effect on commercial-vehicle mileage than it did on all-vehicle mileage. The ranked values for December 1962 and February 1963 are the same for the two indices: fifth and third lowest respectively. However, the autumn minus January difference for commercial vehicles is only the eighth lowest January value in the 1959-79 series, indicating

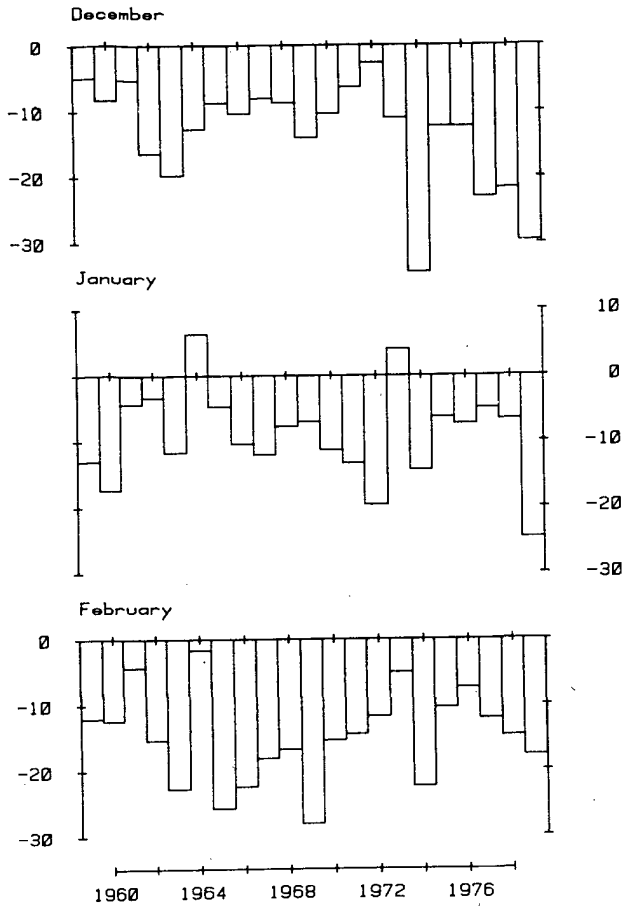


Fig. 2. Winter month Index of Commercial-vehicle Mileage compared to preceding autumn mean, 1959-79.

that mileage was not substantially rest icted in this month. The diagrams in Figure 2 suggest that the impact in the 1978-9 winter was much greater. The ranked orders for the three months are, from the lowest in the time series, December 1978 second, January 1979 first and February 1979 seventh.

These results for commercial-vehicle mileage do not agree well with what would be expected if the effects were weather-induced. The autumn minus January 1963 difference is not low enough, and the December 1978 value is too low, given that the severe weather did not begin until January 1979. There is one further factor which must be considered as a possible influence upon commercial-vehicle mileage, and that is the state of industry. Clearly, if industrial production is depressed by the influence of, for example, industrial disputes and strikes, then this in turn may be reflected in commercial-vehicle mileage. An example of this effect is shown in Figures 1 and 2 by the low index values throughout the winter of 1973-4. A coal miners' ban on overtime working throughout the latter half of 1973 finally put the country's workforce on to a three-day working week from January to March 1974. Palutikof (1983) devised a measure of industrial disputes named the Industrial Disputes Index (IDI). Values of the IDI were very high from October 1978 until the end of winter: October and December 1978 and January 1979 had the highest IDI values of their respective months in the 1958-79 time series; November 1978 had the third highest November IDI value. Allowing for lag effects, the large number of days lost to industry through strikes may well have influenced the 1978-9 winter commercial-vehicle mileage (and thus the all-vehicle mileage also). The relationship between this index and industrial production is investigated more fully by multiple regression analysis in the next section.

The results presented so far indicate that, whereas all-vehicle mileage is substantially reduced in harsh winter conditions, commercial-vehicle mileage is more resilient. The implication is that in bad weather private journeys are likely to be cancelled completely, but that commercial-vehicle journeys can only be postponed. This accords well with what might be expected.

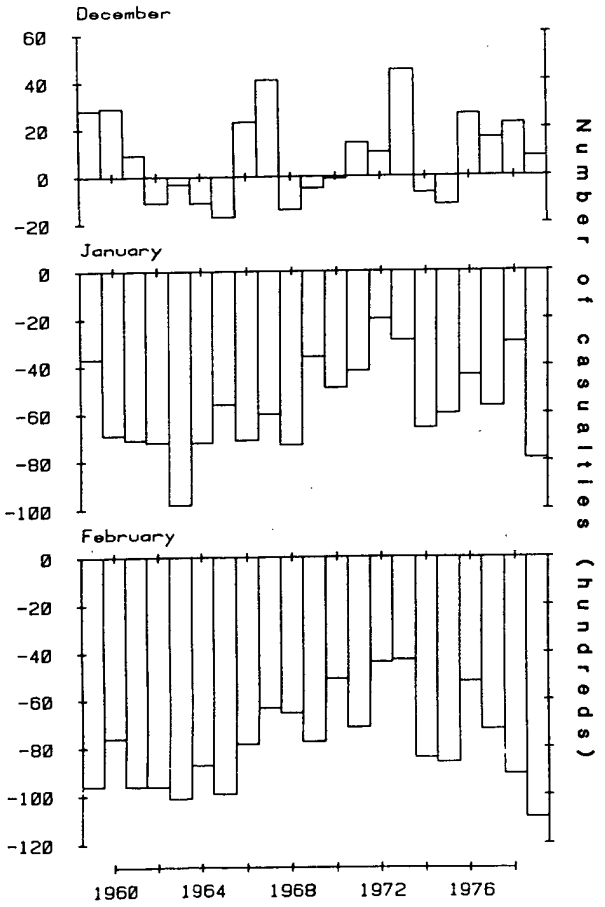


Fig. 3. Winter month number of road casualties compared to preceding autumn mean, 1959-79.

Finally in this section, an analysis was made of the relationship between severe winter weather and the number of road casualties. The autumn mean minus winter month differences for the 1958-79 time series are shown in Figure 3. The December 1962 difference fails to show any influence of the severe weather, but the accident figures for this month are complicated by the effects of the long Christmas vacation, accompanied by a sharp increase in the number of arrests for drunken

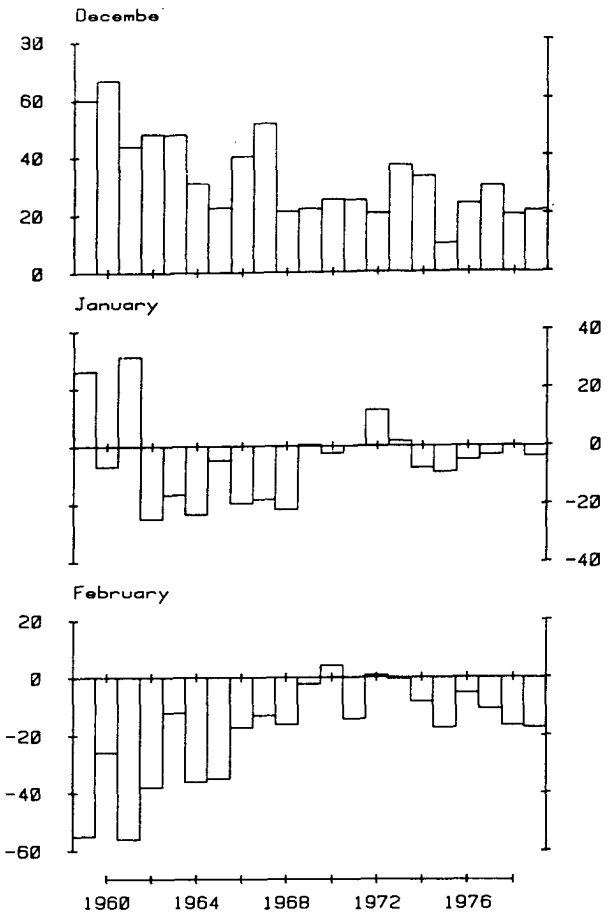


Fig. 4. Winter month number of road casualties controlled for all-vehicle mileage compared to preceding autumn mean, 1959-79. Units: casualties in hundreds (see Fig. 3) divided by mileage index (see Fig. 1).

driving. January and February in both 1963 and 1979 had a very low number of road casualties when compared with the preceding autumn so one might conclude that severe weather reduces road casualties. However, given that vehicle mileages were also relatively low in these months, it is necessary to control for distances travelled. This was done by dividing the road casualty number for each month by the index of all-vehicle mileage for the same month before repeating the autumn mean/winter month comparison. Figure 4 shows these results. It is clear from these diagrams that the severe weather of the 1962-3 and 1978-9 winters had no effect on the number of road casualties per unit distance travelled. This would suggest that the low total number of casualties in harsh winters arises simply because people do not travel, rather than because they drive more carefully.

4. THE IMPACT OF GENERAL WEATHER CONDITIONS

As already shown, the index of all-vehicle mileage and the absolute number of road casualties are strongly affected by severe winter conditions. Little or no relationship appears to exist for commercial-vehicle mileage or for the number of casualties per unit distance travelled. In this section we investigate causal connections between the general run of winter weather and transport using best subset multiple regression analysis.

The predictands were: all-vehicle mileage, commercial-vehicle mileage, total number of road casualties and casualty rate. February was the month chosen for analysis, as being representative of winter conditions yet free of effects from the Christmas and New Year vacations.

We shall consider first the relationship between the two indices of vehicle mileage and general winter weather conditions. The February time series for all-vehicle mileage exhibits a steady upward trend with time. The commercial-vehicle index contains a less regular upward trend from 1958 to 1970, but after this time levels off, as shown in Figure 5. It is, therefore, not possible to consider the raw values for either index. In order to remove the trends for the all-vehicle mileage, a single regression line was fitted to the series and the residuals from this line calculated. These residuals were used as the predictand in the multiple regression analysis. Two regression lines were fitted to the commercial-vehicles time series, the first for the period 1958-70 and the second for the period 1971-79. Again the residuals from these lines formed the predictand.

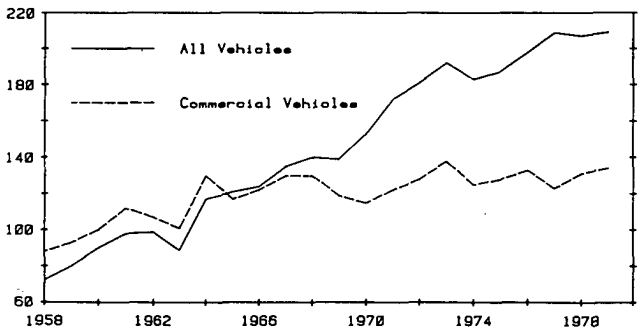


Fig. 5. Indices of All-vehicle and Commercial-vehicle Mileage for February to base 1958 = 100.

The long-term trends in these transport series must be produced by such factors as economic growth and social trends in mobility. It is hypothesised that the fluctuations around the trend, as described by the residuals, are caused in part by short-term variations in weather and in economic factors such as industrial production. These are the predictor variables in the regression analysis. The choice of weather variables was restricted by the fact that the transport series are for the whole of Great Britain. Three measures of weather were used: temperature, rainfall and snow. The Index of Industrial Prodction (IIP) was used to represent short-term economic fluctuations.

The temperature predictor was central England series (CET), on the basis that this is representative of conditions over the most densely populated regions of Britain. For the same reason,

England and Wales rainfall (EWR) was used to describe rainfall (Nicholas and Glasspoole, 1932, updated by the U.K. Meteorological Office). Three predictor variables for snow were examined: number of days with snow falling, number of days with snow lying, and the monthly-mean snow depth. For the first two, it was only possible to obtain a value for the whole winter season (U.K. Meteorological Office, 1979; see Table 1). Both variables are strongly correlated (r = 0.78 over 22 years) and, furthermore, they are strongly correlated with monthly temperature and snow depth data. Because of these correlations (which imply that their variability can be largely explained by other variables), because strongly intercorrelated variables may lead to unstable regression results, and because they are seasonal rather than monthly variables, days with snow falling and days with snow lying were not considered further in the analysis. We recognize, however, that they may be important as primary controlling variables, so we are assuming here that their influence is subsumed by the other predictors.

Raw values of the IIP could not be used to describe short-term economic fluctuations, because this Index contains an upward trend with time. Instead, regression lines were fitted to the (rebased) Index values, and the residuals from these lines used as the predictor (see Palutikof, 1983).

The final set of candidate predictor variables was, therefore,

1 January CET (CETJAN)
2 February CET (CETFEB)
3 January EWR (EWRJAN)
4 February EWR (EWRFEB)
5 Mean snow depth for January (SNOJAN)
6 Mean snow depth for February (SNOFEB)
7 IIP residuals, January (IIPJAN)
8 IIP residuals, February (IIPFEB)

January variables were included to allow for lag effects. For example, a high value for SNOJAN might affect industrial production in such a way that commercial-vehicle mileage in February is substantially reduced.

Even this limited set of variables presents problems in regression analysis. Firstly, with only 22 years of data, the number of degrees of freedom is perilously small. Secondly, the variables themselves are not independent and one might expect problems due to multicollinearity. Thirdly, any results may be largely dependent on a few extreme years and not be representative of more general conditions.

The first difficulty cannot really be avoided given the amount of data available and the nature of the problem with its mesh of possible interactions and potential causal factors. Ideally, any results obtained should be tested on an independent data set, but this is impossible here. Results must, therefore, be taken as indicative rather than definitive.

The second problem, multicollinearity, can affect the stability of the regression results and can make interpretation of regression coefficients hazardous. The strongest intercorrelations are those between temperature and snow depth: for the 22-year study period r(SNOJAN-CETJAN) = -0.87, r(SNOFEB-CETFEB) = -0.76; for the 20-year period (omitting 1963 and 1979) r(SNOJAN-CETJAN) = -0.56, r(SNOFEB-CETFEB) = -0.61. The next strongest correlation between predictor variables is between IIPJAN and IIPFEB (r = 0.80 for the 22-year period, r = 0.84 for the 20-year period). A preliminary scrutiny of the simple correlations between predictands and predictors showed that commercial-vehicle mileage was not noticeably correlated with temperature and so CETJAN and CETFEB were omitted from the commercial-vehicle analyses. All-vehicle mileage is significantly correlated with both CETJAN and CETFEB, but in view of the high correlations between the temperature and snow depth variables we felt that it would be folly to leave both predictors in the regression. The temperature variables were, therefore, also eliminated from the all-vehicle analysis leaving a total of six candidate predictors. We did not feel justified in eliminating any other variables. In interpreting our results, these intercorrelations need to be borne in mind.

Our solution to the third problem has already been alluded to: we simply reanalyzed the data after omitting the most extreme years, 1963 and 1979.

Regression analyses were performed using a best subset routine. Best subset regressions were calculated for 1, 2, 3, etc. predictor variables and the procedure terminated when a variable entered the equation with a t-value less than one. The results were found to be statistically stable. Predictors, once they had been entered into the regression, remained in, and regression coefficient values changed little when new predictor variables were added. The final equations were:

20-year analysis

VEHALL = - 0.190 SNOFEB (R=0.490)

VEHCOM = - 0.303 SNOJAN - 0.190 EWRJAN - 0.051 EWRFEB (R=0.851)

22-year analysis

VEHALL = - 0.098 SNOJAN - 0.199 SNOFEB (R=0.738)

VEHCOM = - 0.150 SNOJAN - 0.170 EWRJAN - 0.036 EWRFEB + 0.361 IIPFEB (R=0.787)

All results are significant at the 1% level or better (F-test) with the exception of the VEHALL-20-year equation which is significant at the 5% level.

The main features of these results are as follows. Firstly, the primary determining factors are climatic rather than economic. Industrial activity (through IIPFEB) only appears in one of the regression equations (VEHCOM-22-year) where it enters third after SNOJAN and EWRJAN. Secondly, the 20-year and 22-year results are similar, indicating that the relationships observed in the complete, 22-year analyses are not simply artefacts of the results for isolated extreme years. For all-vehicle mileage, the relationship is considerably weaker when the extreme years are omitted, but SNOFEB remains the dominant controlling variable. For commercial-vehicle mileage the relationship actually strengthens when the

extremes are omitted, but the change is not statistically significant. Thirdly, the multiple correlation coefficients for the commercial-vehicle equations are higher than those for all-vehicle mileage, pointing to a stronger weather/climate sensitivity for commercial vehicles. This is somewhat surprising since the analysis of the impact of severe seasons (see earlier) suggested that all-vehicle mileage is more affected by climate than commercial-vehicle mileage. However, the major difference in the regression results occurs when the most extreme seasons are omitted so these differences are not necessarily inconsistent.

Multiple regression analyses were also performed to analyse the effects of climate on road casualties. Again February data were used. It has already been suggested, in the section on the impact of extreme winters, that the raw value for casualties is greatly reduced by harsh winter weather, and that this is not a direct weather effect, but the result of fewer journeys being made in harsh conditions. Even if the effect is indirect it should, nevertheless show up as being weather-related in a multiple regression analysis.

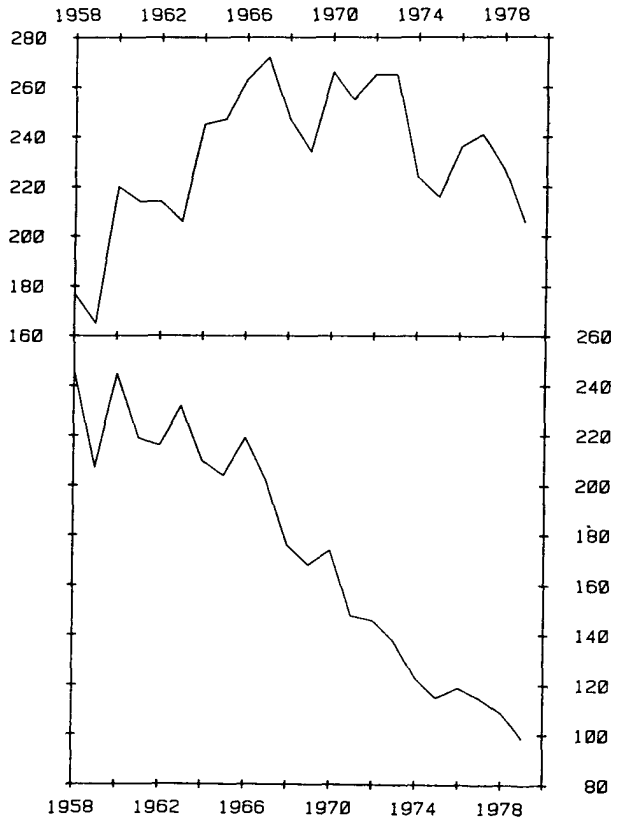


Fig. 6. Numbers of road casualties in February, 1958-79. Upper diagram shows raw value and lower diagram shows casualties controlled for all-vehicle mileage (see Figure 4).

The number of casualties show a non-linear long-term trend (Fig. 6) which must be removed from the data before looking for year-to-year climate effects. Two regression lines were fitted, from 1958 to 1970, and from 1971 to 1979. The residuals from these lines formed the predictand in the regression analysis. Of the set of predictor variables described above, it was felt that only three could be argued to directly influence the number of road casualties: CETFEB, EWRFEB and SNOFEB. As before, since temperature and snow depth are so strongly correlated, CETFEB was excluded from the analysis. Regression equations were developed for both the full 22-year data series and for the 20-year subset with the most extreme years (1963 and 1979) removed. The results obtained confirm our earlier conclusion that the number of casualties diminishes when weather conditions deteriorate. The regression equations are -

20-year analysis

FEBACC = - 57.7 SNOFEB + 14.5 EWRFEB (R=0.578)

22-year analysis

FEBACC = - 48.4 SNOFEB + 13.0 EWRFEB (R=0.593)

Both equations are significant at the 5% level (F-test). These relationships are surprisingly strong given the crudeness of the predictand data (representing the total accident count from a variety of different rural and urban areas over a wide part of the country) and the simplicity of the predictor variables. The similarity of the results indicates that the relationship is a general one and not simply confined to or determined by the most extreme years.

The same analysis was performed for the number of road casualties controlled for vehicle mileage after first removing the trend (see Figure 6). For both the full series and the series with severe winters removed the multiple correlation coefficients were low and not significant. This confirms the evidence presented in Figure 4 which suggested that there is little or no relationship between severe winter weather and the casualty rate, and that the link between the total number of casualties and climate arises largely as a result of weather-related reductions in travel.

5. CONCLUSIONS

An attempt has been made in this paper to quantify the relationship between road transport and weather in Great Britain. Three measures of road transport were examined: distances travelled by all road vehicles, distances travelled by commercial vehicles, and the number of road casualties.

The first section of the analysis dealt with the relationship between these three measures and the severe winters of 1962-3 and 1978-9. By comparing winter mileage with mileage in the preceding autumn it was shown that the distances travelled by all vehicles were substantially reduced in both these severe winters. However, commercial-vehicle mileage proved to be more resilient to harsh conditions.

The number of road casualties drops substantially in a bad winter but when this number is controlled for distances travelled the

relationship breaks down. The implication is that the cause of fewer casualties is the fact that people do not travel, rather than that they drive more carefully.

To test the relationship between road transport and general weather conditions, multiple regression analyses were performed using a best subset routine. Time series of the three transport measures were used as predictand variables, after first removing long-term trends. For the mileage measures, a set of six candidate predictor variables was isolated, representing climatic and economic conditions. Best subset regressions were performed, increasing the number of predictor variables by one at each step until a variable entered the equation with a t-value less than one. The primary determining factors are shown to be climatic rather than economic. When the extreme seasons are removed from the time series, the results remain similar, suggesting that general weather conditions do contribute to determining road-vehicle mileage. The sensitivity of commercial-vehicle mileage is shown to be greater than was suggested by the analysis of the impact of severe winters.

Only two weather variables were used in an attempt to predict the number of road casualties. The relationships proved surprisingly strong and persisted when the extreme seasons were removed. As expected, when casualties were controlled for vehicle mileage the multiple correlation coefficients proved to be low and not significant.

Reasonably strong relationships have been shown to exist between interannual climate fluctuations and road transport. However, this does not necessarily imply that longer time scale climatic variations and road transport would be similarly related. Some societal adjustments could be expected if a gradual climatic change towards, for example, colder weather were to occur, allowing road transport to settle into a new but stable relationship with the weather. Thus a trend in the climate is unlikely to cause long-term disruption, although considerable expenditure might be necessary to adapt to changed conditions. A climatic change towards greater variability may have more significant implications.

Acknowledgements

The research in this paper was funded by the U.K. Department of the Environment under Research Contract No. DGR/480/579. I would like to thank Tom Wigley and Peter Mayes for their advice and assistance in the preparation of this paper.

6. REFERENCES

Central Statistics Office, 1958-79: Monthly Digest of Statistics, 145-408, H.M.S.O., London.

Codling, P.J., 1971: Thick fog and its effect on traffic flow and accidents. Road Research Laboratory Report, LR 397, H.M.S.O., London.

Codling, P.J., 1974: Weather and road accidents. In, Climatic Resources and Economic Activity (J.A. Taylor, ed.), David and Charles, Newton Abbot, 205-222.

Department of Transport, Welsh Office and Scottish Development Department, 1981: Transport Studies Great Britain 1970-1980, H.M.S.O., London.

Doornkamp, J.C., K.J. Gregory and A.S. Burn (eds.), 1980: Atlas of Drought in Britain 1975-76, Institute of British Geographers, London.

Hay, J.S., and C.P. Young, 1972: Weather forecasting for the prevention of icy roads in the United Kingdom. In, Weather Forecasting for Agriculture and Industry (J.A. Taylor, ed.), David and Charles, Newton Abbot, 155-165.

Landsberg, H.E., 1972: The assessment of human bioclimate: a limited review of physical parameters. W.M.O. Technical Note 123, W.M.O. No. 331, World Meteorological Organization, Geneva.

Manley, G., 1974: Central England temperatures, monthly means, 1659-1973. Quart. J. R. Met. Soc., 100, 389-405.

Nicholas, F.J., and J. Glasspoole, 1932: General monthly rainfall over England and Wales, 1727-1931. British Rainfall 1931, 299-306.

Palutikof, J.P., 1983: The impact of weather and climate on industrial production in Great Britain. J. Climatology, 3, 65-79.

Tellings, M.M.J., 1978: The influence of temperature on industry demand for petroleum products and different types of energy. Supply Oil Programmes and Plans, Shell Centre, London.

Tromp, S.W., 1980: Biometeorology, Hyden, London.

Shellard, H.C., 1968: The winter of 1962-3 in the U.K. - a climatological survey. Met. Mag., 97, 129-141.

A REGIONAL COMPARISON OF SEASONAL AFRICAN RAINFALL ANOMALIES*

Thomas L. Puterbaugh, Raymond P. Motha and Norton D. Strommen

USDA, World Agricultural Outlook Board,
Washington, D.C. 20250

Wayne L. Decker

Atmospheric Science Department,
University of Missouri,
Columbia, Missouri 65201

Sharon K. LeDuc

NESDIS, Assessment and Information
Services Center
Columbia, Missouri 65201

1. INTRODUCTION

The 1968-1973 drought in sub-Saharan Africa gained worldwide attention by illustrating the vulnerability of agriculture in semi-arid regions to extremes in rainfall variability. In fact, drought conditions have continued to plague Africa for years (Lamb, 1982). Numerous studies have examined rainfall variability for specific regions within sub-Saharan Africa, most notably the Sahel, and discussed the possible causes of persistent drought.

The extent of the north-to-south migration of the Intertropical Convergence Zone (ITCZ) is a major determinant of rainfall distribution in sub-Saharan, eastern, and portions of southern Africa. The strength and movement of the ITCZ in these areas is a primary cause of rainfall variability (Trewartha, 1961). If it is assumed that mechanisms which control rainfall over Africa operate on a very large scale, then statistical comparisons of rainfall between regions might show that anomalous wet or dry years are related. This information would be highly useful as an early-alert system to identify high risk periods for drought and subsequent crop failure. Thus far, a majority of the analyses of African rainfall patterns have centered around testing for trend, persistence, or periodicity. For example, Winstanly (1973) concluded from extrapolating past trends that future seasonal rainfall in the Sahel Region of West Africa will decrease to a minimum in A.D. 2030.

Other studies have examined rainfall variability in West Africa. Bunting et al. (1976) tested a mean annual rainfall percentage series comprised of five stations in West Africa for 1904-1974. No linear trends and no statistically significant oscillations were found in the series. The pattern of wet and dry years also was concluded to be random. The tendency for persistent drought years in West Africa has also been investigated by Kraus (1977), Katz (1978), and Nicholson (1979). Kraus tested a normalized rainfall series for the time period 1911-1974. Katz further analyzed the rainfall data collected by Kraus, but different statistics were used to analyze the data. Nicholson expanded the work of the previous two researchers using a more detailed data set. A normalized rainfall series similar to Kraus's was created, but for three regional subdivisions of West Africa. The conclusions of the three authors varied. Both Kraus and Nicholson found an indication of persistence while Katz concluded that his time series was highly variable.

East African rainfall data have been studied by Rhode and Virji (1976). Annual rainfall data for selected stations were grouped into eight regions. Areal averaged precipitation series were then constructed for the period 1922 through 1973. Each series was tested for trends and periodicities. No definite long term trends were found. However, spectral analysis indicated that three major cycles corresponded to periods of 2.0, 3.5, and 5.0-5.3 years.

In South Africa, Brooks and Mametse (1970), investigated annual rainfall trends for selected stations for 60, 70, and 80-year periods. Distinct areas with positive and negative rainfall trends were found. With an expanded data set, Tyson et al. (1975), also tested selected stations for the presence of trend. Although the existence of trends was found in some stations, the distribution of stations exhibiting trend showed no spatial clustering.

---

*Contribution from the Missouri Agricultural Experiment Station, Journal Series No. 9365.

Onesta and Verhoef (1976), tested annual rainfall series of district means using the period 1931-1970. No trends were found. Using district means for 1910-1972, Tyson and Dyer (1975) examined the rainfall data for periodic fluctuations. The results indicated that the Summer Rainfall Region of South Africa has been affected predominantly by a quasi-twenty year oscillation in annual rainfall.

Ogallo (1979) analyzed the time series of annual rainfall for each of 69 stations throughout Africa testing for trends and periodicities. Results indicated that most of the annual rainfall series did not have a statistically significant trend. However, prominent cycles in many of the annual rainfalls were 2.0-2.5 years and 2.7-3.3 years, each significantly different from white noise at the 0.05 level assuming white noise.

Attempts to examine the rainfall between different African regions for similar time periods have been made by Nicholson (1980, 1981a and 1981b). In the first study, a correlation technique was used to determine areas of similar climate variation north of 10°S latitude. Similar anomaly patterns characterizing the Guinea Coast, North and East Africa, during periods of very anomalous rainfall in the Sahel and Sudan, were found. In the second study (Nicholson, 1981b), rainfall for regions within northern hemispheric Africa were correlated. The results indicated that regions within the areas of West and East Africa, correlate well with each other. However, when correlations were done on regions outside their respective areas, the correlations were poor. In the third study (Nicholson 1981b), annual rainfall departure patterns were correlated for Africa north of the equator, and then classified into anomaly patterns. Data for regions in southern Africa were then stratified according to the anomaly patterns previously derived for northern hemisphere Africa. These map patterns, suggested that rainfall fluctuations between northern and southern hemisphere Africa are somewhat related. However, departures in opposite directions tend to occur in equatorial Africa.

This paper also examines regional African precipitation series for non-randomness, trend, or periodicity. Precipitation time series are presented for each region and distinct climate anomalies are described. In addition, statistical evidence for the existence or non-existence of teleconnections between the seasonal rainfall series for different regions in northern and southern hemisphere Africa is presented.

All analyses are based on the station network in Figure 1. Data consisted of monthly rainfall data obtained from three sources. The National Center for Atmospheric Research (NCAR) supplied a magnetic tape which contained data for a series of 419 African stations located in an area from 10°S latitude to 37°N latitude. Data were also obtained from the World Meteorological Organization (WMO) Telecommunications network. Lastly, the Weather Bureau in Pretoria, South Africa provided monthly precipitation data for various districts in South Africa.

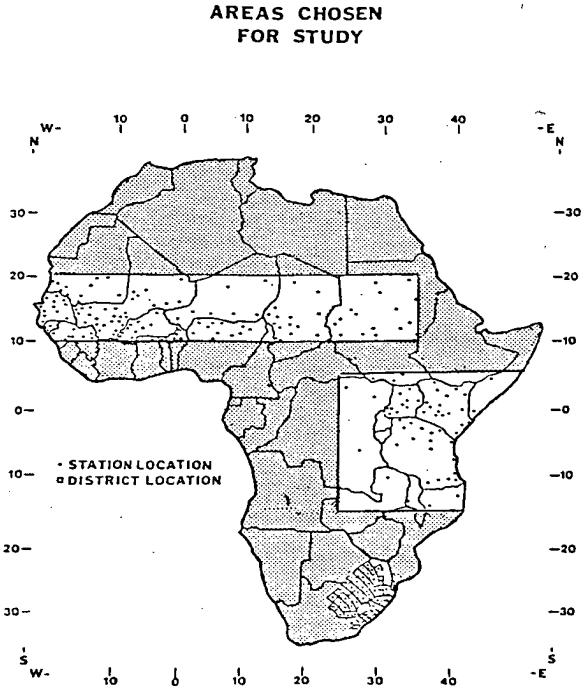


Figure 1. Network of stations and districts which were found to lie within the area boundaries (unshaded areas) and chosen for analysis.

2. CLIMATE FEATURES

Rainfall in sub-Saharan Africa, between 10° to 20°N, is associated with the advancement and retreat of the ITCZ. Most of the rain occurs from May-October. In sub-Saharan Africa, the months of greatest rainfall are July and August when the ITCZ is the farthest north. The length of the rainy season and consequently total annual rainfall varies latitudinally decreasing from 10°N to 20°N latitude. Trewartha (1961), describes the structure of the ITCZ as a shallow sloping layer of air which from its maximum northward surface position increases in thickness southwards towards the equator. Extensive rainfall may occur about 300-600 km south of the northernmost position of the ITCZ provided that a sufficiently deep maritime current exists. These rains often form as squall lines and supply the major part of the total seasonal rainfall. Rainfall is usually below normal in years when squall lines are not observed (Rijks, 1979). The ITCZ retreats southward out of sub-Saharan Africa in October, advancing into eastern Africa. Like sub-Saharan Africa, the ITCZ in eastern Africa is the major rain producer. In eastern Africa, the ITCZ fluctuates north and south of the equator during the year. As a result, regions close to the equator experience two wet and dry periods each year while those further north or south have one wet and one dry season. Unlike sub-Saharan Africa, two atmospheric circulation patterns produce rainfall in eastern Africa: the southeast (SE) monsoon; and the northeast (NE) monsoon (Trewartha, 1961). Rains produced by the the NE monsoon occur during October-December when the ITCZ moves southward. Areas north of the equator experience a dry season during the NE monsoon.

Rains produced by the SE monsoon occur from March-May as the ITCZ moves slowly northward.

In contrast with sub-Saharan and eastern Africa, where the advancement of the ITCZ is the major rain producer, many of southern Africa's rainfall producing systems are associated with mid-latitude circulation patterns. The major climate features are oceanic subtropical high pressure cells, both to the west and to the east of southern Africa. The South Atlantic anticyclone remains relatively fixed to the west while, the position of the Indian Ocean anticyclone to the east fluctuates more (Triegaardt, 1960). In the summer months (November-May), the Indian Ocean anticyclone advances eastward and a shallow thermal trough develops over the interior plateau. The trough promotes the southward influx of moist tropical air from the Congo Basin and Indian Ocean area (Trewartha, 1961), allowing shower activity on the eastern side of southern Africa. The rainy season in this region occurs primarily from November-April. As a result, rainfall analyses in southern Africa were restricted to the South Africa Summer Rainfall Region (Figure 4) to permit comparision with the other summer rainfall regimes of sub-Saharan and eastern Africa.

3. APPROACH

Based on the climatic mechanisms which produce rainfall in sub-Saharan, eastern, and southern Africa, boundaries were established for the station and district networks (Figures 2, 3 and 4). A total of 133 stations were found to lie within the area of sub-Saharan Africa, 84 in eastern Africa, and 42 districts were located within the South Africa Summer Rainfall Region. Two of the three areas were then further sub-divided into regions. The criteria used to determine the regional boundaries were based on rainfall characteristics, circulation features, and between region variability. The 1934-1973 period was chosen as the basic time frame for the analyses to establish an unbroken series of data. The normal seasonal rainfall distributions for each region were derived by averaging the monthly rainfall for all available stations in each region for the 1934 to 1973 period. Table 1 lists the total number of stations used in deriving the normal seasonal rainfall distributions for each region.

Sub-Saharan Africa area was divided longitudinally into the West (20°W to 0°), Central (0°E to 20°E), and East Region (20°E to 35°E). The north and south boundaries were 20°N to 10°N respectively for all three regions (Figure 2). The rainfall distributions indicate that rainfall decreases from a maximum in the West Region, which is adjacent to the North Atlantic Ocean and Gulf of Guinea moisture source, to a minimum in the East Region.

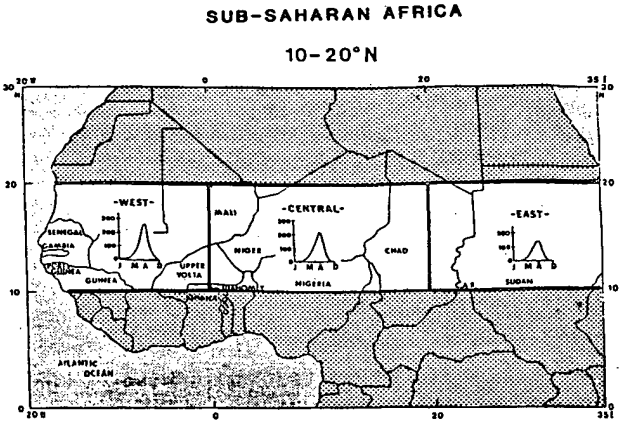


Figure 2. Map of sub-Saharan Africa (unshaded area) showing the regional boundaries (in bold black lines) and normal seasonal rainfall distributions.

Eastern Africa was subdivided into four regions (bold lines in Figure 3). The normal seasonal rainfall distributions for Regions A, B, C, and D are shown. An inspection of the normal seasonal rainfall distribution for Region A reveals that a double maximum of precipitation occurs in this region. This double maximum is a result of the north to south migration of the ITCZ through Region A which is in close proximity to the equator. Since Region A experiences two rainy seasons during the year, annual statistics may mask rainfall deficiences or excesses in one of these periods in a time series analysis. Trewartha (1961) noted that the October-December "short rains," produced by the NE monsoon, are more variable than the March-May "long rains," produced by the SE monsoon. In fact, a failure of the "short rains" can have a detrimental

TABLE 1

Seasons of Maximum Rainfall, Number of Stations Within Each Region of Africa, and Percentage of Annual Rainfall Occurring in the Seasons of Maximum Rainfall for All Regions in the Study

Area	Region	Max Rainfall Season	Number of Stations	Percentage of Annual Rainfall
Sub-Sahara	West	May-October	72	97%
Sub-Sahara	Central	May-October	40	96%
Sub-Sahara	East	May-October	21	98%
Eastern	A1	March-May	26	46%
Eastern	A2	October-December	26	31%
Eastern	B	April-October	13	81%
Eastern	C	April-June	4	59%
Eastern	D	November-April	15	88%
South Africa	Summer Rainfall	October-April	(42) Districts	87%

effect on crop production (FAO, 1980). In Figure 3, Region A, seasonal subdivisions were produced by placing a subscript 1 above the March-May maximum and a subscript 2 above the October-December maximum. In future analyses Region A will be referred to as Region A1 or Region A2, depending on which maximum rainfall season is being represented. Region A1 and A2 contain the same stations. Further, for most of the other regions used in this study, seasonal rainfall in a given year comprises more than 80 percent of total annual rainfall (Table 1). Therefore, seasonal rainfall totals were used in this study for all regions instead of annual rainfall totals. In South Africa, the area bounded by the solid back line in Figure 4 is the South Africa Summer Rainfall Region, represented by 42 districts.

After stations were assigned to their appropriate regions, normalized rainfall departure series were then derived for each region (Figure 5A, 5B and 5C). The procedure is essentially the same one used by Kraus (1977). A normalized departure value of zero (shown by a straight line on the graphs), represents the mean rainfall value during the period 1934 through 1973. The extent and magnitude of dry and wet years are clearly illustrated in the normalized seasonal rainfall series.

A review of Figure 5A reveals a marked similarity between rainfall fluctuations for the West Region, Central Region, and East Region in sub-Saharan Africa. Wet or dry periods on the order of 10 years are evident in the series. In general, a deficit of rain was found in all three series during the 1940's, in the late 1960's and early 1970's. Periods of above-normal rainfall were found in the 1930's, and during the 1950's. These results are similar to those of Kraus (1977), and Nicholson (1979). Similar periods of rainfall fluctuations to those in sub-Saharan Africa have been experienced in the South Africa Summer Rainfall Region shown in Figure 5B. In general, wet periods occurred in the 1930's and 1950's, while periods experiencing below normal rainfall were the 1920's 1940's, and 1960's. However, during peak years of the sub-Saharan drought in the early 1970's, rainfall anomalies in the Summer Rainfall Region were not as extreme and, in some years, were above normal. Similar results for the Summer Rainfall Region were also found by Tyson and Dyer (1978); however, the most extreme years were found not to coincide. Nicholson (1981), also found that on the scale of decades or longer, fluctuations in southern African rainfall series strongly parallel those in northern Africa.

EASTERN AFRICA

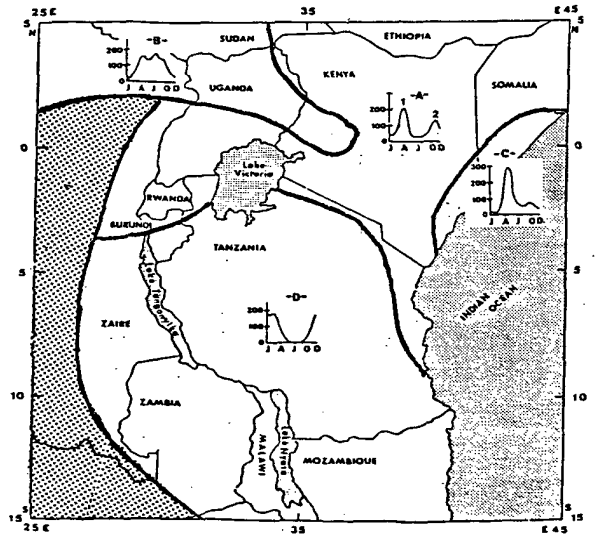


Figure 3. Map of eastern Africa (unshaded area) showing the regional boundaries (bold black lines) and the normal seasonal rainfall distributions.

SOUTH AFRICA

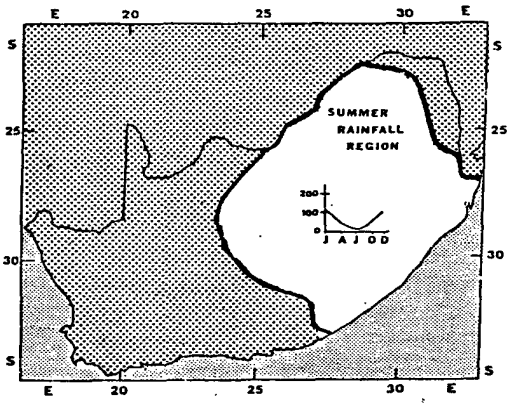


Figure 4. Map of South Africa, showing the boundaries for the Summer Rainfall Region (unshaded area) and the normal seasonal rainfall distribution.

The seasonal standarized precipitation departures series for regions in Eastern Africa is given in Figure 5C. Although periods with similar fluctuations were found to exist between some regions, these periods often did not coincide with wet or dry periods in either the sub-Saharan or southern Africa rainfall series. In fact, rainfall fluctuations for some regions in eastern Africa were highly variable in some periods and, in some cases, were often dissimilar. Periods for which rainfall fluctuations did not coincide with those in sub-Saharan and southern Africa were the 1930's in Region A2, the 1940's in Regions C and B, the 1950's in Region B, and the 1960's in Regions A2, B, and D.

Therefore, visual inspection of the seasonal standardized rainfall departure series for all nine regions indicate that rainfall fluctuations in sub-Saharan Africa tend to parallel those in southern Africa on the scale of 10 years. However, most extremes do not appear synchronous. The apparent similarites between and amongst rainfall series in sub-Saharan and southern Africa suggests that mechanisms which control rainfall variability in Africa operate on a very large scale. Rainfall fluctuations in part of

NORMALIZED SEASONAL RAINFALL SERIES

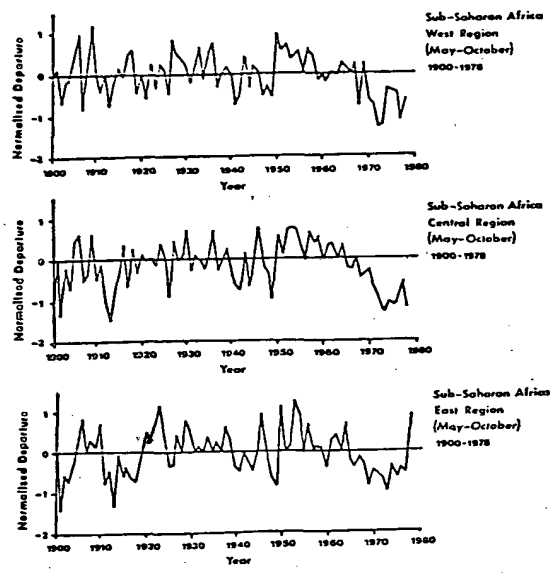


Figure 5a. Rainfall time series for sub-Saharan Africa.

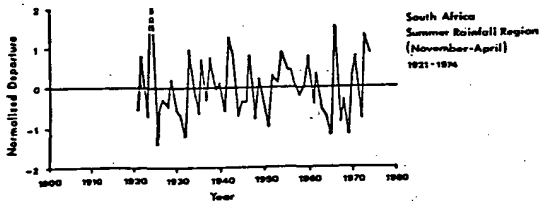


Figure 5b. Rainfall time series for South Africa Summer Rainfall Region.

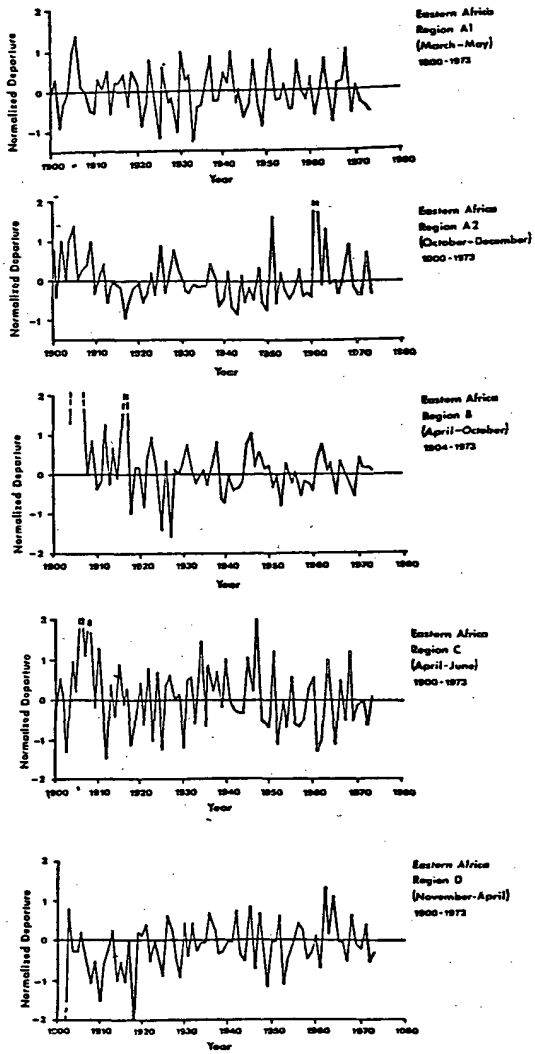


Figure 5c. Rainfall time series for eastern Africa.

eastern Africa tend to depart for those in other areas indicating that local physiographic as well as climatic features may influence rainfall fluctuations in these areas.

4. ANALYSES AND RESULTS

Before statistical comparisons of rainfall between regions were made, the regional normalized rainfall series were tested for the presence of trend, persistence, and periodicities. The purpose of these tests was to determine the likelihood of drought in a region for a certain year, given the climate occurrences in previous years, in the same region. In the trend tests, twenty-year periods beginning with 1904 were tested for short-term trends, while the base period 1934-1973 was tested for the presence of a long-term trend. The 1934-1973 period was used in the persistence tests. Lastly, the time period chosen for the spectral analysis procedure was 1904-1973 for most regions. However, due to a limited number of years with data, the time period 1925-1973 was used in the analysis for the Summer Rainfall Region of South Africa.

The results indicate that the existence of long-term or short-term trends does not exist in the normalized regional departure series. Rainfall in most years was found to be of random occurrence for most regions, except the East Region in sub-Saharan Africa, where persistence is indicated. Finally, in most cases, spectral analysis by the Blackman-Tukey method found no statistically significant oscillations in the regional departure series. These results suggest that the predictive capability of drought occurrences in a region for a certain year, from the climate occurrences in previous years in the same region, does not exist.

The coherence of seasonal rainfall patterns between regions located within their respective areas was tested to determine if drought years between regions are related. The 1934-1973 period of record was used in the analysis. South Africa was omitted since it was not divided into

regions. In Table 2, the intra-area correlation coefficients are presented. The underlining below the various correlation coefficients indicates the level of statistical significance for each. Correlation coefficients which have no underlining were not significant at the 0.10 confidence level.

A review of the intra-area correlations for sub-Saharan Africa (Table 2), indicates that the West and East regions are well correlated with the Central Region. The regional correlations for sub-Saharan Africa were significant at the .001 level, but were found to decline from west to east. These results imply that rainfall departures are statistically coherent throughout sub-Saharan Africa, and wet and dry years are generally synchronous over the entire area. However, like the results of Nicholson (1980), the apparent decline in the correlation coefficients from west to east suggests that this similarity of dry and wet years over the area occasionally breaks down, with droughts in some years affecting only certain longitudinal sectors. The low correlation coefficient for the East Region suggests the tendency for the climate behavior of this region to depart from that of the entire area. This may be due to its relative location away from the North Atlantic moisture source and closer links to the influence of the Arabian anticyclone. Also, the tendency for dry and wet years to persist in the East Region may be a factor.

TABLE 2

Intra-area Correlation Coefficients

SUB-SAHARAN AFRICA

Region -	West	Central
Central	0.628	
East	0.394	0.522

EASTERN AFRICA

Region -	A1	B	C	D
A2	0.167	-0.013	0.015	0.359
D	0.182	-0.013	0.133	
C	0.300	-0.028		
B	-0.236			

Note. Underlining indicates confidence levels: double bar, 0.001; single bar, 0.01; dashed bar, 0.05; dotted, 0.10.

Returning to Table 2, intra area correlations between regions A1, A2, B, C, and D in eastern Africa are given. Most correlations between regions in eastern Africa are weak. Although some type of a positive relationship exists between regions A2 and D, as well as regions A1 and B, these correlations are also weak. In general, little coherence appears to exist between the seasonal rainfall patterns for most regions in eastern Africa.

These results are not suprising as Nicholson (1981b) discovered from previous research that stations in parts of eastern Africa correlate poorly with each other. Therefore, when averaging these seemingly independent stations together to form regions one might also expect to find little coherence between regions. The reasons for the poor correlations include the nature of the tropical storm systems which produce sporadic rains over eastern Africa and two different atmospheric circulations which produce rainfall over the area. The monsoons are composed of different climatic elements each of which are likely responsible for differences in rainfall fluctuations. The influences of physiographic features such as lakes and highlands might also contribute by imposing small scale variations on the larger scale features.

The idea that teleconnections might exist between different regions in Africa has been presented by Nicholson (1980, 1981a and 1981b). However, no known research correlates the rainfall between different African regions both in the northern and southern hemisphere to statistically prove whether such relationships exist. Earlier inspections of the normalized rainfall series revealed that rainfall fluctuations in sub-Sahara and the South Africa Summer Rainfall Region appear to be somewhat synchronous. This may be a hint that mechanisms which produce rainfall over Africa operate on a large scale and climatic teleconnections may exist over Africa. Dyer and Tyson (1978), examined the year-to-year positional changes in the South Atlantic and South Indian Ocean anticyclones, and their relationship with changes in other major components of the general circulation in both hemispheres to determine if links between upper air synoptic features exists. No attempt was made to correlate rainfall anomalies at the surface to upper air pressure systems over Africa. The results implied that during July, shifts in the South Atlantic anticyclone equatorward are associated with a northward displacement in the North Atlantic anticyclone and the Atlantic ITCZ, away from the equator. Furthermore, Kidson (1977), found evidence for links between changes in large scale circulation features over Africa and rainfall anomalies in West Africa. Applying these results to weather over eastern Africa, the idea that climatic teleconnection may exist between regions in this area and other African regions cannot be ruled out. Shifts in the ITCZ from its normal latitudinal position over eastern Africa may produce similar shifts in the ITCZ over sub-Saharan Africa and vice versa. From an agricultural perspective, this result has numerous ramifications, as agricultural productivity in many areas of Africa relies heavily on the short and sometimes erratic rainfall during the wet season. The 1968-1973 drought in sub-Saharan Africa is a well known example. An apparent failure of the ITCZ to move into the area as illustrated by Motha (1980) caused extensive cattle losses and widespread starvation. Similar drought episodes have adversely impacted agricultural productivity in eastern Africa (FAO, 1980) and more recently, a prolonged dry period in the South Africa Summer Rainfall Region during the 1982/83 growing season reduced corn yields to the lowest on recent

record. If statistical evidence of the existence of teleconnections between rainfall anomalies in different Africa regions is found, this information would be highly useful as an early-alert system to identify high risk periods for drought and subsequent crop failure.

In order to find this statistical evidence, regions located within different areas were correlated using the Kendall's tau correlation coefficient. The results of these inter-area regional correlations are given in Table 3. The time period over which the seasonal rainfall was correlated was 1934-1973 for all regions. The results of the correlations between the Summer Rainfall Region of South Africa with the West, Central, and East regions of sub-Saharan Africa indicates that the seasonal rainfall patterns for these regions are independent. The seasonal rainfall for each of the five regions in eastern Africa were also correlated with each region of sub-Saharan Africa (Table 3). Generally, weak or non-significant correlations were found to exist in the seasonal rainfall patterns between the regions in these two areas. Finally, the seasonal rainfall for the Summer Rainfall Region of South Africa does not generally appear to be correlated with the seasonal rainfall for any of the regions in eastern Africa. However, a weak negative correlation of -0.35 was found between the Summer Rainfall Region of South Africa and Region A2 the short rainy season. The low correlation, as well as the possibility of a type I error, reduces the utility of this negative relationship.

TABLE 3

Inter-area Regional Correlations

SUB-SAHARAN AFRICA

Region -	West	Central	East
SOUTH AFRICA			
(Summer Rainfall Region)	-0.069	0.000	0.059
EASTERN AFRICA			
Region			
A1	0.092	0.038	0.122
A2	-0.003	0.051	0.122
B	-0.185	-0.100	0.017
C	-0.161	-0.133	0.032
D	-0.079	-0.103	-0.012

EASTERN AFRICA

Region -	A1	A2	B	C	D
SOUTH AFRICA					
(Summer Rainfall Region)	0.033	-0.354	0.074	0.133	0.163

Note. Underlining indicates confidence levels: single bar, 0.01; dotted bar, 0.10.

In summary, the predictive capabilities for using the seasonal rainfall in one region to predict the climatic occurrences in other regions do not appear promising. The lack of strong correlations was not completely surprising given the rather simplified assumption that large-scale weather features are related over Africa and that recurrent displacements in these upper air features would produce the same anomalous rainfall patterns at the surface year after year. Also, the simple reasoning that the displacement of the ITCZ is solely responsible for wet or drought years over Africa may also be a false one in some years. Nicholson (1981a), found that the position of the ITCZ in sub-Saharan Africa is not a critical factor in drought, but its northward displacement is responsible for anomalously wet years. During drought years the ITCZ was found to be at its normal position but of reduced intensity. In eastern Africa, the concept that the seasonal migration of rains across the equator takes place along a zonal rainbelt might also be misleading. In fact, rains tend to fall in irregular patches separated by areas of scattered or little if any rainfall (Johnson, 1962).

These results suggest that a complex interaction of climatic mechanisms are likely responsible for yearly rainfall patterns in Africa. In fact, Lamb (1978) found that distinct anomalies in sea surface temperatures can be linked to wet and dry occurrences in Africa. In addition, the influences of local physiographic features such as lakes and highlands, the strength of prevailing ocean currents, as well as other regional climatic factors serve to influence rainfall throughout Africa. Therefore, although evidence of significant interactions exist between large-scale weather circulation features, regional climatic factors serve to break down such interactions at the surface.

Recent drought occurrences over Africa illustrate the unrelenting evidence of large-scale regional drought patterns. During the 1981-82 summer rainy season over southern Africa, rains were erratic or in some cases failed to materialize. Furthermore, the worst drought in recent history has also plagued these same areas during the 1982-83 growing season. According to FAO (1982), rains in eastern Africa during 1982 were generally beneficial but in parts of West Africa, particularly Mauritania, Ghana, Togo, and Chad drought conditions damaged crops and stressed livestock.

5. CONCLUSION

The highly vulnerable nature of agricultural productivity in many areas of Africa to the sometimes meager, erratic, and even unreliable rainfall during the growing season has been given much publicity. The devastating drought in sub-Saharan Africa which began in 1968 and continued to the present in some areas sparked this worldwide concern. Numerous research articles analyzing both surface and upper-air weather phenomenon have been performed to determine the nature of rainfall variability over Africa. Like other research, rainfall analyses in this study tested for trend, patterns, and periodicities in seasonal rainfall times series for selected African regions. In addition, statistical evidence for the existence or non-existence of teleconnections between the seasonal rainfall series for the various regions in Africa is presented.

These results suggest that rainfall variabilty for these semi-arid regions should be considered random. The formulation of a solution to the problem raised by the delicate balance between the capacity for agricultural production to meet the needs of the increasing human population, and highly variable nature of the rainfall throughout the region, require a comprehensive analysis. An in-depth analysis of large-scale upper atmospheric and surface data, including sea surface temperatures, may provide further insight into possible connections between the anomalies in the conditions of the ocean surface and large-scale atmospheric pressure patterns as related to seasonal rainfall in Africa. However, these initial results suggest local controls play a major role in determining African rainfall.

6. REFERENCES

Brooks, G.A. and M.N. Mametse. 1970. "Rainfall Trend Patterns in South Africa." South African Geographical Journal., 52, 134-138.

Bunting, A.H., M.D. Dennet, J. Elston and J.R. Milford. 1976. "Rainfall Trends in the West African Sahel." Quarterly Journal of the Royal Meteorological Society., 102, 59-64.

Dyer, T.G.J. and P.D. Tyson. 1978. "Inter-hemispheric Associations in the Year-to-Year Variations of Major Components of the General Circulation of the Atmosphere: 1910-1963." South African Journal of Science., 74, 325-328.

FAO. 1980 and 1982. Special Report - Foodcrops and Shortages. "Global Information and Early Warning System on Food and Agriculture." Food and Agricultural Organization of the United Nations, Rome.

Johnson, D.H. 1962. "Rain in East Africa." Quarterly Journal of the Royal Meteorological Society., 88, 1-19.

Katz, R.W. 1978. "Persistence of Subtropical African Droughts." Monthly Weather Review., 106, 1017-1021.

Kidson, J.W. 1977. "African Rainfall and its Relation to the Upper Air Circulation." Quarterly Journal of the Royal Meteorological Society., 103, 441-456.

Kraus, E.B. 1977. "Subtropical Droughts and Cross-Equitorial Energy Transports." Monthly Weather Review., 105, 1009-1018.

Lamb, P.J. 1978. "Large-scale Tropical Atlantic Surface Circulation Patterns Associated with Sub-Saharan Weather Anomalies." Tellus., 30, 240-251.

Lamb, P.J. 1982. "Persistence of SubSaharan Drought." Nature., 299, 46-48.

Motha, R.P., S.K. LeDuc, L.T. Steyaert, C.M. Sakamoto, and N.D. Strommen. 1980. "Precipitation Patterns in West Africa." Monthly Weather Review., 108, (10), 1567-1578.

Nicholson, S.E. 1979. "Revised Rainfall Series for the West African Subtropics." Monthly Weather Review., 107, (5), 620-623.

Nicholson, S.E. 1980. "The Nature of Rainfall Fluctuations in Subtropical West Africa." Monthly Weather Review., 108, (4), 473-487.

Nicholson, S.E. 1981a. "Rainfall and Atmospheric Circulation During Drought Periods and Wetter Years in West Africa." Monthly Weather Review., 109, 2191-2208.

Nicholson, S.E. 1981b. "A Comparative Study of Drought In the Semi-Arid Subtropical Zones of Northern and Southern Africa." WMO Symposium on Meteorological Aspects of Tropical Droughts. (R.P. Pearce, Ed.).

Ogallo, L. 1978. Rainfall in Africa. East African Institute for Meteorological Training and Research, Republic of Kenya Meteorological Department Research Report No. 5/78.

Onesta, P.A. and P. Verhoef. 1976. "Rainfall Trends in 80 Rainfall Districts in South Africa." South African Journal of Science., 72, 237-239.

Rijks, D. 1978. "Experience of Drought in the Sahel Region." Symposium on Drought in Botswana. (M.T. Hinchey, Ed.). Botswana Society in Collaboration with Clark University Press, Hanover, New Hampshire.

Rodhe, J. and Virji, H. 1976. "Trends and Periodicities in East African Rainfall Data." Monthly Weather Review., 104, 307-315.

Trewartha, G.T. 1961. The Earth's Problem Climates. The University of Wisconsin Press, Madison, Wisconsin.

Triegaardt, D.O. 1960. Tropical Meteorology in Africa. Preceedings of the Symposium Jointly Sponsored by the World Meteorological Organization and the Munitalp Foundation (D.J. Bargman, Ed.). Published by Munitalp Foundation.

Tyson, P.D., T.G.J. Dyer, and M.N. Mametse. 1975. "Secular Changes in South African Rainfall: 1880 to 1972." Quarterly Journal of the Royal Meteorological Society., 101, 817-833.

Tyson, P.D. and T.G.J. Dyer. 1975. "Mean Annual Fluctuations of Precipitation in the Summer Rainfall Region of South Africa." South African Georgraphical Journal., 57, 104-110.

Tyson, P.D. and T.G.J. Dyer. 1978. "The Predicted Above-Noraml Rainfall of the Seventies and the Likelihood of Droughts in the Eighties in South Africa." South African Journal of Science., 74, 372-377.

Winstantley, D. 1973. "Recent Rainfall Trends in Africa, the Middle East and India." Nature., 243, 464-465.

CYCLIC COMPONENTS IN PRECIPITATION AND INSOLATION TIME SERIES IN ALTO-ALENTEJO (PORTUGAL)

João A.M. Corte-Real, Carlos Portugal da Câmara and Maria Emilia Costa

Department of Physics, Faculty of Sciences
University of Lisbon
Lisbon-Portugal

1. INTRODUCTION

The main purpose of this paper is to present evidence in favour of the existence of cycles with periods of about 10 years in the time series of meteorological elements and to relate such cycles with solar activity. Other well established periods are also considered.

Indications that this study might be worth doing came from the observation (Serra and Albuquerque de Matos, 1983) that in the Alto-Alentejo Province of Portugal good and bad years of wool yield (merino wool) seem to alternate with a period of about 11 years (see Figs. 1a and 1b) whose importance in solar activity has been well established.

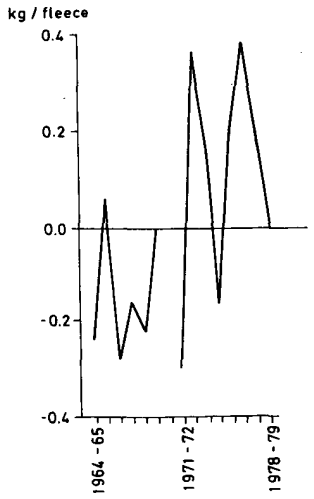


Fig.1a-Deviations from the mean of total merino wool yield (kg/fleece), as a function of year, for Elvas.

Taking into account that precipitation and insolation are the most important meteorological elements for the growth of natural pastures and, consequently, for wool production, the above mentioned fact suggested the possibility of the existence of cycles with a similar period in the time series of those elements. Moreover, from a strict meteorological point of view, other parameters such as relative humidity and temperature, would also deserve to be investigated in this respect.

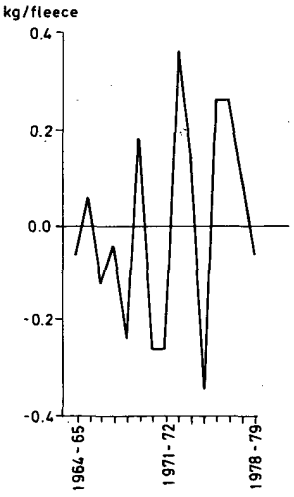


Fig.1b-As in Fig.1a, for Portalegre.

The analysis here reported refers only to the meteorological stations of Elvas and Portalegre, located in the aforementioned province of Portugal (see Fig.2). We intend to extend the work to other stations so as to cover as much as possible of Portugal and especially its South Provinces, where an epoch of drought might be developing.

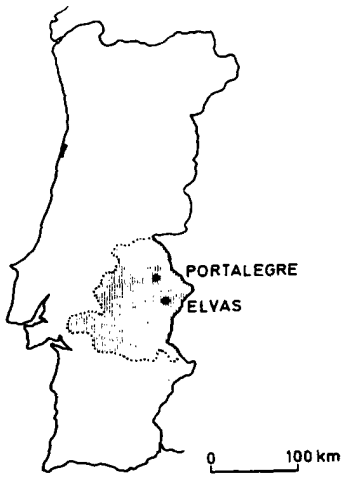


Fig.2-Elvas and Portalegre in the Alto-Alentejo Province of Portugal (shaded area).

2. DATA AND PROCEDURES

Meteorological data used in the present work are the monthly mean values, for Elvas and Portalegre, from January 1947 till December 1980, of precipitation (P), relative sunshine duration (RSD), temperature (T) and relative humidity (RH). Data which refers to the period 1947-1971 were extracted from the Anuário Climatológico, a publication of the National Institute of Meteorology and Geophysics (INMG), the remaining data, not yet published, having been kindly put at our disposition by the INMG.

Solar activity data, specifically the monthly values of sunspot numbers (SSN), for the period 1947-1980, were taken, through courtesy of Observatório Astronómico da Ajuda, from the Quarterly Bulletin on Solar Activity, a publication of the International Astronomical Union (I.A.U.).

Values of the woolyield, for Elvas and Portalegre, were kindly communicated to us by Professor Serra. These values refer to fifteen annual periods or wool years, from April 1 to March 31, starting in 1964.

Procedures employed were those appropriate to describe the components of the four considered meteorological time series, assuming a multiplicative decomposition of the form

$$Y=T_1S_1C_1I_1 \qquad (2.1)$$

where Y stands for the time series variables and the symbols T_1, S_1, C_1 I_1 represent, respectively, the trend and the cyclical, seasonal and irregular index numbers or factors.

These factors, as well as the trends, were computed following Bowerman & O'Connell, 1979. A centered moving average of 12 months was first used in order to remove, from the original data, the seasonal and irregular movements, S_1 I_1, which were then derived from the ratios Y/T_1C_1. The obtained values of S_1I_1 were then grouped by months and averaged for each month, eliminating, in this way, the irregular factors I_1, allowing the isolation of the seasonal indexes S_1. Trend values were next computed by a least-squares linear regression applied to the deseasonalized data. Cyclical factors were obtained by means of a neither too large nor too small 7- month moving average used to smooth out irregular fluctuations in the data previously adjusted for trend and seasonal variations. Irregular indexes were finally estimated as the residuals of equation (2.1).

For each of the considered time series, an additive decomposition of the form

$$Y=T_2+S_2+C_2+I_2 \qquad (2.2)$$

where the symbols on the right hand side respectively denote the trend and seasonal, cyclical and irregular components, has been attempted in order to estimate the percentages of total variance explained by each of them. The additive components T_2, S_2, C_2, I_2, were obtained, from the previously computed factors of the multiplicative decomposition (2.1), through the expressions

$$T_2=T_1=T \qquad (2.3)$$

$$S_2=TS_1-T \qquad (2.4)$$

$$C_2=TS_1C_1-TS_1 \qquad (2.5)$$

$$I_2=Y-T_2-S_2-C_2 \qquad (2.6)$$

whose validity rests on the confirmed assumption that in (2.1) the factors are uncorrelated.

Detection of cycles in the meteorological time series here selected for study, was attempted by two procedures: the so-called Schuster's method (cf., for example, Takahashi, 1980) and another procedure in which the spectra of the time series of the cyclical index numbers C_1, were first computed for each of the four meteorological elements, for Elvas and Portalegre, as the Fourier transforms of the corresponding autocorrelation functions; peaks in the resulting spectra would then reveal the existence of important cycles.

3. DISCUSSION OF RESULTS

3.1 Seasonal, cyclical and irregular contributions for the considered parameters.

Table 1 gives a general perspective of the variability and quality of the meteorological and solar data used in this study, for the period 1947-1980. From this Table it is apparent that, for both Elvas and Portalegre, the four meteorological parameters precipitation, relative sunshine duration, relative humidity and temperature may be arranged by this order with respect to the values of their variances. It may be noted that the variances of temperature and precipitation are respectively one order of magnitude smaller and one or more orders of magnitude larger than those associated with the relative sunshine duration and the relative humidity, which are very similar. This fact can be explained by the observation that precipitation is the meteorological element more affected by non-local phenomena, na-

mely by the irregularly spaced synoptic systems which traverse the country and their associated cloud systems that greatly influence both relative sunshine duration and relative humidity.

Table 1. Parameters respecting to meteorological and solar source data used in this study for the period 1947-1980.

	ELVAS			
PARAMETER	RECORD LENGTH	MISSING VALUES	MEAN	VARIANCE
RSD (%)	408	7	64.8	250.4
P_2 (L/m^2)	408	2	49.9	2701.0
T (^oC)	408	4	15.9	34.7
RH (%)	408	3	63.9	234.8

	PORTALEGRE			
PARAMETER	RECORD LENGTH	MISSING VALUES	MEAN	VARIANCE
RSD (%)	408	145	58.9	296.6
P_2 (L/m^2)	408	1	74.3	5465.7
T (^oC)	408	1	15.2	31.4
RH (%)	408	2	63.3	178.5

PARAMETER	RECORD LENGTH	MISSING VALUES	MEAN	VARIANCE
SSN	408	0	79.9	3417.2

In Table 2 the values of some statistical parameters respecting to the four time series here studied, are presented, for Elvas and Portalegre, and for each index number of the multiplicative decomposition (2.1); the numbers in parantheses are the slopes of the linear regression least squares trend lines. Figs. 3 and 4 show, respectively for Elvas and Portalegre, the time series of the four cyclical factors C_1.

Table 3 has a similar organization as Table 2, but refers to the seasonal, cyclical and irregular components of the additive decomposition (2.2).

Table 2. Mean values and variances, for Elvas and Portalegre, of the index numbers time series, respecting to relative sunshine duration, precipitation, temperature and relative humidity. Slopes of trend lines are indicated in parentheses.

		ELVAS		PORTALEGRE	
PARAMETER	IND. NR.	MEAN	VAR.	MEAN	VAR.
RSD (%)	T	64.9(-0.01)	2.500	59.4(-0.01)	0.600
	S_1	1.0	0.032	1.0	0.041
	C_1	1.0	0.005	1.0	0.011
	I_1	1.0	0.026	1.0	0.047
P_2 (L/m^2)	T	49.1(-0.00)	0.300	72.3(0.01)	2.800
	S_1	1.0	0.309	1.0	0.288
	C_1	1.0	0.219	1.0	0.091
	I_1	1.0	0.774	1.0	0.620
T (^oC)	T	16.0(-0.00)	0.000	15.3(-0.00)	0.300
	S_1	1.0	0.130	1.0	0.124
	C_1	1.0	0.002	1.0	0.002
	I_1	1.0	0.008	1.0	0.011
RH (%)	T	63.9(0.01)	1.000	63.3(0.01)	2.100
	S_1	1.0	0.047	1.0	0.027
	C_1	1.0	0.006	1.0	0.004
	I_1	1.0	0.008	1.0	0.012

Table 3. Mean values and variances for Elvas and Portalegre, of seasonal, cyclical and irregular components respecting to relative sunshine duration, precipitation, temperature and relative humidity.

		ELVAS		PORTALEGRE	
PARAMETER	COMP	MEAN	VAR.	MEAN	VAR.
RSD (%)	S_2	-0.0	136.3	-0.0	143.0
	C_2	0.1	19.7	-0.2	33.6
	I_2	-0.1	86.5	-0.0	124.6
P_2 (L/m^2)	S_2	0.0	743.6	-0.0	1504.5
	C_2	0.5	483.6	0.9	656.7
	I_2	-0.1	1791.7	0.8	3161.4
T (^oC)	S_2	-0.0	33.3	-0.0	28.9
	C_2	0.0	0.4	0.0	0.5
	I_2	0.0	1.3	0.0	2.0
RH (%)	S_2	-0.0	191.1	-0.0	107.1
	C_2	-0.1	19.4	0.0	16.1
	I_2	-0.0	30.3	-0.0	48.7

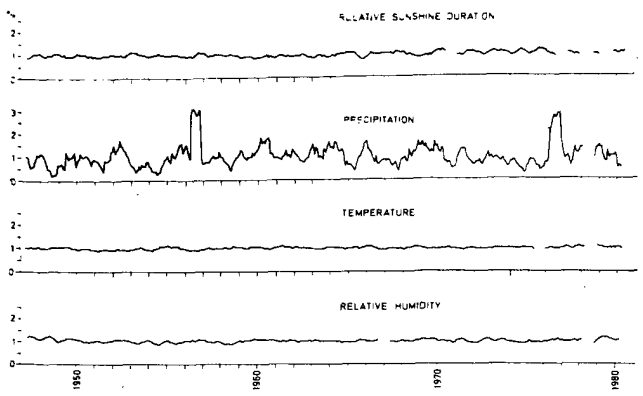


Fig.3-Monthly values of the cyclical index numbers C_1, respecting to relative sunshine duration, precipitation, temperature and relative humidity, for Elvas.

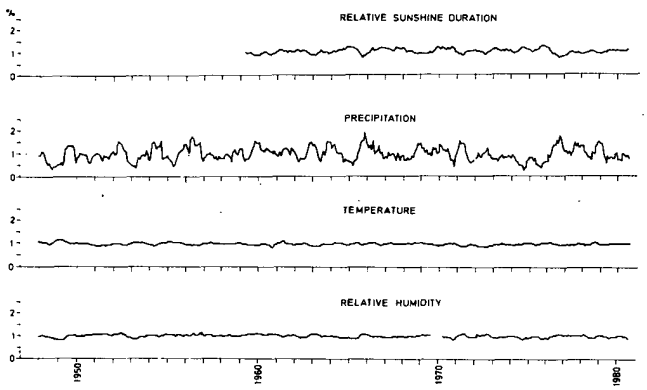


Fig.4-As in Fig.3, for Portalegre.

The computed correlation coefficients between these components proved to be non significant at the 5% level, which guarantees that, for the meteorological time series here considered, expression (2.2) gives, in fact, a partitioning into four statistically independent components, each of which explains a certain fraction of the total variance. Besides, for all the meteorological time series, the irregular index numbers I_1, as well as the irregular components I_2, were tested for its statistical significance (5% level) and found to be white noise.

In Table 4 the percentages of total variances explained by the trend and the seasonal, cyclical and irregular components are indicated, for each meteorological parameter, for Elvas and Portalegre.

Table 4. Percent values of total variance, respecting to relative sunshine duration, precipitation, temperature and relative humidity, explained by trend and the seasonal, cyclical and irregular components, for Elvas and Portalegre.

PARAMETER	COMPONENT	ELVAS	PORTALEGRE
RSD (%)	T	0.99	0.20
	S_2	54.43	48.21
	C_2^2	7.86	11.32
	I_2^2	34.54	42.00
P (L/m^2)	T	0.00	0.05
	S_2	27.53	27.52
	C_2^2	17.90	12.01
	I_2^2	66.33	57.84
T (^oC)	T	0.00	0.95
	S_2	95.96	92.03
	C_2^2	1.15	1.59
	I_2^2	3.74	6.36
RH (%)	T	0.42	1.17
	S_2	81.38	60.00
	C_2^2	8.26	9.01
	I_2^2	12.90	27.28

From Table 2, it can be concluded that the slopes of the least squares regression lines, determined to compute the trend values, are very small; a statistical test, performed for each slope, revealed that all of them are clearly non-significant at the 5% level; hence, the four time series, for Elvas and Portalegre, can be safely considered as stationary.

As shown in Tables 2 through 4 and Figs. 3 and 4, it can be concluded that, for both temperature and relative humidity, the most important contributions for the total variance come from the seasonal and irregular oscillations, the former being however dominant; these two components also account for almost all the variance of relative sunshine duration, each one contributing with fractions of the same order of magnitude. Unlike the other parameters considered, precipitation, which is also the most important meteorological element for the growth of natural pastures and indirectly for wool production and for animal production in general, shows a different temporal behaviour, both in Elvas and Portalegre. In effect, not only there are irregular oscillations responsible for more than half of its variance, but also this parameter shows cyclical components with markedly increased contribution to variance relatively to their homonymous of the other three. Although, for precipitation, predominance of the irregular component could be expected due to the non-local character of the phenomena that condition this meteorological element, as referred above, increased importance of cyclical oscillations is an interesting fact which probably cannot be completely explained by the same reasoning.

This suggests further investigation of periodic components that give significant contributions to the cyclical movements of precipitation time series. For this purpose, spectral analysis and statistical methods certainly are useful tools. A combination of both methods, as in the present study, seems to be also appropriate, especially when the available time series are not sufficiently long to guarantee clear identification of certain periods solely by the methods of spectral analysis.

3.2 Detection of important cycles

Detection of cycles, especially the 10 to 11 years one, in the meteorological time series of relative sunshine duration, precipitation, temperature and relative humidity, was attempted employing the two procedures previously referred to in section 2: the statistical significance test known as Schuster's method and examination of the spectra of the selected meteorological parameters.

We first give a brief review of Schuster's method, as described in Takahashi, 1980, presenting subsequently the results obtained in this study for the considered time series.

Let $X(k\Delta t)$ for $k=0,1, \ldots, N-1$ be a sample record of any meteorological parameter, where Δt is the constant time interval between successive sample values. An harmonic component of period P in X can be written as

$$X_P(k\Delta t)=a_P \cos\frac{2\pi k\Delta t}{P} + b_P \sin\frac{2\pi k\Delta t}{P} \qquad (3.1)$$

The Fourier coefficients a_P and b_P are given by the well known formulas

$$a_P=\frac{2}{N}\sum_{k=0}^{N-1} X(k\Delta t) \cos\frac{2\pi k\Delta t}{P} \qquad (3.2)$$

$$b_P=\frac{2}{N}\sum_{k=0}^{N-1} X(k\Delta t) \sin\frac{2\pi k\Delta t}{P} \qquad (3.3)$$

The amplitude, A, and the phase, θ, of X_P can be obtained from the expressions

$$A=(a_P^2+b_P^2)^{1/2} \qquad (3.4)$$

$$\theta=\tan^{-1}(b_P/a_P) \qquad (3.5)$$

If X is a random time series, with standard deviation σ_X the expected value of the amplitude A is

$$\bar{A}=2\sigma_X/N^{1/2} \qquad (3.6)$$

Hence, if the computed value of A largely exceeds its expected value, in the hypothesis of X being random, it can be concluded that the cycle of period P is statistically significant. Moreover, since the probabilities that the ratio

$$R=A/\bar{A} \qquad (3.7)$$

is greater than 1.5 and 2.0 are respectively, 0.105 and 0.018, if X is a random time series, 1.5 and 2.0 can be used as critical values in a significance test.

Values of A, θ, $\bar{A}$ and the ratio R were computed for the monthly time series of sunspot numbers, relative sunshine duration, precipitation, temperature and relative humidity, respecting to harmonic components of periods ranging from 1 to 12 years. The available yearly time series of wool yield values, for Elvas and Portalegre, were also subject to a similar analysis, bearing in mind however that the results obtained merely constitute an indication, given the shortness of the two series.

In Figs. 5 through 8 the values of the ratio R are plotted against period, for solar activity, the four meteorological parameters and wool yield.

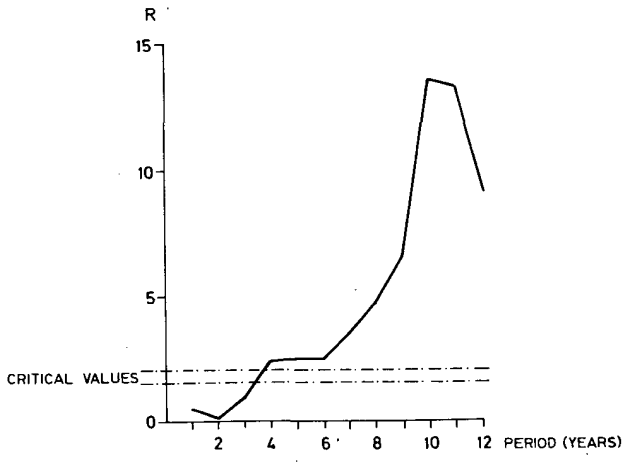


Fig.5-Values of the ratio R, as a function of period, for the sunspot time series.

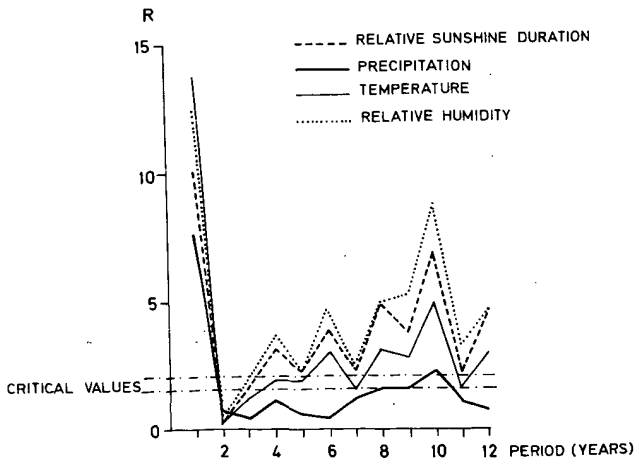


Fig.6-As in Fig. 5, for the time series of relative sunshine duration, precipitation, temperature and relative humidity, at Elvas.

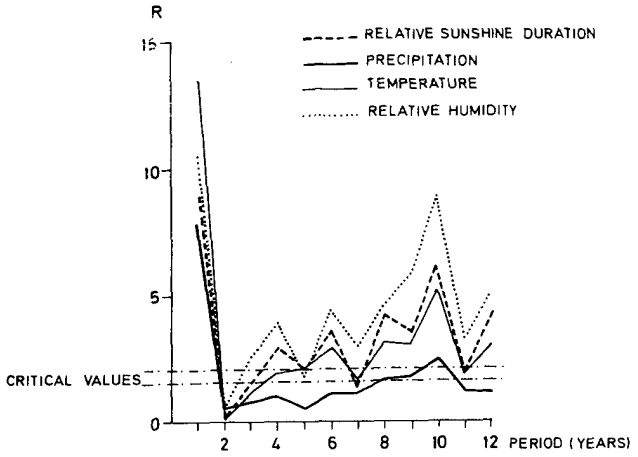


Fig.7-As in Fig. 6, for Portalegre.

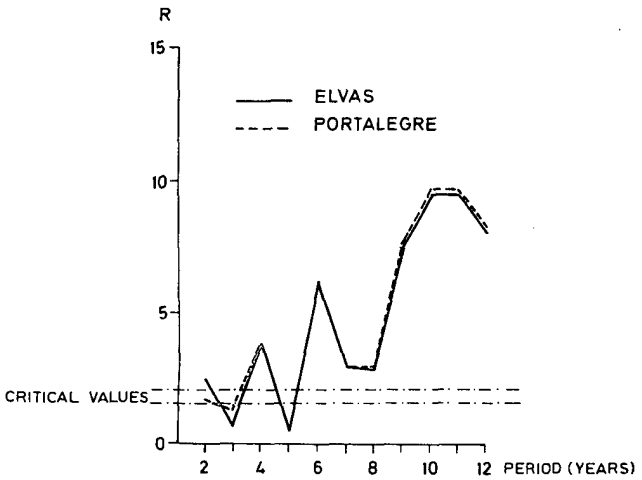


Fig.8-As in Fig. 5, for the wool yield, at Elvas and Portalegre.

From these figures it is clear the existence, in solar activity, of a fundamental 10 to 11 year cycle as well as of a weaker 4 to 6 year cycle. All meteorological time series exhibit a sharply defined component of period 1-year, being also evident for some of them the existence of 4, 6 and 8 periodic components. The 4-year cycle in solar activity, as well as in meteorological elements, has already been reported by Takahashi, 1980. However, the most important conclusion provided by Figs. 5 to 8, is the existence of a cycle of about 10 years in all meteorological parameters which can probably be attributed to solar activity. To clarify this point it is convenient to consider the phases of the several harmonics which seem more likely to be present in atmospheric behaviour. In Table 5 are shown the values of the phase angle θ, in degrees, for the 4, 6, 8 and 10 years periodic components of solar activity and of relative sunshine duration, precipitation, temperature, relative humidity and the wool production, respectively for Elvas and Portalegre.

Table 5. Phase angles, in degrees, of harmonic components of sunspot, relative sunshine duration, precipitation, temperature, relative humidity and wool yield time series for Elvas and Portalegre.

PERIOD (YEARS)	SSN	RSD		P		T		RH		WOOL YIELD	
4	86	87	80	62	46	86	-84	-85	79	77	69
6	-16	-59	-47	78	82	-61	-58	-64	-63	58	59
8	88	51	56	-10	-9	46	49	37	30	-38	-46
10	73	71	71	66	63	70	69	73	70	69	69
		E	P	E	P	E	P	E	P	E	P

From Table 5, it can be concluded that only the approximately 10-year cycle presents values of the phase angle θ with the same sign and order of magnitude for all the considered time series. This fact strongly indicates that this cycle probably depends upon solar effects exerted on the earth's atmosphere. Nevertheless, this indication needs to be confirmed by testing longer time series than those used in the present work, respecting to more stations that significantly cover certain climatic regions. Whether the effects can be generalized to a planetary scale is another question. In any case, it must always be kept in mind that the risk of studying pseudo-cycles can only be eliminated when a physical mechanism, for the probable link between solar activity and atmospheric phenomena, becomes firmly established.

In order to consolidate the above conclusions the smoothed spectral estimates of the cyclical index numbers, relative to the sunspot and the four meteorological time series, at Elvas and Portalegre, were computed from their respective autocorrelation functions, truncated at different lags. For smoothing purposes, the so-called Hanning procedure was chosen. The results obtained are shown in Figs. 9 through 17, where the logarithms of the smoothed spectral density estimates are plotted against period. It should be noted that, due to the shortness of the samples' lengths, truncation points of the autocorrelation functions had to be kept small to avoid exaggerated confidence intervals, resulting in large bandwidths and, consequently, in appreciably smoothed spectra. Hence, it cannot be expected that peaks, in the frequency band of interest in this study, should appear clearly. However, the results shown in Figs. 9 through 17 do not invalidate, rather confirm, that it is in the same frequency band that the spectral contributions, both for solar activity and meteorological data, are higher in magnitude. Moreover, the spectra of the cyclical additive components, as well as of the original observed time series, respecting to the same parameters, were computed and the results, not shown here, led to the same conclusions. These latter spectra also revealed that the spectral density functions in the frequency band corresponding to

long period oscillations, although not being high, seem to be sufficiently important not to be ignored in a description of the global processes.

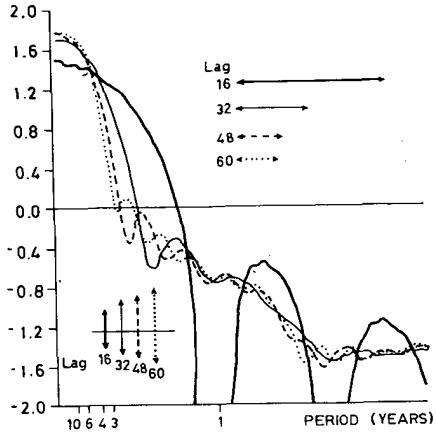


Fig.9-Logarithm of the sample spectral density function of cyclical index numbers C_1, respecting to the sunspot time series, as a function of period, for different truncation points (16,32, 48 and 60 months).

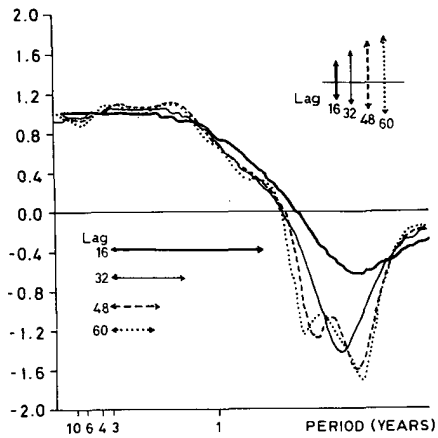


Fig.10-As in Fig. 9, for relative sunshine duration, at Elvas.

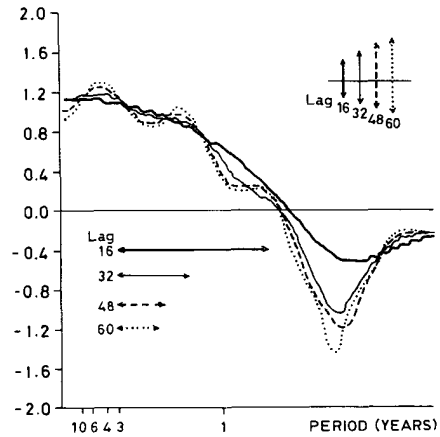


Fig.11-As in Fig. 10, for precipitation.

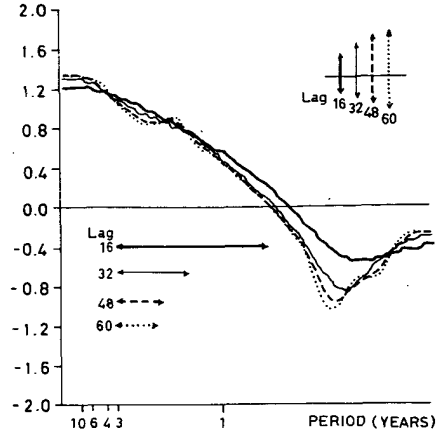


Fig.12-As in Fig. 10, for temperature.

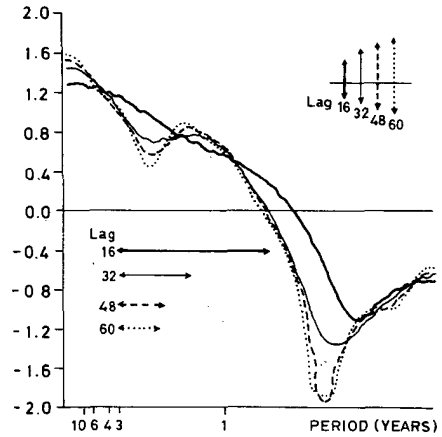


Fig.13-As in Fig. 10, for relative humidity.

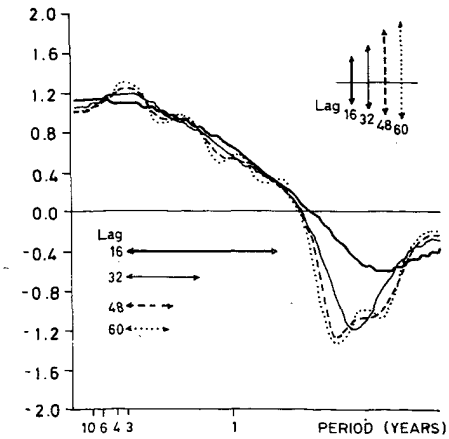


Fig.14-As in Fig. 10, for Portalegre.

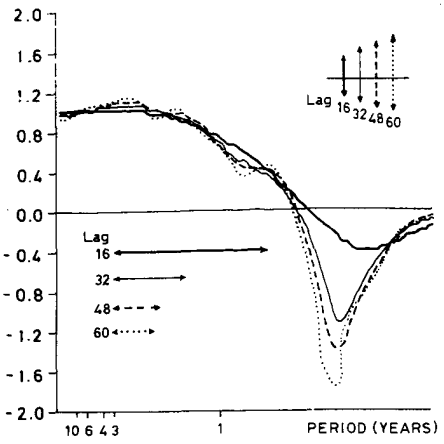


Fig.15-As in Fig. 14, for precipitation

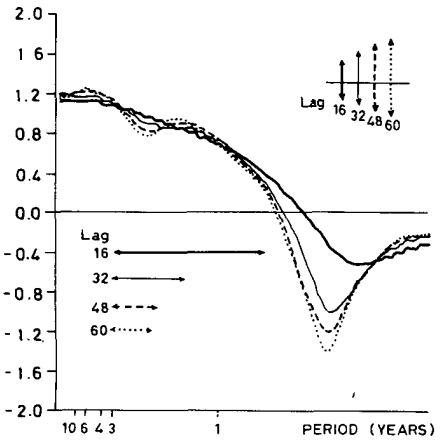


Fig.16-As in Fig. 14, for temperature.

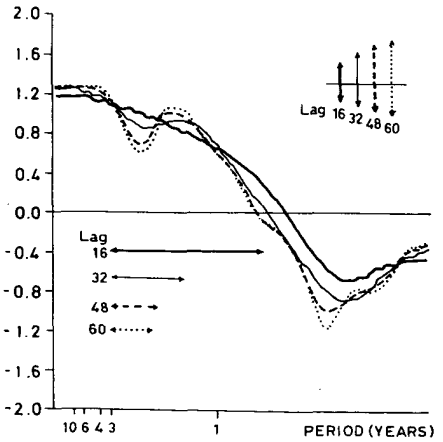


Fig.17-As in Fig. 14, for relative humidity.

3.3 The Climatic Year in the Alto-Alentejo Province of Portugal respecting to precipitation

Precipitation is one of the meteorological elements of decisive importance in agriculture and especially in the Alto-Alentejo, whose production of wheat is of significance to the Country's economy. Hence, it would be desirable to try find what could be the beginning of the climatic year respecting to precipitation. For this purpose, a statistical criterion which could be proposed, is that beginning of the climatic year occurs in the season for which the time series of precipitation values show higher correlations with the homonymous series of subsequent seasons. Of course, the general validity of this criterion would stand, only if the conclusions drawn following it do not depend either on the length of the chosen time interval, provided long enough to guarantee the significance of a statistical test, or in its starting point. However, it could be asked whether the criterion could apply locally for limited periods, during which certain climatic effects predominate.

Correlations between precipitation seasonal mean values were calculated for sections of 15,20 and 25 consecutive years, differently distributed along the complete available record of 34 years, this one being also considered.

However, the significance t-test at the 5% level led, in practically all cases, to the non rejection of the null hypothesis of the considered series being uncorrelated. Moreover, examination of dispersion diagrams did not indicate the possible existence of non-linear correlations between the time series of seasonal precipitation mean values. Interestingly, the same analysis, carried out respecting the other three meteorological parameters considered in this study, also pointed to an identical conclusion. If the few cases, where appreciable correlations were found to exist, are merely a sample effect or have a deeper meaning, is a question which remains to be answered.

However, the failure of the attempt of defining the beginning of the climatic year, based on the proposed statistical criterion, has the positive counterpart of showing, contrary to empirical wisdom, that anomalies observed to occur in one season do not determine the behaviour of subsequent seasons.

ACKNOWLEDGEMENTS

The authors are greatly indebted to Professor J. A. Serra for having suggested this study, for his stimulating comments and ideas, namely those concerned with the last section of this paper, as well as for his assistance in the preparation of the manuscript.

REFERENCES

Bowerman, B.L. & R.T. O'Connell, 1979: Time Series and Forecasting, Duxbury Press.

Serra, J.A. and R.M. Albuquerque de Matos, 1983: Produção Portuguesa de lãs finas. Public. Junta Nacional dos Produtos Pecuários, Lisboa, Série B, nº 13 (in publication).

Takahashi, K., 1980: The Four-Year Cycle in Atmospheric and Solar Phenomena. Statistical Climatology, Developments in Atmospheric Science, 13, Edited by S. Ikeda et. al., Elsevier Scientific Publishing Company.

APPROACH TO ESTIMATING CLIMATE FLUCTUATION IN RELATION TO AGRICULTURE

Felix N. Kogan

Department of Atmospheric Sciences
University of Missouri
Columbia, MO 65201, U.S.A.

1. INTRODUCTION

Fluctuations of weather in the world and the tendency towards a negative balance between food production and consumption during the last 10-15 years has stimulated research on the problem of climate change and its effect on agricultural production. Presently, there is no satisfactory general theory of climate that can be used to reduce uncertainty in climate forecasting (Marchuck et al., 1981). Accordingly, judgements on the possible climate change (or climate fluctuation) in the near future are usually based on the historical temperature records of the past 100 years. The mean annual temperature and sometimes mean winter and summer temperatures averaged over the entire Earth, hemisphere, continent or certain band of latitudes are normally investigated. Unfortunately, it is difficult to apply the results of such investigations to the assessment of the impact of climate change on agriculture. The point is that the amount of agricultural production is the result not only of climate fluctuations but also of the impact of the whole ecological complex including soil, vegetative biocenosis, and also economic and human activities. Concerning the climate itself, it is necessary to note that the reaction of crops to climate impact is normally reflected over a period shorter than one year. Calculations show that substantial changes in the annual temperature for some areas result from changes in the temperature for part of a year which sometimes play little or no role in altering crop production. It is suggested that climate fluctuation be analyzed in relation to agriculture over the periods of the year when the weather has the most pronounced impact on crops, and for large regions which are relatively homogeneous.

2. METHOD

Periods of the year in which weather has a strong impact on crops were identified, and the values of this impact were calculated based on an environment-crop model. This model was constructed by means of studying relations between crop yield and environmental factors over the period 1945-1980. After model verification, it was run on the weather data available for the period 1891-1944. Finally, analysis of climate fluctuation and crop response was conducted for the past 90 years (1891-1980).

The approach of constructing an environment-crop model and isolating the weather-related input into yield involves the elimination of deterministic components from crop production, particularly areal, ecological, and technological (Kogan 1975, 1983).

To eliminate an areal component, crop production per unit of sown area (yield) was used. The impact of soil and other ecological factors on crop were diminished in the process of data aggregation. These data were collected for a region with a certain type of soil and vegetative biocenosis so that their impact on yield was considered approximately constant over the years. In an attempt to eliminate the most important long term input upon yield variance, a time-trend in the yield series was used as a surrogate to quantitatively estimate the technology input into crop yield.

Correct separation of weather-related yield variance depends very much on the accuracy of the trend approximation. This requires a long enough yield series (at least 25 years) to facilitate a proper choice of trend shape and hence correct estimation of the rate of trend change. This estimate gives qualitative assessment of climate-technology interaction impact on yields. Normally, a declining rate indicates that the applied technology is not fully compatible with the existing climate; conversely, an increasing rate shows compatibility between technology and climate. This thesis can be demonstrated by Fig. 1, where the regional USSR barley yields over the period 1945-1978 are presented. In Kazakh and North Caucasus regions where climate, with quite limited water resources, has not been very complimentary to the applied technology, the rate of yield growth declined over time (convex trend). In Belorussia and Volgo-Vyatka regions with less limiting water resources, climate appears to have been more compatible to the applied technology and has resulted in an increase in the rate of yield growth (concave trend).

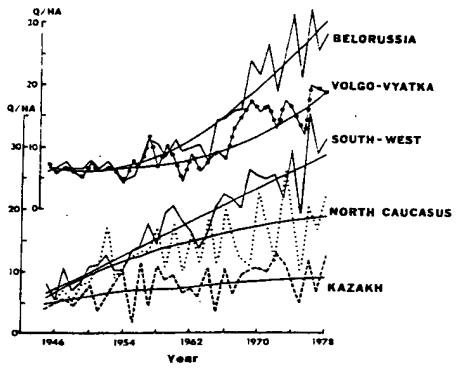


Fig. 1. Barley yield and trend in USSR regions, 1945-1978.

Climate-technology interaction and the nature of its impact on yields suggests the need for a non-linear approximation of trend. For yield series which cover one epoch of changes in technology (for example, "green revolution"), the second degree polynomial approximation of trend is considered satisfactory. Correct estimation of the weather related yield variance requires also the more accurate approximation of trend in the latest few years, particularly in cases where yields in these years were very changeable due to weather fluctuations. A technique was developed to stabilize regression coefficients of the polynomial approximating trend if extreme variations of yields took place in the last few years.

Under the aforementioned conditions, the weather related values of yield can be presented by the following equation:

$$\Delta Y = Y/\hat{Y} \qquad (1)$$

$$\hat{Y} = \alpha_0 + \alpha_1 \tau + \alpha_2 \tau^2 \qquad (2)$$

where Y is actual yield; $\hat{Y}$ is calculated yield as a function of year number τ (trend); ΔY is weather related ratio of the deviation of yield from trend (detrended yield); the α's are regression coefficients in polynomial approximating trend.

The relationship between detrended yields and weather can be presented in the following general form:

$$\Delta\hat{Y} = f\,(P_j, T_k) + e \qquad (3)$$

where P_j, T_k are weighted average precipitation and temperature correspondingly for a season of the year with the same sign of the impact of weather variable on crop yield; j and k are season's number; e is error.

To calculate P_j and T_k the following equations are applied:

$$P_j = \sum_s P_{js} * V_{js} \qquad (4)$$

$$T_k = \sum_q T_{kq} * V_{kq} \qquad (5)$$

$$V_{js} = r^2_{js} \Big/ \sum_s r^2_{js} \qquad (6)$$

$$V_{kq} = r^2_{kq} \Big/ \sum_q r^2_{kq} \qquad (7)$$

where P_{js}, V_{js} are precipitation and their weight correspondingly for month number s of subperiod j; T_{kq}, V_{kq} are temperature and weight correspondingly for month number q of subperiod k; r^2 is the coefficient of determination obtained by regressing the detrended yield ratio ($\Delta\hat{Y}$) on regional total monthly precipitation and regional average monthly temperature; subperiods contain several months from the period from January (number 1) through December (number 12) of the year preceding the year of crop harvest and continued from January (number 13) through July (number 19) of the crop harvest year.

Simulated yield deviation from trend (equation 3) and weighted weather (equations 4 and 5) were smoothed out by five-year moving average over the whole period of simulated data (1891-1981). The results were analyzed in terms of:

* determining periods of the year when climate has a strong impact on crop;
* identification of tendencies in the dynamics of detrended yields and in the weighted precipitation and temperature variables;
* determining the range of climate change able to alter a tendency in yield dynamics;
* comparison of the results with the results obtained in statistical analysis of the global and latitudinal climate change over the past 100 years available from recent publications.

The proposed technique was applied to analyze climate and cereal yield changes in Kazakh, one of the largest grain producing regions in the USSR. Grain here occupies around 25 million hectares of land. This grain consists mostly of spring wheat, spring barley and oats. The sown area extends from 47° to 53° N and from 50° to 80° E. This region belongs to the steppe zone which has mostly chernozem and chesnute soils. Climate of the area is sharply continental. A specific feature of climate in relation to agriculture is inadequate annual water balance, particularly in the first half of the year.

Regional weather was aggregated from twelve meteorological stations which had records of observations on total monthly precipitation and average monthly temperature since 1891.

4. RESULTS

4.1 Modeling.

Analysis of the cereal crops yield series in Kazakh, shown in Figure 2, indicated that in the period from 1945 to 1980 technology input (trend) explained about 28 percent of the yield variance. The average annual yield increase resulting from the advancing technology amounted to only 0.2 quintals per hectare annually. Whereas weather and other factors changed yields between adjacent years up to 8 quintals per hectare. The interaction between climate and technology over time caused the rate of yield growth to decrease with the speed 0.0021 $q.HA^{-1}.YEAR^{-2}$. The absolute variation of yields around the trend seems to be gradually increasing as time goes by.

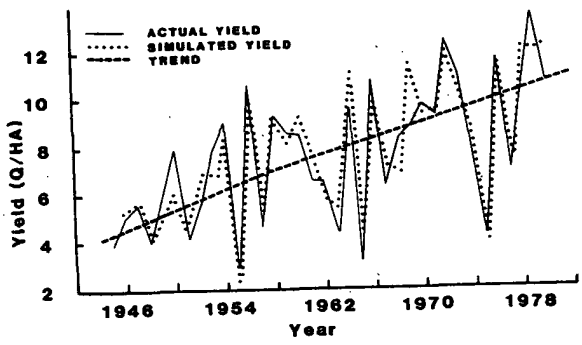


Fig. 2. Cereal crops yields in Kazakh, USSR, 1945-1980.

The constructed cereal crop-weather-trend model was verified and showed a good correspondence between simulated and actual yields (Fig. 2). The weather related part of the equation describing yield was presented in the following numerical form:

$$\Delta\hat{Y} = 94.65 + 2.88P_A - 0.80P_B - 5.32T_a - 5.03T_b - 0.36T_b^2 \quad (8),$$

where

$$P_A = 0.175P_{12} + 0.1P_{13} + 0.207P_{16} + 0.211P_{17} + 0.306P_{18} \quad (9)$$

$$P_B = 0.099P_5 + 0.696P_6 + 0.205P_9 \quad (10)$$

$$T_a = 0.167T_{13} + 0.074T_{15} + 0.362T_{17} + 0.207T_{18} + 0.191T_{19} \quad (11)$$

$$T_b = 0.565T_3 + 0.089T_4 + 0.346T_{11} \quad (12)$$

where P_A, P_B, T_a, T_b are weighted average precipitation (P) and temperature (T) for subperiods A, B, a and b.

As seen in equation (8), P_A and T_a have considerable input into yield fluctuation around trend. A distribution of this input during the year is shown in Fig. 3. As seen, this input is not uniform over months. The amount of precipitation is very important for yield in April through June (month numbers 16-18). This period should be taken into consideration first in analysis of the climate change impact on cereals in Kazakh region. Also, additional information can be obtained through preseason precipitation, particularly in December and January (months 12 and 13). The most pronounced temperature impact on cereal crops normally occurs in May through July (Fig. 3 and month numbers 17-19 in equation 11). Preseason temperature also makes some contribution to the cereals yield ratio departure from trend.

4.2 Long Term Yield and Climate Change

Figure 4 shows the past 90 years simulated cereal yields departure from trend, and also weighted precipitation (P_A) and weighted temperature (T_a) for the periods in which these crops

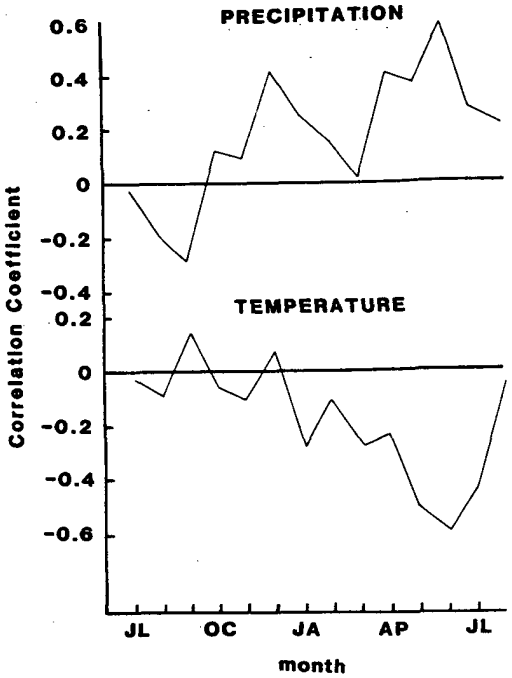


Fig. 3. Pearson correlation coefficients between cereals yields departure from trend with monthly total precipitation and average temperature, Kazakh, 1945-1980.

are the most sensitive to weather change. These data were smoothed by means of five-year moving averages.

As seen, the general tendency in the five-year moving average indicates considerable decrease in yield deviation from trend beginning at the end of the previous century to the mid 1940's. During the period (around 50 years) yield dropped by 20 percent. After a short period of stability in the general tendency, a slight increase in simulated yield (around 5 to 8 percent) was observed in the most recent 15-20 year period. This general tendency in cereal crop yield departure from trend growth was primarily associated with the corresponding general tendency in precipitation change during the most important period of the year (Fig. 4). The amount of precipitation in this period of the year in Kazakh dropped by almost 20 percent from the 1890's to the 1940's. In the past 15 to 20 years the amount of precipitation in Kazakh started to increase slightly causing a corresponding increase in simulated cereal yields. The general temperature dynamics were also in agreement with yield dynamics although this agreement is less pronounced as compared with precipitation. Concerning the short periods (5 to 15 years), changes in weather (mostly precipitation) of a more than 10 percent deviation from normal cause detectable changes in the simulated yield ratio departure from trend (Fig. 4).

Unlike the seasonal precipitation dynamics discussed earlier, the annual precipitation dynamics shown in Fig. 5 are less similar to the dynamics of simulated yields (Fig. 4). In the first part of the 1891-1980 period, simulated

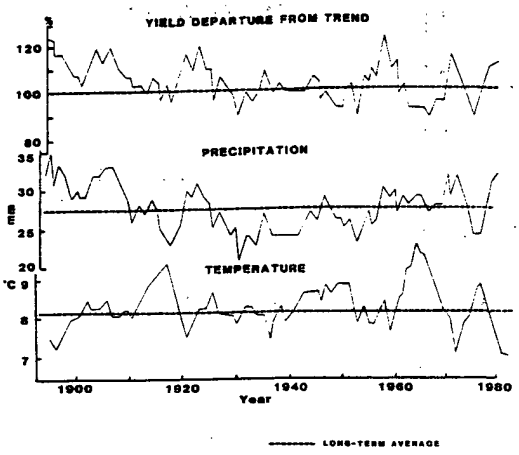


Fig. 4. Five-year moving average for yield deviation from trend, weighted precipitation in April-June and temperature in May-July, Kazakh, 1891-1982.

yields had sharp decrease (Fig. 4) whereas the rate of annual precipitation decrease was slow (Fig. 5). The opposite picture was observed in the second part of the period: yields had a slow increase but annual precipitation had a fast increase. Considerable difference between these two curves exists in the period from 1930 to 1948 when annual precipitation was much below normal (up to 20 percent), but yields were near normal or just above normal. This lack of correspondence between simulated yields and annual precipitation is connected with the specific distribution of precipitation during the seasons of the year. As seen in Figure 5, a sharp increase in annual precipitation in the 1950s and 1960s was connected with the increase of precipitation in the second half of the year when cereals are less responsive to water deficits (Fig. 3). Low annual precipitation in 1930-1948 was primarily associated with below normal precipitation in October through December.

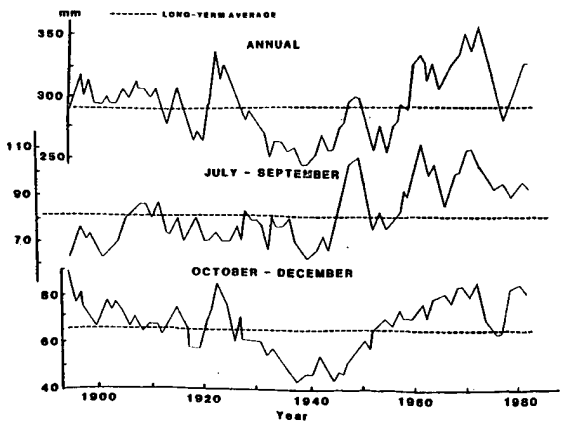


Fig. 5. Five-year moving average, total precipitation, Kazakh, 1891-1982.

The same problems exist in the case of application of annual temperature to explain long-term simulated yield dynamics. For example, in the 1960s the annual temperature in Kazakh was 1° to 1.5° higher than normal with simulated cereal crop yield departures near normal. This high annual temperature had a low impact on crops because the increase in annual temperature was primarily caused by higher temperatures in January through March, which does not normally limit cereal crop yields in Kazakh.

More problems appear when using global (latitudinal) annual temperature dynamics to explain its possible impact on crop production change in a particular big region. Comparison of the annual temperature dynamics in temperate zone (57.5° - 37.5° N) (Vinnikov et al., 1980) with the dynamics of simulated yield departure from trend indicates that the only period with corresponding dynamics was from the end of the previous century until the 1940s. But this happened because the annual temperature increase observed in this period was connected primarily with the temperature increase in the period of year when temperature has the strongest impact on cereal crops (Fig. 3). In the past ten years, the average temperature of the Northern Hemisphere and different latitudinal zones is getting warmer (Vinnikov et al., 1980). So far, there has been no indication of the change of cereal crop yield dynamics in Kazakh. Unfortunately, it was impossible to make similar analysis for precipitation as the average global precipitation is not available. But comparison of regional precipitation data with average global precipitation for separate months (Gruza et al., 1981) showed many more contradictions in their dynamics than the temperature records showed.

5. CONCLUSIONS

The developed technique allows us to identify periods of the year when a crop is the most sensitive to fluctuations of weather and estimate quantitatively the impact of cumulative weather on crops. Application of this technique to long-term weather records permits one to extend yield records over longer periods (even though they must be qualified as simulated yields) and thus examine long-term dynamics in yields. Long-term dynamics in yields provide opportunity to analyze possible ranges of climate fluctuations responsible for changes in yield dynamics.

6. ACKNOWLEDGEMENTS

This work is a part of the project supported by a cooperative research agreement with the Statistical Reporting Service of the USDA. The author has benefited from useful comments on the topic of this paper given by Wendell W. Wilson with whom the author is working in the Yield Evaluation Section (USDA).

7. REFERENCES

Gruza, G. V., E. G. Apasova, 1981. Climatic Variability in the Monthly Sums of Precipitation in the Northern Hemisphere. Soviet Meteorology and Hydrology No. 5, pp. 1-8.

Kogan, F. N., 1975. On the Assessment of Weather Influence on Cereal Crops in Different Soil and Climatic Zones of the USSR. In "Programming of the Crop Production," published by Kolos Moscow, pp. 236-244 (in Russian).

Kogan, F. N., 1983. "Environment-Crop Models and Considerations for Their Utilization." 16th Conference on Agriculture and Forest Meteorology, American Meteorological Society, April 26-28, Fort Collins, Colorado, pp. 106-109.

Marchuck, G. I., K. Kondratiev, V. Dimnikov, 1981. Achievements in Meteorology and Climatology, Vol. 7, Moscow, p. 130.

Vinnikov, K. Ya., G. V. Gruza, V. F. Zakharov, A. A. Kirillov, N. P. Kovyneva, E. Ya. Ran'kova, 1980. Current Climate Changes in the Northern Hemisphere. Soviet Meteorology and Hydrology, No. 6, pp. 1-10.

ARIDITY FLUCTUATIONS OVER CENTRAL SOUTH INDIAN REGION

A.A.L.N. Sarma and M. Ravindranath
Dept. of Meteorology & Oceanography,
Andhra University, Waltair
Visakhapatnam 530 003, S. India.

1. INTRODUCTION

Studies on the prevalence of aridity alongwith its magnitude in different climates is of utmost importance in understanding the potentialities of any region. In the present paper an attempt has been made to obtain the aridity index, from the water balance concept, at selected stations of central south Indian region. The aridity index is defined as the percentage ratio of the annual water deficiency to the annual water need. The variations in aridity are as a result of variations in the strength and depth of the circulation pattern of the S-W and N-E monsoons that vary in space and time. The realm of N-E monsoon is small. The root cause of the aridity lies in the magnitude of water deficit. The geographical setting of the south Indian region is such that it encompasses the complete climatic spectrum (perhumid, humid, moist subhumid, dry subhumid, semi-arid and arid climates) as one marches from west to east (Subrahmanyam and Sarma, 1975). The present paper is an endeavour to study the trends in aridity at selected stations of the region.

Stable climatic provinces together with the appropriate land use and land management practices yield good returns. The moisture status of the different climatic zones depend upon the relative magnitudes of the corresponding humidity and aridity indices which in turn entirely depend on the vagaries of the monsoon systems. The humidity index is defined as the ratio of annual water surplus to annual water need, expressed in percentage. The moisture index serves as the basic input not only in knowing the distribution and occurrence of the different climatic types but also in differentiating the stable climatic zones from the unstable ones.

2. MATERIAL AND METHODS

The revised bookkeeping procedure of Thornthwaite and Mather (1955) is followed in obtaining the basic water budget elements at about forty (40) selected stations that are well distributed over the region. Humidity, aridity and moisture indices are obtained for all the selected stations. And in studying the trends in aridity in the climatic spectrum of the region six (6) stations are chosen among the 40 stations by selecting not less than one station from each of its zones. To reveal the aridity trends at these six stations, the exponential smoothing technique is followed as suggested by Sarma and Ravindranath (1982). The details of this technique are omitted here for space considerations and the expression for exponential smoothing is

$$V_S = V_{S-1} + (D - V_{S-1})$$

where V_S = Newly smoothed value,

V_{S-1} = Previously smoothed value,

D = Next data point and

α = Smoothing factor.

The so obtained smoothed values of aridity index from the above expression are considered and departures of the aridity index from the mean are computed and expressed as a percentage ratio of the range of aridity index at a station. Standard deviation is used to differentiate the different grades of aridity and the scheme that was followed by Sarma and Ravindranath (1982) is made use of. The scheme of classification is as follows:

$< \sigma$	Slight
σ to $< 2\sigma$	Moderate
$\geq 2\sigma$	Large

The occurrence and frequency of different grades of moisture status (climatic types) at a station is computed by following the year climate concept. Yoshino and Urishibara (1981) reported the distribution and occurrence of year climates over Eastern Asia through the year climate concept based upon the Koppen's scheme of climatic classification. In the present report the occurrence and the pattern of the different

climatic types over the central south Indian region is found by adopting the rational classification of Thornthwaite (1948). In delineating the different climatic types, the revised expression of Carter and Mather (1966) is made use of. To determine the frequency of a particular grade of moisture status at a station, the number of occurrences of that grade of moisture status has been expressed as a percentage ratio of the total number of observations. Accordingly the frequency of the different grades of moisture status during the period of investigation have been found for all the selected stations.

3. DISCUSSION AND RESULTS

3.1 Aridity trends in moist climates

Fig.1 depicts the march of aridity index in moist climates.

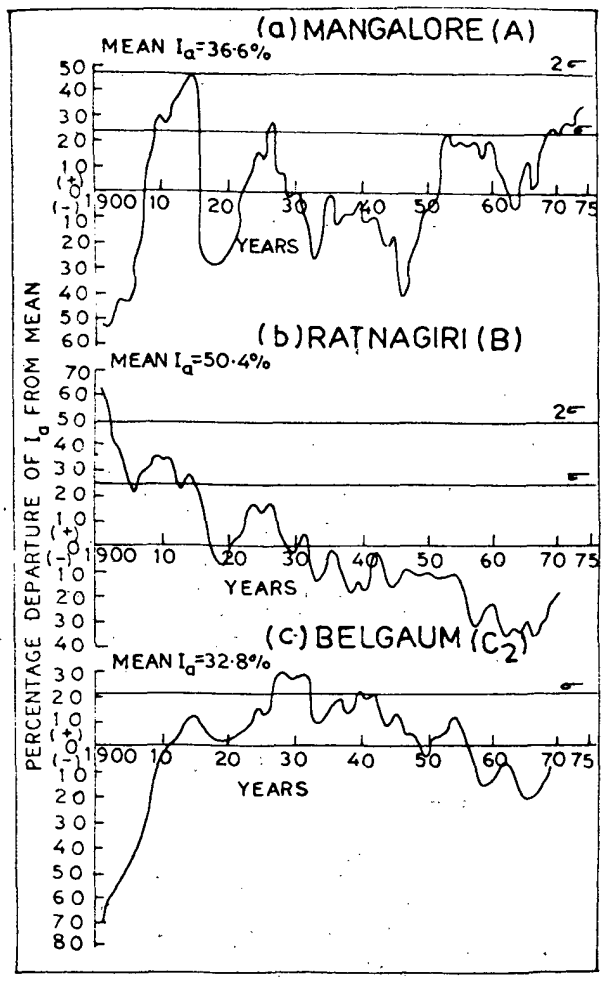


FIG.1. MARCH OF ARIDITY INDEX- MOIST CLIMATES.

Out of the three stations that are drawn from these zones, only Ratnagiri (Fig.1b) with its highest mean aridity (Table) recorded a decreasing trend from the beginning of the study period to the end. But on the otherhand Mangalore (Fig.1a) with the lowest standard deviation value (Table) exhibited wide undulations about the mean. Very interestingly, Belgaum (Fig.1c) somewhat an interior station exhibited an increasing, levelling and decreasing aridity trend during 1901 to 1927, 1928 to 1942 and 1943 to 1969 respectively.

Table. Aridity index - Climatic spectrum

Station	Climatic type	Symbol	Aridity index		Standard deviation expressed as % of range
			Mean	Range	
Moist Climates					
Mangalore	Perhumid	A	36.5	5.6	23.6
Ratnagiri	Humid	B	50.4	9.4	24.0
Belgaum	Moist subhumid	C_2	32.7	14.0	21.2
Dry Climates					
Nizamabad	Dry subhumid	C_1	44.9	6.8	30.4
Kakinada	Semiarid	D	47.6	10.0	24.1
Bellary	Arid	E	69.0	6.8	25.7

3.2 Aridity trends in dry climates

The dry zones are characterised by violent variations accompanied by sudden jumps in aridity (Fig.2). Inspite of the fact that Nizamabad and Bellary belong to different climatic zones, they recorded the same range of aridity (Table). Features such as increase, dip, levelling, jump and decline in aridity could be known from Fig.2. A remarkable jump in aridity is noticed at Nizamabad and Kakinada from 1950 onwards, whereas the same feature is observed at Bellary from 1918 to 1932. The jumps in aridity are substantial in magnitude and is varied from -55 to +40, -45 to +55 and -39 to +42 at Nizamabad, Kakinada and Bellary respectively.

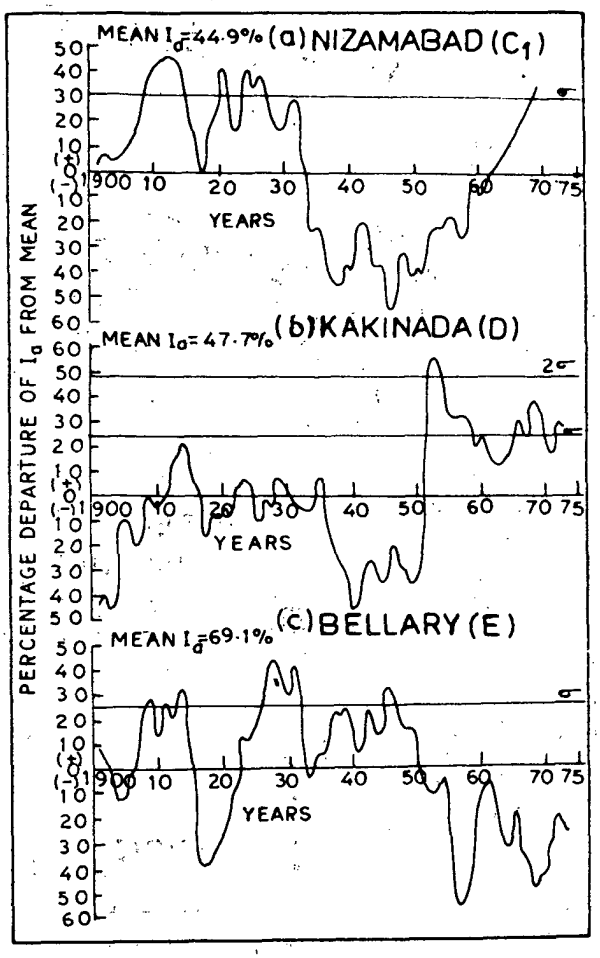


FIG.2. MARCH OF ARIDITY INDEX-DRY CLIMATES.

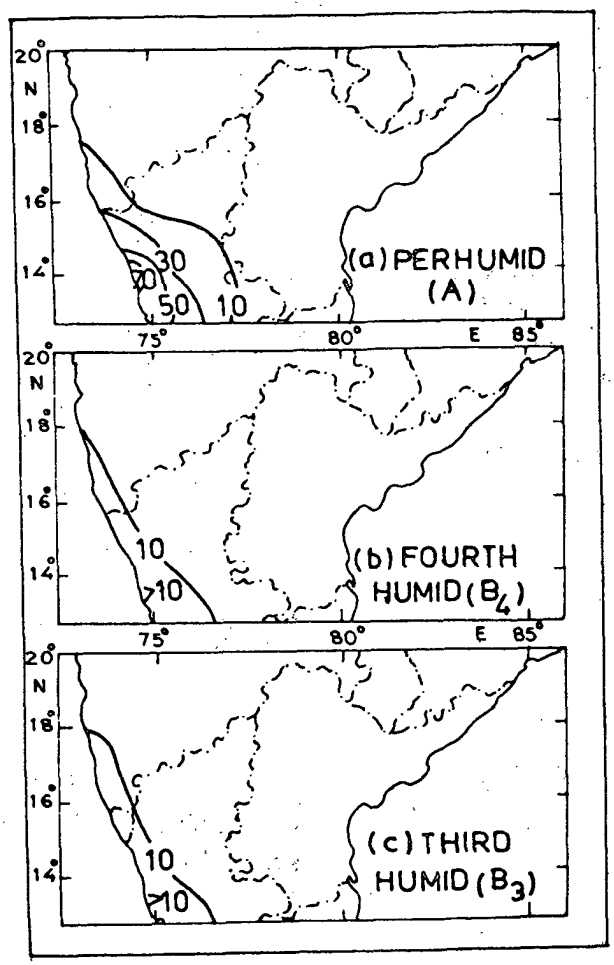


FIG.3. DISTRIBUTION AND OCCURRENCE OF CLIMATIC TYPES.

4. OCCURRENCE AND DISTRIBUTION OF CLIMATES

4.1 Perhumid climates

Fig.3 describes the pattern of various climates over the central south Indian region. The perhumid zones prevailed right on and along the west coast and the frequency is very rapidly decreased as one moves away from the coast. An isolated pocket of high frequency (> 70%) is found at Hanover (Fig.3a). Hanover is considered as a stable one since more than 70% of the time the station experienced perhumid type of climate. Any station that belongs to any climatic type if experiences more than 70% of the time, the most prevailing climatic type, could be marked as a stable one. The entire east coast and its adjoint interior is free from perhumid climates.

4.2 Humid climates

A narrow zone of 4th humid category of climates is observed along the west coast with a frequency that varied from 10 to 20% (Fig.3b). The 3rd humid type climates extended in the form of a tube running north to south longitudinally along the west coast is prevailed with a frequency of 10 to 19% (Fig.3c). The 2nd humid climates occurred with a frequency of 39% at Ratnagiri (Fig.3d). The frequency of 1st humic type climates is low and appeared as isolated pockets (Fig.3e). A noticeable feature of Fig.3e. is that the B_1 type climates is also observed on the extreme northeastern portion of the region.

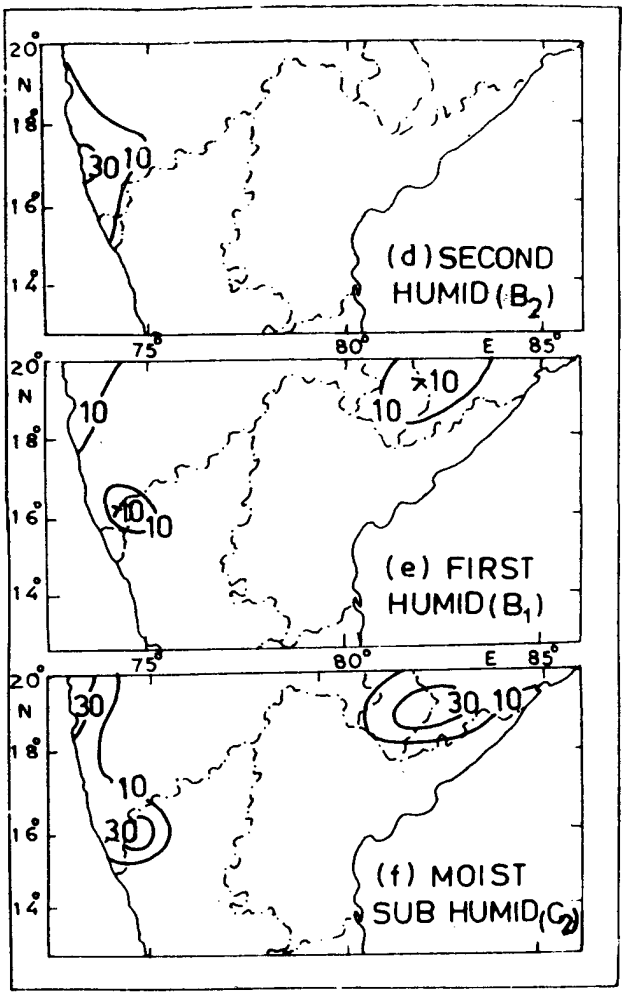


FIG.3. CONTINUED.

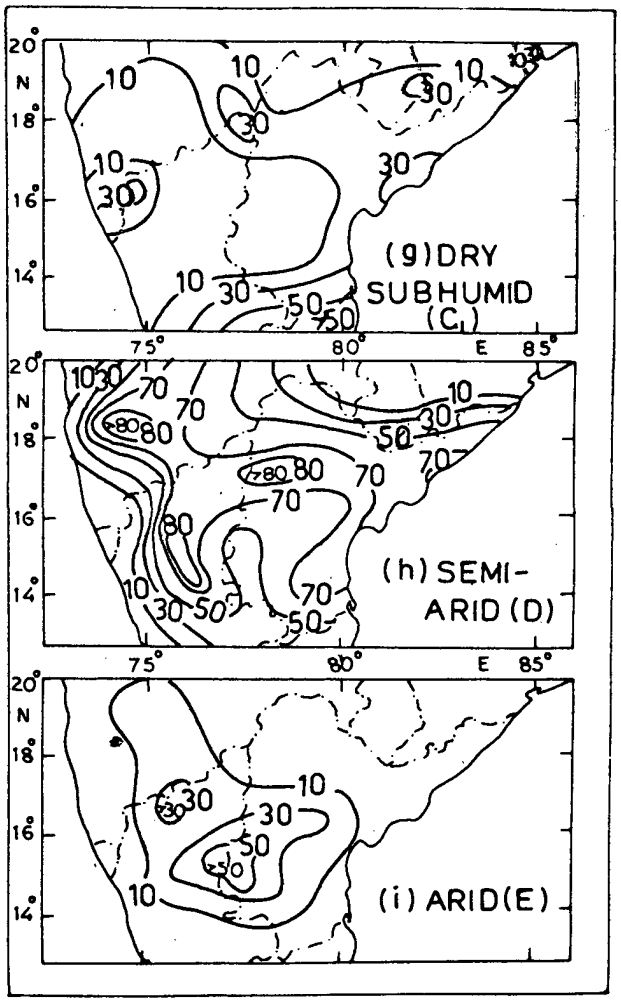


FIG.3. CONTINUED.

4.3 Subhumid climates

The realm of moist subhumid type is not only extended from north of Karwar to Bombay along the east coast but also found at and around Jagadalpur (Fig.3f). The spread of dry subhumid climates is all over the region and the frequency varied from 10% to greater than 50%. It is worthwhile to mention, the prevalence of dry subhumid climates, with a mild frequency on the west coast. But the frequency of the dry subhumid climates along the east coast and its adjoins is very high compared to west coast.

4.4 Semiarid climates

A very narrow zone that is very close to the west coast is free from the clutches of semiarid climates (D) (Fig.3h). The frequency of D type climates is very high as one moves to the interior of the region. The frequency of D type climate is greater than 80% at Poona, Gadag and Hyderabad and appeared as isolated pockets. The area bounded by 70% and 80% frequency is considered as a stable domain for D type climate (Fig.3h). Visakhapatnam a solitary station on the east coast maintained a stable moisture status that belongs to D type climate.

4.5 Arid climates

The core of arid climates is located at and around Bellary with a west to east spread together with a protuberence into the north-west direction (Fig.3i). The core which is situated at and around Bellary with a frequency that is greater than 50% is diminished very rapidly to 10% at the periphery.

5. REFERENCES

Carter, D.B. and Mather, J.R., 1966: Climatic classification for environmental biology. Publ. in Clim., XIX, 341-352.

Sarma, A.A.L.N. and Ravindranath, M., 1982: Incidence of aridity waves in relation to southwest monsoon. Presented in the semirar on Hydrological investigations during the last 25 years in India, held at Dept. of Geophysics, Andhra University, Waltair, India.

Subrahmanyam, V.P. and Sarma, A.A.L.N., 1975: Incidence of aridity and droughts in the climatic spectrum of the south Indian region. Trop. Ecol., 16, 120-127.

Thornthwaite, C.W., 1948: An approach toward a rational classification of climate. Geogr. Rev., 38, 55-94.

Thornthwaite, C.W. and Mather, J.R., 1955: The water balance. Publ. in Clim., 8, 100-104.

Yoshino, M.M. and Urishibara, K., 1981: Regionality of climatic change in East Asia, Geo. J., 5, 123-132.

OPTIMAL USE AND ECONOMIC VALUE OF PRIOR DATA, CATEGORICAL, AND PROBABILISTIC CLIMATE FORECASTS*

Roman Krzysztofowicz

Department of Systems Engineering
University of Virginia
Charlottesville, Virginia 22901

1. INTRODUCTION

The hypothesis that climatological information has economic value, though not always explicit, is basic to climate research, data collection, and forecasting. In a strict sense, information acquires value when it affects decisions having economic impact. The value acquired depends not only upon the quality and timeliness of information, but also upon economic and behavioral characteristics of decisions. Therefore, study of information value is tantamount with study of information-decision systems.

In this paper, we discuss some of the fundamental aspects of information-decision systems from a Bayesian viewpoint. The focus is on the optimal use and economic value of climatological information such as prior (historical) data, categorical, and probabilistic long-range forecasts. Probabilistic forecasts are the ones wherein the forecaster quantifies explicitly his degree of uncertainty, for instance, in terms of credible intervals about the median. Each forecasting scheme, (1) categorical and (2) probabilistic, is evaluated in a coupling with two decision procedures: (1) an optimal (Bayesian) procedure which accounts for forecast uncertainty and makes the full use of prior climatic data, and (2) a popular, but suboptimal, procedure which disregards forecast uncertainty and prescribes decisions as if the forecasts were perfect or perfectly calibrated. Thus, four forecast-decision systems are evaluated. The bounds on the values of these systems are established by the values of naive and perfect forecasts coupled with optimal decisions.

The main conclusions from this paper shed some light on answers to the following three sets of questions: (1) What opportunity losses are incurred when users of climate forecasts ignore forecast uncertainty? How do these losses vary with the forecast skill? (2) Why users who want to make optimal decisions on the basis of climate forecasts need also prior climatic data, or probabilistic models of such data? What is the economic value of prior data? How is this value related to the forecast skill? (3) Why is it important to explicitly quantify uncertainty in climate forecasts? What economic gains are to be expected from probabilistic forecasts? How are these gains impacted by the forecast skill?

*This material is based upon work supported by the National Science Foundation under Grant No. CEE-8300928.

2. FORECAST-DECISION SYSTEMS

2.1 Decision Problem

The following very general decision problem will serve as a vehicle for our analyses. Suppose an uncertain future value of a climatic variable θ (such as monthly or seasonal average temperature, total precipitation) is of concern to a decision maker who must make a decision a and whose loss will be determined according to a real-valued function $\ell(\theta,a)$, where θ and a are continuous, possibly constrained, variables. A long-range forecast of θ with a specified and fixed lead time is available every time the decision is made. The decision maker is assumed to minimize the loss or the expected loss (risk) depending upon whether he perceives the forecast to be certain (error-free) or uncertain. It is postulated that normatively optimal decision behavior is to account for uncertainty in the forecast.

2.2 Naive and Perfect Forecasts

Naive and perfect forecasts are considered first to establish an upper bound on the value of any real-time forecast. A naive forecast is simply the prior probability density function g of θ estimated on the basis of a historical record. The minimum expected loss (risk) under this density is

$$RN = \min_a \int \ell(\theta,a)\ g(\theta)\ d\theta \qquad (1)$$

A forecast is perfect if it indicates the actual value of θ with probability one. Under such circumstances, the optimal decision minimizes the loss for any given value of θ; the risk is

$$RF = \int \min_a \ell(\theta,a)\ g(\theta)\ d\theta \qquad (2)$$

2.3 Categorical Forecast - Optimal Decision

Let t denote a categorical forecast of θ. To optimally account for the possible forecast error, the forecast has to be processed through Bayes theorem. Towards that end, forecast uncertainty is described by a family of conditional density functions $\{f(\cdot|\theta)$:all $\theta\}$ of t. For any given forecast t, the likelihood function $f(t|\cdot)$ of θ provides a probabilistic description of the forecast error. The prior density g of θ can now be revised to obtain a posterior density

$$\eta(\theta|t) = \frac{f(t|\theta)\ g(\theta)}{\xi(t)} \qquad (3)$$

where

$$\xi(t) = \int f(t|\theta)\ g(\theta)\ d\theta \qquad (4)$$

The prior density g can be estimated on the basis of a historical record $\{\theta_i: i=1,\ldots,n\}$. The likelihood functions f can be estimated from a joint historical record of the forecast and actual values $\{(t_i,\theta_i): i=k,\ldots,n\}$; this record is typically much shorter than the entire historical record of θ ($1 \ll k < n$). Consequently, if this joint record were used to estimate directly a family of conditional density functions $\{\psi(\cdot|t)$:all $t\}$ of θ, one would have to ignore the prior record $\{\theta_i: i=1,\ldots,k-1\}$. Thus, one would find $\psi(\theta|t) \neq \eta(\theta|t)$ since the posterior density η conveys more information than the density ψ. For the implications of using ψ instead of η, see Krzysztofowicz (1983).

Having determined the family of the posterior densities $\{\eta(\cdot|t)$:all $t\}$ of the variable θ, we can solve the decision problem prior to real-time decision making. The optimal solution is a decision function α^* specifying for every forecast t an optimal decision $a^* = \alpha^*(t)$ which minimizes the expected loss. The associated risk is

$$RC = \int [\min_{\alpha} \int \ell(\theta,\alpha(t))\ \eta(\theta|t)\ d\theta]\ \xi(t)\ dt \qquad (5)$$

2.4 Categorical Forecast - Nonoptimal Decision

If the decision maker perceives the categorical forecast t of θ to be error-free, then he employs the criterion:

$$\min_{\alpha} \ell(t,\alpha(t)) \qquad (6)$$

Thus his decision function is $\hat{\alpha}$ ($\hat{\alpha} \neq \alpha^*$), and the associated risk is

$$\overline{RC} = \iint \ell(\theta,\hat{\alpha}(t))\ \eta(\theta|t)\ \xi(t)\ d\theta\ dt \qquad (7)$$

2.5 Probabilistic Forecast - Optimal Decision

A probabilistic forecast of a continuous variable θ specifies a density function of θ. It is assumed that this forecast density belongs to a parametric family $\{\rho(\cdot|\underline{\mu})\}$ with the parameter vector $\underline{\mu}$. The density $\rho(\cdot|\underline{\mu})$ encodes the forecaster's uncertainty about the actual value of θ, and provides a means of conveying this uncertainty to the decision maker. The smaller the uncertainty, the smaller the variance of $\rho(\cdot|\underline{\mu})$.

A measure of the reliability of probabilistic forecasts is their calibration: the forecasted frequency of θ should match the observed frequency. (For precise definition of a calibration function for forecasts of continuous variables see Morris (1977); for examples and discussions see Winkler and Murphy (1979), and Alpert and Raiffa (1982).) To optimally account for imperfect calibration, the probabilistic forecast $\rho(\cdot|\underline{\mu})$ should be processed through Bayes theorem, and the optimal decision should minimize the expected loss under the posterior density of θ given the forecast $\rho(\cdot|\underline{\mu})$.

Parameterization of the forecast densities enables us to treat probabilistic forecasts in a manner similar to that for categorical forecasts, with $\underline{\mu}$ replacing t. Accordingly, from a joint historical record of the forecast densities $\rho(\cdot|\underline{\mu})$ and the actual values of θ, a likelihood function $f(\underline{\mu}|\cdot)$ of θ must be estimated and then processed, together with the prior density g of θ, through Bayes theorem to obtain the posterior density $\eta(\cdot|\underline{\mu})$ of θ.

The optimal solution is in the form of a decision function α^* specifying for every forecast $\rho(\cdot|\underline{\mu})$ an optimal decision $a^* = \alpha^*(\underline{\mu})$. The associated risk is

$$RP = \int [\min_{\alpha} \int \ell(\theta,\alpha(\underline{\mu}))\ \eta(\theta|\underline{\mu})\ d\theta]\ \xi(\underline{\mu})\ d\underline{\mu} \qquad (8)$$

2.6 Probabilistic Forecast - Nonoptimal Decision

If forecasts are not perfectly calibrated, but yet are perceived as such by the decision maker, then his criterion of optimality is

$$\min_{\alpha} \int \ell(\theta,\alpha(\underline{\mu}))\ \rho(\theta|\underline{\mu})\ d\theta \qquad (9)$$

and the minimizing decision function is $\hat{\alpha}$ ($\hat{\alpha} \neq \alpha^*$). The risk is specified by

$$\overline{RP} = \iint \ell(\theta,\hat{\alpha}(\underline{\mu}))\ \eta(\theta|\underline{\mu})\ \xi(\underline{\mu})\ d\theta\ d\underline{\mu} \qquad (10)$$

The above results are summarized in Table 1.

3. OPPORTUNITY LOSSES AND FORECAST VALUES

3.1 Ordinal Relations Between Risks

In the order

$$RF \leq RP \leq RC \leq RN \qquad (11)$$

the suborders $RF \leq RP \leq RN$ and $RF \leq RC \leq RN$ hold always; the order $RP \leq RC$ holds if the probabilistic forecast is sufficient for the categorical

Table 1. Risks of forecasting schemes and decision procedures.

FORECAST	DECISION	
	OPTIMAL	NONOPTIMAL
PERFECT	RF	
PROBABILISTIC	RP	$\overline{RP}$
CATEGORICAL	RC	$\overline{RC}$
NAIVE	RN	

forecast in the sense of Blackwell's (1953) definition of sufficient experiments. A special case of sufficiency arises when the categorical forecast t appears as an element of μ. For more elaborate discussions of the notion of sufficiency see DeGroot (1970, p.433), Hilton (1981), Alexandridis and Krzysztofowicz (1982). The orders

$$RP \leq \overline{RP} \quad , \quad RC \leq \overline{RC} \tag{12}$$

are obvious. No order can be asserted a priori between RP and $\overline{RC}$, between RC and $\overline{RP}$, between RN and $\overline{RC}$, and between RN and $\overline{RP}$.

3.2 Decision Expected Opportunity Losses

If the decision maker implements a nonoptimal decision, he incurs an opportunity loss. The decision expected opportunity losses, DOL, for probabilistic and categorical forecasts are defined, respectively, as

$$\begin{aligned} DOLP &= \overline{RP} - RP \\ DOLC &= \overline{RC} - RC \end{aligned} \tag{13}$$

Two situations can be distinguished to interpret DOL. The first situation is that of a decision maker who is unaware of the normative rules for optimal processing of information and decision making (Slovic, 1982). Accordingly, DOLC is the expected opportunity loss due to ignoring uncertainty of the categorical forecast; DOLP is the expected opportunity loss due to presuming that the probabilistic forecast is perfectly calibrated. In short, DOLC and DOLP represent the costs of not using the optimal decision procedures.

The second situation is that of a decision maker who, for whatever reason, does not have climatic data necessary for the estimation of the prior density and the likelihood functions. Now, DOLC and DOLP represent the values of prior climatic data to the decision maker. Naturally, prior climatic data are also used by the forecaster in preparation of a forecast. The value of prior data in this role is not separated herein, but it constitutes a part of the forecast value.

3.3 Forecast Values

The economic values, V, of perfect, probabilistic, categorical, and naive forecasts are defined, respectively, as

$$\begin{aligned} VF &= RN - RF \\ VP &= RN - RP \quad , \quad \overline{VP} = RN - \overline{RP} \\ VC &= RN - RC \quad , \quad \overline{VC} = RN - \overline{RC} \\ VN &= 0 \end{aligned} \tag{14}$$

The ordinal relations between the risks stated in Section 3.1 translate to the following conclusions concerning the values of forecasts. (1) When a forecast is used in an optimal decision procedure, then the value of the forecast is always nonnegative. Furthermore, (2) a perfect forecast is always at least as valuable as a probabilistic forecast; a probabilistic forecast, which is sufficient for a categorical forecast, is always at least as valuable as the categorical forecast; a categorical forecast is always at least as valuable as a naive forecast:

$$VF \geq VP \geq VC \geq VN \tag{15}$$

(3) When a forecast is used in a nonoptimal decision procedure, then the value of the forecast may be negative. Furthermore, (4) a probabilistic forecast may be less valuable than a categorical forecast, and a categorical forecast may be less valuable than a naive forecast.

4. ANALYSIS OF CATEGORICAL FORECASTS

This and the next sections present simple numerical examples. Their purpose is to illustrate the economic value of prior climatic data, values of both types of forecasts (categorical and probabilistic), gains from probabilistic forecasts, and expected opportunity losses to be incurred by suboptimal users who do not take full advantage of all climatological information. Relationships between these economic measures and statistical measures of the forecast skill are established. Implications of these relationships for both climate forecasters and climate information users are discussed.

4.1 Elements of the Model

The elements of the decision model are specified as follows. The loss function:

$$\ell(\theta,a) = (\theta - a)^2 \tag{16}$$

The prior density:

$$g(\theta) \sim N\ (M_\theta, S_\theta^2\ ) \tag{17}$$

The likelihood function:

$$f(t|\theta) \sim N\ (\bar{\varepsilon} + \theta, \sigma^2) \tag{18}$$

N denotes a normal probability law with mean and variance as indicated in parentheses. The model for the likelihood function arises from two assumptions: (1) the error of the forecast

$\varepsilon = t - \theta$ conforms to a normal density: $h(\varepsilon) \sim N(\bar{\varepsilon}, \sigma^2)$, and (2) ε is independent of θ so that $f(t|\theta) = h(t-\theta)$. The resulting expressions for the risks and the decision expected opportunity loss are as follows: $RN = S_\theta^2$, $RF = 0$, and

$$RC = \frac{\sigma^2 S_\theta^2}{\sigma^2 + S_\theta^2} \qquad (19)$$

$$DOLC = \frac{\sigma^4}{\sigma^2 + S_\theta^2} + \bar{\varepsilon}^2 \qquad (20)$$

The remaining elements follow immediately from the relations specified in Section 3. Details of the derivations can be found in Krzysztofowicz (1983).

4.2 Results and Implications

For numerical examples, we assumed that the predictand θ is the average seasonal temperature with parameters: $M_\theta = 50^oF$, $S_\theta^2 = 100^oF^2$, and $\bar{\varepsilon} = 3^oF$. The value of the risk $RN = S_\theta^2$ was chosen conveniently 100 so that other risks can be interpreted in terms of a percent of RN. (Note that the units of the risk can be changed arbitrarily, by multiplying the loss function by a positive constant, without affecting the optimal decision and the risk ratios.)

The most revealing result is the relationship between the economic value of the forecast and the variance of the forecast error σ^2. The variance σ^2 represents a statistical measure of the forecast skill. Figure 1 illustrates the fact that a decision maker who optimally accounts for uncertainty is always better off with a categorical forecast than without it (VC > VN), no matter the skill of the forecast. On the contrary, a decision maker who ignores forecast uncertainty may be worse off with the forecast than without it ($\overline{VC}$ < VN). Thus the implication of not using prior climatic data and a Bayesian decision procedure on the economic value of a categorical forecast is that the forecast value may be negative.

Of particular interest to users of long-range forecasts, which usually have relatively low skill, should be the robustness of the Bayesian decision procedure. For instance, even if the variance of the forecast error σ^2 were 200, doubled value of the climatological variance S_θ^2, such a forecast would still be worth 33% of its potential value to a user who employs an optimal decision procedure. At the same time, a user who employs a nonoptimal decision procedure would be losing an amount, $\overline{VC}$, equal to 109% of the potential forecast value.

The difference DOLC = VC - $\overline{VC}$ constitutes the expected opportunity loss incurred by the user whose decision procedure ignores forecast uncertainty. Alternatively, DOLC represents the economic value of the prior climatic data without which Bayesian decision procedure could not be applied. Figure 1 indicates that this value

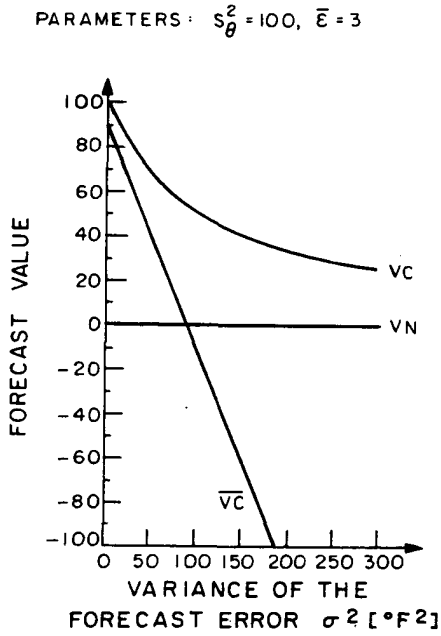


Fig. 1. Sensitivity of the value of a categorical forecast to the variance of the forecast error.

increases as the forecast skill decreases.

5. ANALYSIS OF PROBABILISTIC FORECASTS

5.1 Elements of the Model

The probabilistic forecast of θ is specified herein by $\underline{\mu} = (t, \lambda)$, where t is the median:

$$Pr(\theta \leq t) = Pr(\theta \geq t) = \frac{1}{2} \quad , \quad -\infty < t < \infty \qquad (21)$$

and λ is the width of a p % central credible interval:

$$Pr(t - \frac{\lambda}{2} \leq \theta \leq t + \frac{\lambda}{2}) = p \quad , \quad \lambda > 0 \qquad (22)$$

On each forecast occasion, the forecaster assesses, first, the median t and, next, the width λ of the p % central credible interval. Thus, (t, λ) defines three points on the forecast distribution P of θ (Figure 2).

The model for the likelihood function

$$f(\underline{\mu}|\theta) = f(t|\lambda,\theta)\ f(\lambda|\theta) \qquad (23)$$

is derived from the following assumptions. (1) The density h of the error $\varepsilon = t - \theta$ is conditioned by λ via the variance: $h(\varepsilon|\lambda) \sim N(\bar{\varepsilon}, b^2\lambda)$, where b^2 is a parameter. (2) ε is independent of θ. Thus

$$f(t|\lambda,\theta) = h(t - \theta|\lambda) \sim N(\bar{\varepsilon} + \theta, b^2\lambda) \qquad (24)$$

(3) The credible interval λ is independent of θ,

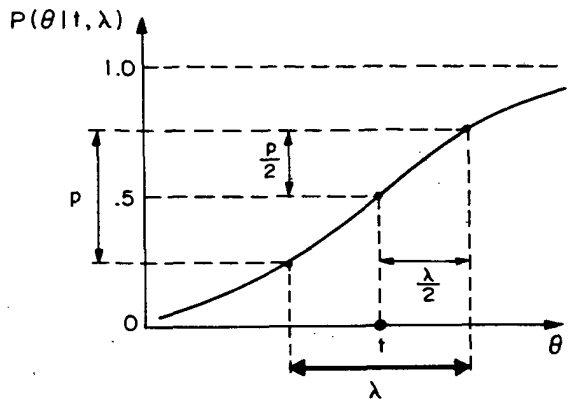


Fig. 2. Example of a variable-width, fixed-probability central credible interval forecast.

and the density ξ of $1/\lambda$ is gamma with parameters α and β:

$$f(\lambda|\theta) \sim \xi(\tfrac{1}{\lambda}) \sim G(\alpha,\beta) \qquad (25)$$

With the loss function and the prior density as in Section 4.1, the expressions for the risk and the decision expected opportunity loss take the form:

$$RP = \frac{\beta^{\alpha} b^2 S_\theta^2}{\Gamma(\alpha)} \int_0^{\infty} \frac{z^{\alpha-1}}{b^2 + S_\theta^2 z} e^{-\beta z}\, dz \qquad (26)$$

$$DOLP = \bar{\varepsilon}^2 + \frac{\beta^{\alpha} b^4}{\Gamma(\alpha)} \int_0^{\infty} \frac{z^{\alpha-2}}{b^2 + S_\theta^2 z} e^{-\beta z}\, dz \qquad (27)$$

For further analysis, it is convenient to express the parameters α and β in terms of the mean $E[1/\lambda] = \nu$ and the unconditional variance $Var(\varepsilon) = \sigma^2$. The relations are

$$\alpha = \frac{\nu\sigma^2}{\nu\sigma^2 - b^2}, \qquad \beta = \frac{\sigma^2}{\nu\sigma^2 - b^2} \qquad (28)$$

under the condition $\nu\sigma^2 - b^2 > 0$. For the derivations see Krzysztofowicz (1983).

5.2 Results and Implications

The values of the parameters were: $M_\theta = 50^oF$, $S_\theta^2 = 100^oF^2$, $\bar{\varepsilon} = 3^oF$, $\sigma^2 = 100^oF^2$, $b^2 = 10^oF^2/^oF$, $1/\nu = 1.82^oF$. Sensitivity analyses were performed by varying the values of the unconditional variance of the error of the median, σ^2, and the "average" width of the credible interval, $1/\nu = 1/E[1/\lambda]$. Figures 3 and 4 relate these statistical measures of the forecast skill to an economic measure of the forecast value. As expected, the less skillful the median t (i.e., the larger σ^2), or the less certain the forecaster (i.e., the larger $1/\nu$), the less valuable the forecast; this is true for both the optimal (VP) and nonoptimal ($\overline{VP}$) decision procedures.

For a fixed $1/\nu$, Figure 3 reveals a striking insensitivity of the forecast value VP (and $\overline{VP}$ to a lesser degree) to σ^2. That is, if the forecaster's degree of confidence is justifiably high, and he is able to consistently specify relatively narrow credible intervals λ within which he expects the actual value of θ to fall with probability p, then the forecasts have a high economic value VP even though the medians t

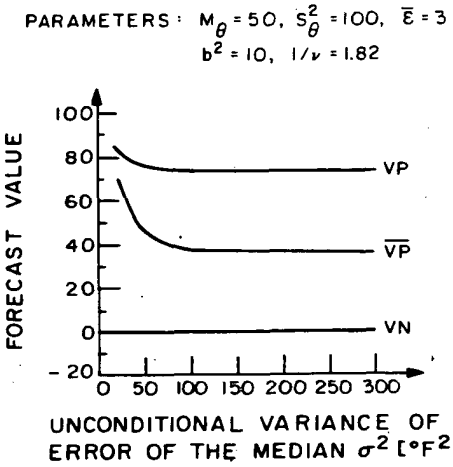


Fig. 3. Sensitivity of the value of a probabilistic forecast to the unconditional variance of the error of the median.

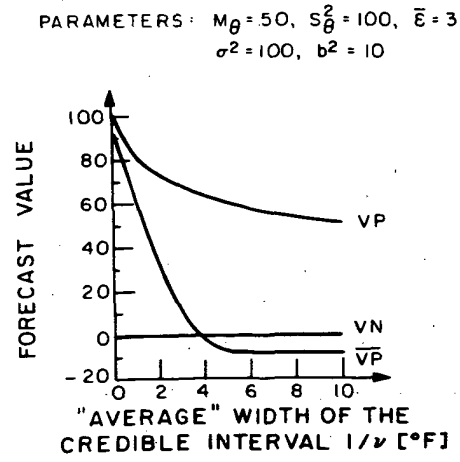


Fig. 4. Sensitivity of the value of a probabilistic forecast to the "average" width of the credible interval.

themselves may be poor categorical (point) predictors of θ. On the other hand, Figure 4 indicates that for a fixed σ^2, the forecast value VP is sensitive (and $\overline{VP}$ is very sensitive) to the "average" width of the credible interval $1/\nu$.

In summary, the economic value of a probabilistic climate forecast is much more sensitive to a statistic which reflects the skill in assessing the confidence ($1/\nu$) than to a statistic which reflects the skill in assessing a point predictor (σ^2). Inasmuch as forecasts having a small "average" width $1/\nu$ and somewhat large variance σ^2 could be associated with a skillful but unlucky forecaster, the insensitivity of VP to σ^2 seems plausible from a normative standpoint.

These results provide an economically-based argument concerning a desirable direction of research in climate forecasting. This research should be aimed primarily at narrowing the confidence bounds of forecasts and at quantitative assessment of these bounds within a probability framework rather than on precise estimation of the "best" point predictors such as medians, means, etc.

The expected opportunity loss DOLP = VP - $\overline{VP}$ incurred by a user who employs the nonBayesian decision procedure may be substantial: up to 70% of the potential value of the forecast (Figure 4) Moreover, for large $1/\nu$, the value of the forecast to such a user, $\overline{VP}$, is negative. On the contrary, the value of the same forecast to a user employing Bayesian decision procedure, VP, remains greater than 50% of the potential value. According to an alternative interpretation, DOLP represents the economic value of the prior climatic data. Figures 3 and 4 suggest that this value generally increases as the forecast skill decreases. (An exception: for $1/\nu > 5$, DOLP is convex and slowly decreasing.)

5.3 Economic Gains From Probabilistic Forecasts

What is to be gained from quantification of the forecast uncertainty? To answer this question, let us compare Figures 1 and 3 under the assumption that the categorical forecast coincides with the median of the probabilistic forecast. One can clearly see that for both decision procedures (optimal and nonoptimal), the probabilistic forecast is always at least as valuable as the categorical forecast. The value difference VP - VC and $\overline{VP}$ - $\overline{VC}$, plotted in Figure 5, measures the economic gain from the probabilistic forecast. This gain is to be attributed specifically to the forecaster's ability to quantify uncertainty in terms of the credible interval. The gain increases with the variance of the error σ^2 and is larger for the nonoptimal decision procedure than for the optimal procedure.

Two conclusions can be drawn from Figure 5. First, recall from Section 3.2 that an alternative explanation for the use of the nonoptimal decision procedure is the lack of prior climatic data. The relation $\overline{VP}$ - $\overline{VC}$ > VP - VC implies that the quantification of uncertainty in climate forecasts is particularly important when the prior data record is nonexistent (case $\overline{VP}$ - $\overline{VC}$) or relatively short. (If we assume that the

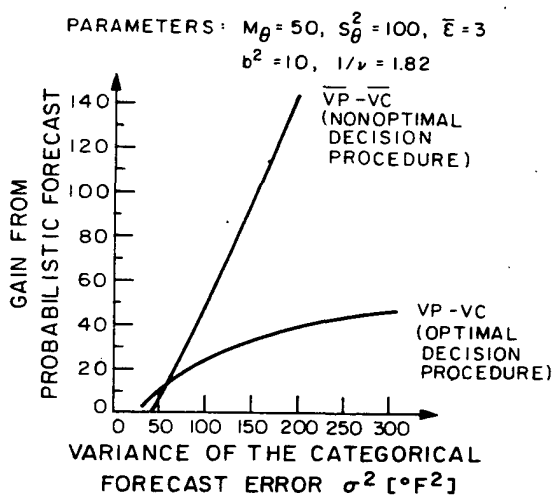


Fig. 5. Gains from a probabilistic forecast (relative to gains from a categorical forecast) for optimal and nonoptimal decision procedures.

case VP - VC corresponds to an "infinite" prior data record, then cases with shorter and less informative data records would be plotted above the curve VP - VC but below the curve $\overline{VP}$ - $\overline{VC}$.) Second, inasmuch as a large variance of the categorical forecast error σ^2 can be associated with low skill of categorical forecasts, the following conclusion emerges: the lower the skill of a categorical forecast, the higher the economic gain from the explicit quantification of the forecast uncertainty.

In summary, two arguments can be provided in support of probabilistic forecasting of climate. (1) An economic argument: information about uncertainty conveyed by probabilistic forecasts, even in such a simple form as a central credible interval, can be transformed into large tangible gains by economically motivated users even if the decision procedures they employ are suboptimal. (2) A behavioral argument: in view of the mounting evidence compiled by behavioral scientists that people tend to ignore uncertainties, particularly those not directly apparent or pointed out explicitly, probabilistic forecasts should be considered superior to categorical forecasts because they do apprise users that forecasts are inherently uncertain. This, in turn, may impel the users to refine their decision procedures so as to take the forecast uncertainty into account.

6. SUMMARY AND CONCLUSIONS

1. Probabilistic forecasts, which explicitly quantify the forecaster's uncertainty about a future value of a climatic variable, are more valuable than categorical (point) forecasts when compared via the same (optimal or nonoptimal) decision procedure.

2. An opportunity loss is always incurred by a user whose decision procedure does not account for errors of categorical forecasts or for imperfect calibration of probabilistic forecasts. The expected opportunity loss increases as the

forecast skill decreases. Consequently, the economic value of a forecast may be negative.

3. In order to apply an optimal (Bayesian) decision procedure, the user must have access not only to climate forecasts but also to prior (historical) climatic data or to probabilistic models of such data. The economic value of prior climatic data generally increases as the skill of a forecast (categorical or probabilistic) decreases.

4. The economic gain from using probabilistic climate forecasts instead of categorical forecasts increases as (a) the skill of categorical forecasts decreases, or (b) the informativeness of the prior climatic data diminishes.

5. The economic value of a probabilistic climate forecast is much more sensitive to the forecaster's skill in assessing the confidence (i.e., quantifying uncertainty) than to the skill in assessing a "best" point predictor (e.g., median, mean). From the viewpoint of economically motivated users of climate forecasts, this finding provides an argument for a desirable direction of research in climate forecasting: quantification of uncertainty.

The above conclusions are predicated on the structure of the exemplary decision models. They are, therefore, general to a degree to which the decision models are representative of reality. At the very least, the above conclusions suggest the kinds of analyses that are likely to produce significant insight into the behavior and economics of climatological information-decision systems.

7. REFERENCES

Alexandridis, M.G. and Krzysztofowicz, R., May 1982: Value of Categorical and Probabilistic Temperature Forecasts for Scheduling of Power Generation, Report No. 282, Ralph M. Parsons Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Alpert, M. and Raiffa, H., 1982: A Progress Report on the Training of Probability Assessors. In Kahneman, D., Slovic, P. and Tversky, A. (Eds.), Judgment under Uncertainty: Heuristics and Biases, Cambridge University Press, New York, 294-305.

Blackwell, D., 1953: Equivalent Comparison of Experiments, Annals of Mathematical Statistics, 24, 265-272.

DeGroot, M.H., 1970: Optimal Statistical Decisions, McGraw-Hill, New York.

Hilton, R.W., 1981: The Determinants of Information Value: Synthesizing Some General Results, Management Science, 27(1), 57-64.

Krzysztofowicz, R., 1983: Why Should a Forecaster and a Decision Maker Use Bayes Theorem, Water Resources Research, 19(2), 327-336.

Morris, P.A., 1977: Combining Expert Judgments: A Bayesian Approach, Management Science, 23(7), 679-693.

Slovic, P., 1982: Toward Understanding and Improving Decisions. In Howell, W.C. and Fleishman, E.A. (Eds.), Human Performance and Productivity: Vol. 2, Information Processing and Decision Making, Erlbaum, Hillsdale, NJ.

Winkler, R.L. and Murphy, A.H., 1979: The Use of Probabilities in Forecasts of Maximum and Minimum Temperatures, Meteorological Magazine, 108, 317-329.

STATISTICAL MODELS FOR ECONOMICALLY OPTIMUM APPLICATION OF CLIMATOLOGICAL DATA

E.E.Zhukovsky

Agrophysical Research Institute
Leningrad USSR

1. INTRODUCTION

For the past ten years there was carried out a great complex of investigations at the Agrophysical Research Institute. The studies concern elaboration of methods for optimal use and evaluation of potential economic efficiency of various kinds of meteorologic information (Zhukovsky, Chudnovsky, 1978; Zhukovsky, 1981). Probability approaches developed and widely used in the theory of statistic decisions (Ventsel, 1972; Chernoff and Moses, 1959; a.o) form the basis of the present work. This paper describes some of the results obtained in these investigations dealing with optimal use of climatologic information. It shows economic utility of the climatologic probability information whose role and significance, unfortunately, is often underestimated by many users up to now.

2. BASIC PRINCIPLES

The analysis is based on the following meteorologic-economic model. It is assumed that a weather-sensitive object or some technologic process is affected by a continuous (as to its possible values) meteorologic element X, the climatic variability of which characterizing variations of weather conditions is described by a certain distribution law g(x). Here and further actual values of x are denoted by X.

The essence of the decision-making process is to plan economic measures corresponding to some definite X=a value which is preassigned and is not further changed from one situation to another. Decisions connected with long-term national-economic planning, choice of project parameters for various constructions, district division of crops, etc., are of such a character.

It is considered that due to departure of actual X=x values from the chosen one X=a there arise certain losses characterized by r(x,a) function further called "the function of meteorologic losses". The term "meteorologic" is introduced because of the fact that the concerned function represents losses caused by uncontrolled variations of the affecting meteorologic conditions. In the general case r value depends on each of x and a variables taken separately. At the same time there exists a wide class of problems where meteorological losses r depend only on the difference h = x-a, so that r(x,a) = r(x-a). Some examples of such functions are given in Fig. 1. The main relations will be deduced for the difference linear function r(x-a) shown graphically in Fig. 1a. Analytic expression of this function is

$$r(x,a) = \begin{cases} A_1 (a-x) & \text{at } x \le a \\ A_2 (x-a) & \text{at } x > a \end{cases} \qquad (1)$$

where A_1 and A_2 are constants characterizing economic significance of negative or positive departure h, respectively.

Adopting Bayesian approach followed by most authors dealing with the problems of optimal use of meteorologic information (Bagrov, 1966; Monin, 1962; Gruza, 1967; McQuigg, 1966; Murphy and Thompson, 1977; Murphy, 1977; a.o.), let us call planning of economic measures corresponding to such X = a∈I as climatologically optimum strategy. In this case statistically averaged losses

$$R_a = \int_{(x)} r(x,a)\, g(x)\,dx \qquad (2)$$

are minimal. Substituting expression(1)

into (2) and differentiating with respect to a, it is easy to show that a_{cl}, the value of interest to us, is the root of equation

$$\int_{x \le a} g(x)dx = \frac{1}{1+k} \qquad (3)$$

where k denotes A_1/A_2 ratio.

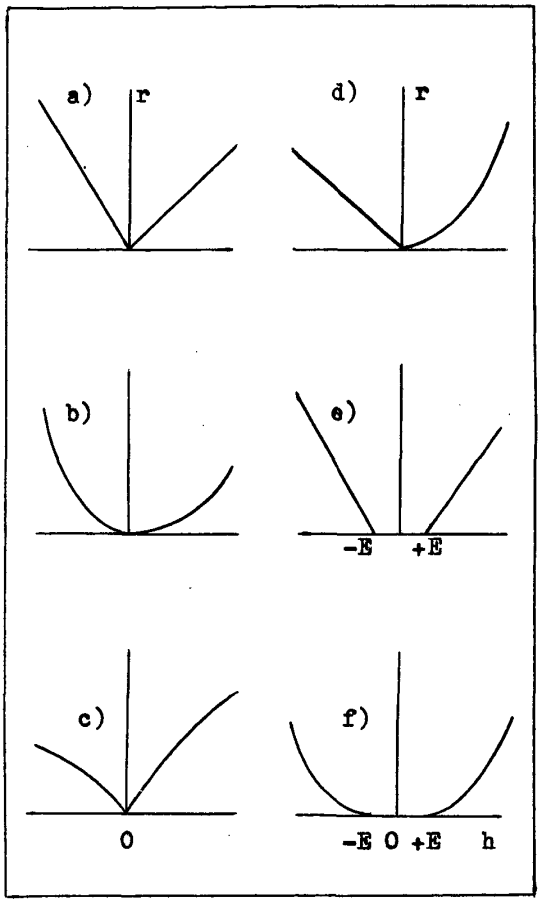


Fig. 1. Some kinds of difference functions of meteorologic losses: a - linear; b - quadratic; c - exponential; d - linear-quadratic; e - linear with the insensitive zone; f - quadratic with the insensitive zone.

Equation (3), valid for any distribution law X, shows that if negative and positive departure of actual x from a have the same economic 'weight' (k=1), then climatologically optimum strategy is orientation on the median of distribution.

Assuming that in the absence of probability climatologic characteristics of X or due to inability of using them economic decisions are made corresponding to the simplest climatologic characteristic for which $X = \bar{x}$ is the norm. The obtained gain owing to climatologic optimization of decisions may be estimated by a relative index

$$y_{cl} = \frac{R_x - R_{a_{cl}}}{R_x} \times 100\% \qquad (4)$$

It is obvious that y_{cl} value may simultaneously be regarded as an index of potential efficiency of probability climatologic information being estimated with relation to a particular user.

3. DERIVING CLIMATOLOGICALLY OPTIMAL ECONOMIC STRATEGIES AND THEIR EFFICIENCY

3.1 The Case of Gaussian Distribution

Gauss law is widely used for approximation of distribution of various meteorological elements such as numerous temperature characteristics, precipitation sums, etc. (Kobysheva, 1971; Zhukovsky et al., 1976).

In case of Gaussian distribution law defined, as it is well-known, merely by two parameters, namely, by the norm $\bar{x}$ and RMS departure s, equation (3) takes the form of

$$F(t) = \frac{1}{2}\,\frac{1-k}{1+k} \qquad (5)$$

where

$$t = (a - \bar{x})/s \qquad (6)$$

Symbol F denotes probability integral as follows

$$F(t) = \frac{1}{\sqrt{2\pi}} \int_0^t \exp(-z^2/2)dz \qquad (7)$$

For any k one can find $t=t_o$ value satisfying (5), and then use equation (6) to calculate the required a_{cl} values which equals

$$a_{cl} = \bar{x} + t_o s \qquad (8)$$

Table 1 shows particular values of t_o for different $k > 1$. For $k < 1$ (1/2, 1/3, etc.) t_o quantities are the same in their absolute value, but positive.

Table 1. Optimal parameter t_o as a function of k for users with a difference linear function of losses (1)

k ...	1	2	3	5	10	20
t_o...	0	-0.43	-0.67	-0.97	-1.34	-1.67

One may easily see that at $k<1$ economically optimal strategy comes to planning economic measures calculating on some X values above the norm, and et $k>1$ vice versa, i.e. on some X values lower than the norm. For $k = 1$ from (5) it follows that $t_o = 0$ and $a_{cl} = \bar{x}$. It is connected with the fact that for Gaussian law, as for any symmetrical distribution, the median and the norm coincide.

As it is seen from a more comprehensive analysis, at an arbitrary a there exists the following equation

$$R_a = (A_1+A_2)s\left[tF(t)+f(t) - \frac{1}{2}\frac{1-k}{1+k}t\right] \quad (9)$$

where f(t) is Gaussian probability density defined as

$$f(t) = \frac{1}{\sqrt{2\pi}} \exp(-t^2/2) \quad (10)$$

Substitution of $t_o = (a_{cl} - \bar{x})/s$ taking into account equation (5) gives

$$R_{a_{cl}} = (A_1 + A_2)\, sf(t_o) \quad (11)$$

Similarly it is easy to show that at $a = \bar{x}$

$$R_{\bar{x}} = (A_1 + A_2)\, s/\sqrt{2\pi} \quad (12)$$

Comparison of (11) and (12) allows to conclude that in case of Gaussian distribution law of X transition from orientation on the norm to climatologically optimal strategy lowers average losses of the users $\exp(t_o^2/2)$ times, and index of potential efficiency y_{cl} may be calculated from

$$y_{cl} = \left[1 - \exp(-t_o^2/2)\right] \times 100\% \quad (13)$$

Fig. 2 demonstrates this value as a function of k, and it should be borne in mind that $y_{cl}(k) = y_{cl}(1/k)$. The resulting gain due to climatologic optimization or, which is the same, potential economic utility of probability climatologic information appears to be higher, the more differ the weighting factors A_1 and A_2. For example, at $k = 2$ (or 1/2), transition from planing corresponding to the norm to climatologically optimal strategy makes it possible to lower average losses caused by weather changes by about 9% and at $k = 5$ (or 1/5) - by 37%.

3.2 The Case of Log-Normal Distribution

In this case the climatologic variability of X is described by the following distribution function:

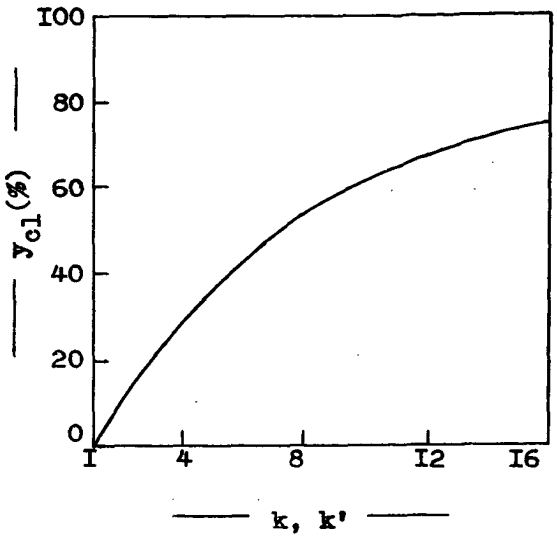


Fig. 2. Relative decrease of average losses by transition from orientation on the norm to climatologically optimal strategy for Gaussian distribution of the influencing meteorologic element.

$$g(x) = \begin{cases} 0 & \text{at } x \le 0 \\ \dfrac{1}{x\,c\sqrt{2\pi}} \exp\left[-\dfrac{(\ln x - b)^2}{2c^2}\right] & \text{at } x > 0 \end{cases} \quad (14)$$

where b and c are parameters in terms of which the norm and variance X are expressed by:

$$\left.\begin{aligned} \bar{x} &= \exp(b + c^2/2) \\ s^2 &= \bar{x}^2(\exp c^2 - 1) \end{aligned}\right\} \quad (15)$$

It is to be noted that for meteorologic elements characterized by log-normal distributionthe norm and the median, the latter being denoted by m, do not coincide. The value of m is determined according to equation $m = \exp b$ and it always appears to be less than $\bar{x}$.

Substituting equation (14) into initial equation (3) the following relations may be obtained for calculating a_{cl}:

$$a_{cl} = \exp(b + ct_o) \quad (16)$$

Here, as before, t_o is the root of equation (5). In full agreement with general equation (3) valid for any distribution law it follows from (16) that at $A_1 = A_2$

$$a_{cl} = \exp b = m \quad (17)$$

The relation between a_{cl} and m values is given in Fig. 3, where family of curves $w = a_{cl}/m$ for various c is shown. Unlike the Gaussian law, the

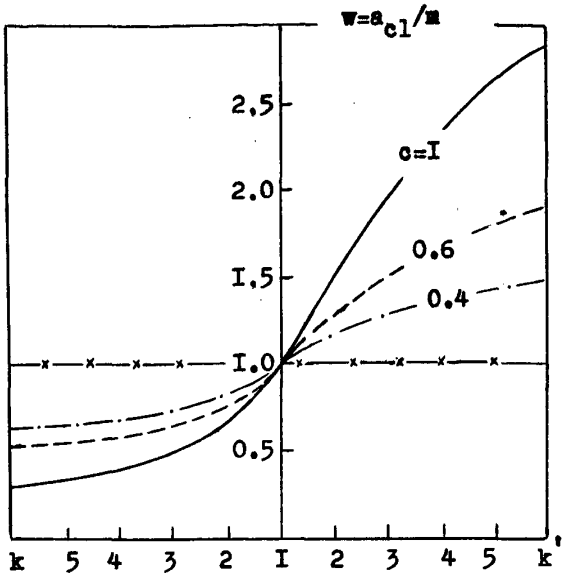


Fig. 3. Index w as a function of relation of weights of positive and negative departures for log-normal distribution of the influencing meteorologic element.

shifts of a_{cl} with respect to the point m at $k > 1$ and at the same $k' = 1/k > 1$ are here different. The difference between these shifts is the stronger, the higher is c value, i.e. the more is the distribution asymmetry X. In the discussed case average losses, when planning economic measures corresponding to a_{cl}, obtained from equation (16), are as follows:

$$R_{a_{cl}} = (A_1+A_2)\bar{x}\left[F(t_o)-F(t_o-c)\right] \qquad (18)$$

It represents the theoretic minimum to which average losses of the user may be brought up when information on log-normal X distribution is optimally used.

Supposing, as previously, that in the absence of proper information decisions are made corresponding to the norm of X, it is possible to find that using the simplest strategy mentioned above, average losses of the same user will be equal to

$$R_{\bar{x}} = 2(A_1 + A_2)\bar{x}\,F(c/2) \qquad (19)$$

In the case of orientation on the median average losses may be determined from the following equation:

$$R_m = (A_1+A_2)\bar{x}\left[F(c)+(1-m/\bar{x})F(t_o)\right] \qquad (20)$$

whence, in particular, for $k = k' = 1$ ($A_1 = A_2 = A$) one obtains

$$R_m = 2A\bar{x}\,F(c) \qquad (21)$$

The knowledge of $R_{a_{cl}}$ and $R_{\bar{x}}$ values allows to calculate y_{cl} index. Appropriate calculations have been made for different c. Fig. 4 shows the relation of $y_{cl} = y_{cl}(A_1, A_2)$ for the case of $c = 1$.

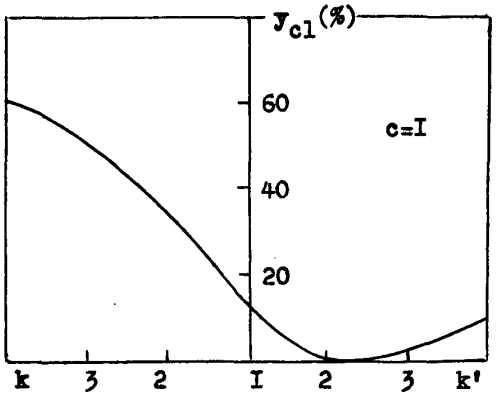


Fig. 4. Relative decrease of average losses by transition from orientation on the norm to climatologically optimal strategy for log-normal distribution of the influencing meteorologic element.

Here two things are of interest. First, unlike Gaussian distribution, transition from orientation on the norm to climatologically optimum strategy gives a certain gain (in the above example it makes about 10%) even at $k = k' = 1$. It results directly from the fact that according to (3) optimal strategy at $k = k' = 1$ is, in the general case, orientation not on the norm but on the median. Since $m < \bar{x}$ orientation on the norm appears to be optimal strategy for users having ratio $k' = A_2/A_1$ somewhat above one. This k' may be calculated from equation

$$k' = k'_o = \frac{1 + 2\,F(c/2)}{1 - 2\,F(c/2)} \qquad (22)$$

and for the example shown in Fig. 4 $k' = 2.3$.

Second, a marked asymmetry of the relation $y_{cl} = y_{cl}(A_1, A_2)$ with respect to point $k' = k'_o$. It shows that transition to optimal strategy will be essentially different for users with $k' < k'_o$ and $k' > k'_o$, namely, the effect is more for the first ones.

3.3 Regularities for Other Distribution Laws

Some equations for obtaining optimal strategies and evaluation of potential efficiency of climatologic infor-

mation at uniform and Rayleigh distributions of X will be given without any additional explanations.

For the uniform law determined by the relation

$$g(x) = \begin{cases} 0 & \text{at } x \le x_1,\ x \ge x_2 \\ \dfrac{1}{x_2 - x_1} & \text{at } x_1 < x < x_2 \end{cases} \qquad (23)$$

the following equations

$$a_{cl} = \bar{x} + \frac{1-k}{1+k} l \qquad (24)$$

$$R_{a_{cl}} = (A_1 + A_2) l \frac{k}{(1+k)^2} \qquad (25)$$

$$R_{\bar{x}} = (A_1 + A_2) l/4 \qquad (26)$$

are valid. Here x_1 and x_2 are minimal and maximal values of X, respectively; $l = (x_2 - x_1)/2$ is half of X variation range; $\bar{x}$ is the norm of X equal to $(x_1 + x_2)/2$.

The Rayleigh law is characterized by probability density

$$g(x) = \begin{cases} 0 & \text{at } x \le 0 \\ \dfrac{x}{u^2} \exp(-x^2/2u^2) & \text{at } x > 0 \end{cases} \qquad (27)$$

where u is a distribution parameter. In this particular case a_{cl}, $R_{a_{cl}}$ and $R_{\bar{x}}$ may be obtained from the following equations:

$$a_{cl} = u\sqrt{-2 \ln \frac{k}{1+k}} \qquad (28)$$

$$R_{a_{cl}} = (A_1 + A_2)\bar{x}\left[\frac{1+k\ a_{cl}/\bar{x}}{1+k} - 2F\left(\frac{a_{cl}}{u}\right)\right] \qquad (29)$$

$$R_{\bar{x}} = 0.21(A_1 + A_2)\bar{x} \qquad (30)$$

where $\bar{x} = u\sqrt{\pi/2}$

3.4 Some Additional Results

An analysis similar to the performed one has been carried out with respect to the users characterized by the functions of meteorologic losses which differ from difference linear function (1). If losses due to the departure of actual values from x used as a basis for the accepted economic plan are expressed by quadratic function

$$r(x,a) = \begin{cases} B_1(x-a)^2 & \text{at } x \le a \\ B_2(x-a)^2 & \text{at } x > a \end{cases} \qquad (31)$$

then climatologically optimal value a_{cl} is obtained from the following equation:

$$(1-k)\int_{x \le a} (a-x)g(x)dx = a - \bar{x} \qquad (32)$$

where $k = B_1/B_2$.

In case of Gaussian distribution X this equation may be presented as follows

$$\frac{t}{tF(t)+f(t)} = 2\frac{t-k}{1+k} \qquad (33)$$

(Zhukovsky, 1974; Belenky, Zhukovsky, 1974; Belenky, Brunova, Zhukovsky, 1976) and, as it may be shown, potential economic gain provided by the transition from orientation on the norm to climatologically optimal strategy ($a = a_{cl} = \bar{x} + t_o s$) is determined by

$$y_{cl} = 2\frac{1-k}{1+k} F(t_o) \cdot 100\% \qquad (34)$$

Particular values of the above parameter t_o which correspond to different $k > 1$ are presented in Table 2. For the inverse k (1/2, 1/3, etc.) the values of t_o are the same but positive.

Table 2. Optimal parameter t_o as a function of k for users with the difference quadratic function of losses (31)

k ...	1	2	3	5	10	20
t_o ...	0	-0.28	-0.44	-0.64	-0.90	-1.16

For many users small departures of x from a are of no economic importance, and unfavourable consequences of climatic variations of X begin to affect them only when $h = x-a$ surpasses a certain threshhold $E > 0$ by absolute value. In other words, there exists a definite insensitive zone. The loss functions of such a type have been shown in Fig.1e,f.

In Fig. 5 a family of curves $t_o(k)$ is shown to illustrate effects caused by the existence of the insensitive zone. The curves correspond to the case of Gaussian distribution law and of the difference linear loss functions

$$r(x,a) = \begin{cases} A_1(a-x-E) & \text{at } a-x > E \\ 0 & \text{at } |x-a| \le E \\ A_2(x-a-E) & \text{at } x-a > E \end{cases} \qquad (35)$$

with various width of the insensitive zone E (Belenky, Brunova, Zhukovsky, 1977).

As it follows from Fig. 5, the general character of climatologic optimal strategy in the presence of the insensitive zone remains the same as in the absence of the latter, namely, at $k > 1$ the user should always orientate on a_{cl} values which are lower than the norm, and

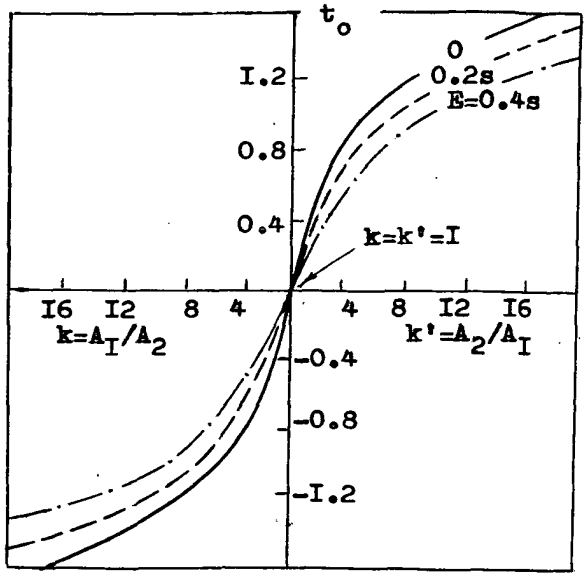


Fig. 5. Optimal parameter of t_o shift as a function of the relation between the weight factors A_1 and A_2 at different width of the insensitive zone.

at $k < 1$ vice versa, i.e. on a_{cl} which are higher than the norm. As to the absolute values of optimal a_{cl} shifts as related to $\bar{x}$, with the expansion of the insensitive zone they get somewhat smaller. A similar regularity takes place for the values of y_{cl}: the presence of the insensitive zone results in a certain decrease of the potential effect that can be obtained when using climatologic information. At the same time, as it follows from computations, the stated decrease of y_{cl} is small and the transition from orientation on the norm to climatologically optimal strategy should be considered as advantageous up to very high E.

3.5 Practical Applications

In the works by the author and his colleagues there were shown some possibilities of using the developed theory for solving several problems of agricultural meteorology. These problems concerned such points as fixing optimal irrigation rates, reasoning insurance grain stocks designed for reseeding winter crops, planning yielding abilities, etc. (Zhukovsky, Chudnovsky, 1973; Zhukovsky, 1974; Zhukovsky, Braginskaya, 1981, a.o.). The suggested decisions always took into account climatologic variability of the affecting agrometeorologic conditions, the cost of the obtained agricultural production and expenditure for carrying out appropriate agrotechnical measures.

4. DECISION MAKING IN THE PRESENCE OF AREA VARIABILITY OF CLIMATOLOGIC CHARACTERISTICS

In all meteorologic-economic models considered so far random variations of the affecting meteorologic factors were taken into account only in time and not in space. However, the above factors possess similar and sometimes even greater variability in space.

Let us assume that in each point of the territory being of interest to us the meteorologic element X affecting economic activity changes in climatic regime according to Gaussian law with the norm $\bar{x}$ and RMS departure s. From one point to another s remains constant but $\bar{x}$ varies following a certain distribution law $q(\bar{x})$. In accordance with that let us introduce area average norm which can be determined as

$$\bar{\bar{x}} = \int_{(\bar{x})} q(\bar{x})\, d\bar{x} \qquad (36)$$

It is considered that economic planning should be carried out in each point of the territory, the losses due to the departure of actual values of x from the values of a forming the basis of the accepted plan, being expressed by a difference linear function of losses of the form (1) which is the same for all the points.

Let $\bar{R}$ be average losses per single user of the territory in question. It is obvious that if in each point climatologically optimum decision is made $\bar{R}$ reaches its minimal level and coincides with $R_{a_{cl}}$ calculated from (11). Thus

$$\bar{R} = R_{a_{cl}} \qquad (37)$$

Suppose there is no information on the norm values in single points of the region and decision is made for the whole territory. The simplest situation is when economic activity of all users (in all the points of the territory in question) is orientated on space average norm $\bar{\bar{x}}$. According to equation (9) average losses of the user located in the point with a particular value of the norm $\bar{x}$ will be

$$R_{\bar{x}} = (A_1+A_2)\, s \left[t_{\bar{x}} F(t_{\bar{x}}) + f(t_{\bar{x}}) - \frac{1}{2}\frac{1-k}{1+k} t_{\bar{x}} \right] \qquad (38)$$

where $t_{\bar{x}} = (\bar{\bar{x}} - \bar{x})/s$, while mean losses averaged over territory are derived from

$$\bar{R} = \bar{R}_{\bar{x}} = \int_{(\bar{x})} R_{\bar{x}}\, q(\bar{x}) d\bar{x} \qquad (39)$$

Let us assume that the space distri-

bution of the norm $q(\bar{x})$ as well as time distribution of X in the point is Gaussian, and it is characterized by mean $\bar{X}$ known to us, and by some RMS departure $s_{\bar{X}}$. In this case equation (39) acquires the form of

$$\bar{R}_{\bar{X}} = R_{\bar{X}} \sqrt{1 + v^2} \qquad (40)$$

where $v = s_{\bar{X}}/s$. Thus, with orientation on the area average norm average losses increase $\sqrt{1 + v^2}$ times if compared to the case when decision is made corresponding to the point value of the norm.

It is clear, however, that orientation on $\bar{X}$ value is not the best economic strategy. Even in the absence of information on particular norm values at individual points one may suggest a more rational variant for economic planning. Namely, similar to the way it was done for single users, it is possible to find such a value of $\bar{X}$ which, generally speaking, differs from $X = a'_{cl}$, which would provide minimization of area-averaged mean losses. In the general case the value a'_{cl} will be the root of equation

$$\frac{\partial \bar{R}}{\partial a} = \int_{(\bar{x})} \frac{\partial R_a}{\partial a} q(\bar{x}) d\bar{x} = 0 \qquad (41)$$

which in the case of Gaussian distribution $q(\bar{x})$ may be written

$$\int_{-\infty}^{\infty} F(T - vz) f(z) dz = \frac{1}{2} \frac{1 - k}{1 + k} \qquad (42)$$

This is the same as

$$F\left(\frac{T}{\sqrt{1+v^2}}\right) = \frac{1}{2} \frac{1 - k}{1 + k} \qquad (42a)$$

In those equations $T = (a - \bar{X})/s$.

It follows that optimal T, denoted as T_o, equals

$$T_o = t_o \sqrt{1 + v^2} \qquad (43)$$

and consequently

$$a'_{cl} = \bar{X} + t_o s \sqrt{1 + v^2} \qquad (44)$$

Parameter t_o in (43) and (44) is the root of equation (5) given in Table 1.

The obtained relations show that a common for the whole territory climatologically optimum decision taking into consideration space norm variability of the affecting meteorologic element bears in general the same characteras climatologically optimal decisions in individual points, i.e. at $k > 1$ economic decisions should be planned corresponding to the values of a'_{cl} lower than $\bar{X}$, and at $k < 1$ - on the values of a'_{cl} higher than $\bar{X}$, the absolute value of the corresponding shift being $\sqrt{1+v^2}$ times higher than that at optimization of decisions in individual points.

To evaluate efficiency of the discussed optimal strategy it is enough to introduce the expression obtained for a'_{cl} into general equation

$$\bar{R} = \int_{(\bar{x})} R_a \, q(\bar{x}) d\bar{x} \qquad (45)$$

Particularly for Gaussian distribution $q(\bar{x})$

$$\bar{R} = \bar{R}_{a'_{cl}} = \bar{R}_{\bar{X}} \exp(-t_o^2/2) \qquad (46)$$

This equation shows that similar to the case of optimizing decisions in individual points, transition from orientation on area averaged norm $\bar{X}$ to the decision optimal for the whole territory $a'_{cl} = \bar{X} + t_o s \sqrt{1 + v^2}$ provides an $\exp(t_o^2/2)$-times decrease in average losses.

5. CONCLUSION

The developed theory and numerical evaluations carried out on its basis show that rational use of climatological probability information consisting in the transition to climatologically optimal economic strategies makes it possible to increase significantly the efficiency of decision making as compared to planning economic measures corresponding to mean conditions. An additional effect may be obtained by optimal taking into account information on space variability of climatologic characteristics.

Only main methodologic principles of the statistical approach being developed have been discussed in this paper. Realization of the obtained results requires a detailed analysis of each particular economic situation and a subsequent construction of proper meteorologic-economic models based on the above analysis. At present attention should be attracted to construction of such models providing participation of specialists working in different branches of applied climatology.

6. REFERENCES

Bagrov, N.A., 1966: On economic usefulness of forecasts. Meteorologia i gidrologia, 2, 3-12.

Belenky, D.H., Zhukovsky, E.E., 1974: On some approaches to the problems of optimal use of the hydrometeorologic information. - In: Hydrometeorological provision of national economy, Moscow, Gidrometeoizdat, 125-131.

Belenky, D.Kh., Brunova, T.M., Zhukovsky, E.E., 1976: On criteria of optimal use of meteorological data. Gidrometeorologia i narodnoe khoziaystvo. Moscow, Gidrometeoizdat, 41-47.

Belenky, D.Kh., Brunova, T.M., Zhukovsky E.E., 1977: On statistical minimization of losses for one class of users of meteorological information. Trudy Kazakhskogo NIGMI, 62,31-39.

Chernoff, G., Moses L., 1959: Elementary decision theory. New York.

Gruza, G.V., 1967: Some general problems of theory of forecast based on statistical data. Trudy SANIGMI, 29 (44), 3-41.

Kobysheva, N.V., 1971: Indirect methods for calculation of climatic characteristics. Leningrad, Gidrometeoizdat,

McQuigg, J.D., Thompson, R.G., 1966: Economic value of improved methods of translating weather information into operation terms. - Mon.Wea. Rev.,94,2, 83-87.

Monin, A.S., 1962: On application of unreliable forecasts. Izvestia Akademii nauk SSSR, seria geofisicheskaya, 2, 218-228.

Murphy, A.H., 1977: The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation. Mon. Wea. Rev., 105,7, 803-816.

Murphy, A.H., Thompson, J.C., 1977: On the nature of the nonexistence of ordinal relationships between measures of the accuracy and value of probability forecasts: an example. J. Appl. Met., 16, 10, 1015-1021.

Ventsel, E.S., 1972: Operation research. Moscow, Sov. radio.

Zhukovsky, E.E., Chudnovsky, A.Ph., 1973: On planning optimal irrigation rates. Sbornik trudov po agronomicheskoy fisike, 31, 101-108.

Zhukovsky, E.E., 1974: On comparison of different variants of using meteorological information for economic decision making. Meteorologia i gidrologia, 2, 38-48.

Zhukovsky, E.E., 1975: On rational use of climatological information and meteorological forecasts when programming yields. Selskostopanska nauka (Bolg.), Sofia, 13, 2,71-81.

Zhukovsky, E.E., Kiseleva, T.L., Mandelstam, S.M., 1976: Statistical analysis of random processes as applied to agrophysics and agrometeorology. Leningrad, Gidrometeoizdat.

Zhukovsky, E.E., Chudnovsky, A.Ph.,1978: Methods for optimal use of meteorological information by decision making. Leningrad, Gidrometeoizdat.

Zhukovsky, E.E., Braginskaya, L.L.,1981: On climatologically optimal planning the rates of reseeding winter crops. Nauchno-tekhnichesky bulleten po agronomicheskoy fisike, 45, 20-24.

Zhukovsky, E.E., 1981: Meteorological information and economic decisions. Leningrad, Gidrometeoizdat.

USE OF CLIMATOLOGICAL INFORMATION IN RIVER BASIN PLANNING UNDER RISK*

Lucien Duckstein, Donald R. Davis and Erich J. Plate

Departments of Hydrology and Water Resources
and Systems and Industrial Engineering
University of Arizona, Tucson, Arizona 85721, USA

Institut Wasserbau III, University of Karlsruhe
Kaiserstr · 12, 7500 Karlsruhe, West Germany

1. INTRODUCTION

The purpose of this paper is to demonstrate how climatological forecasts can be used to reduce risk in hydrologic operation and water resources system design. In the first part of the study, the release of a dual purpose reservoir is calculated under natural and sample uncertainty; climatological information may be used to reduce the effect of sample uncertainty. In the second part, the multicriterion ranking of alternative water resources development plans is considered under expected precipitation value, then increased and decreased values. These precipitation changes are estimated from climatological considerations; the criteria include risk-related ordinal and cardinal measures. In two different planning situations, the ranking is found to be affected to a certain extent by climatological forecasting but the effect is more pronounced in the case of short-term hydrologic operation, and then in a semi-arid climate.

2. RISK IN HYDROLOGIC OPERATION: RESERVOIR CONTROL

2.1 The Problem

Consider a reservoir being controlled for the dual purpose of flood control and water supply. The optimum releases corresponding to each objective is determined under natural and sample uncertainty. Then various approaches to reach a compromise release are given, including a "satisficing" release obtained by equating the expected opportunity losses corresponding to the two objectives. Calculations are given in greater detail in Duckstein et al (1983).

2.2 Flood Control Objective

First consider control for the single objective of minimizing flood damages. Let the reservoir with total storage capacity S be full at time zero. As in Russell (1974) or Krzystofowicz and Duckstein (1979), a flood of volume $\underset{\sim}{Q}$ is forecasted, and a release z must be made to alleviate damages downstream. The release causes a damage kz(k<1), while the flood damage is taken as the excess of $\underset{\sim}{Q}$ over available storage space S-z (if any).

The objective is thus:

$$\text{minimize } \phi_1(z,Q) = \begin{cases} kz & \text{if } \underset{\sim}{Q} \le z \\ kz+(Q-z) & \text{if } z \ge \underset{\sim}{Q} \end{cases} \qquad (1)$$

with $0 \le z \le S$.

Let $F_{\underset{\sim}{Q}}(z|\theta)$ be the distribution function of $\underset{\sim}{Q}$ given parameter vector θ, and $f_{\underset{\sim}{Q}}(q|\theta)$ the corresponding pdf. Then the expectation $\theta_1(\cdot)$ with respect to $\underset{\sim}{Q}$ yields the following goal function.

$$E_{\underset{\sim}{Q}}(\theta_1(z,Q)) = g_1(z,\theta) =$$

$$z[k-1+F_{\underset{\sim}{Q}}(z|\theta)] + \int_z^\infty q dF_{\underset{\sim}{Q}}(q|\theta) \qquad (2)$$

To remove the sample uncertainty on θ, a prior pdf $\pi(\theta)$ is assessed (Berger, 1980) either subjectively, or on the basis of information, in particular climatological information (Matalas and Fiering, 1977; Plate, 1982). The Bayes risk is thus obtained:

$$R_1(z) = z[k-1 + E_{\underset{\sim}{\theta}}\{F_{\underset{\sim}{Q}}(z|\theta)\}] + E_{\underset{\sim}{\theta}}[\int_z^\infty q dF_{\underset{\sim}{Q}}(q|\theta)] \qquad (3)$$

The Bayes action z* minimizes Bayes risk, the necessary condition for a minimum is found to be (under the constraint $0 \le z \le S$)

$$E_{\underset{\sim}{\theta}}\{F_{\underset{\sim}{Q}}(z|\theta)\} = 1-k \qquad (4)$$

As an example, let the flood volume $\underset{\sim}{Q}$ be distributed as a one-parameter gamma probability

*This research was supported in part by a grant from the United States National Science Foundation, #CEE 8110778, and funds from the Alexander von Humboldt Foundation in West Germany.

density function (pdf), with distribution function (DF) equal to:

$$F_{\underset{\sim}{Q}}(q|\theta) = 1 - (1+\theta q)e^{-\theta q}, \quad \theta>0 \qquad (5)$$

Eq. 2 becomes

$$g_1(z,\theta) =$$

$$z[k-(1+\theta z)e^{-\theta z}] + \frac{1}{\theta}\{(\theta z)^2+2\theta z+2\}e^{-\theta z} \qquad (6)$$

Let the prior pdf be chosen from the conjugate family of the distribution (5) which is itself the gamma II family:

$$\pi(\theta|a,b) = \frac{b^a \theta^{a-1} e^{-b\theta}}{\Gamma(a)}, \quad \theta>0 \qquad (7)$$

Note that this prior does not hold in a diffuse information case, that is, when a = 1, because the Bayes risk, which is calculated by taking the expectation of Eq. 6 under prior $\pi(\theta|b)$, then turns out to be unbounded: the integral

$$\int_0^\infty g_1(z,\theta)\pi(\theta|b)d\theta \qquad (8)$$

contains the term $\int_0^\infty \frac{e^{-\theta(b+z)}}{\theta} d\theta$ which does not converge.

Assume, next, that climatological data makes it possible to estimate that a=2 in the prior pdf of Eq. 7:

$$\pi(\theta|b) = b^2\theta e^{-b\theta}. \qquad (9)$$

The Bayes risk is calculated by integrating Eq. 6 with respect to the prior $\pi(\theta|b)$ as given in Eq. 9. This integral converges, and the result is:

$$R_1(z) = zk + \frac{2b^2}{b+z} + \frac{b^2 z}{(b+z)^2}. \qquad (10)$$

This function is represented in Figure 1 with numerical values indicated.

The value z* of z which minimizes the Bayes risk is given by the equation

$$u^3 - \frac{3b^2}{k}u + \frac{2b^3}{k} = 0 \qquad (11)$$

in which u = z+b. Using Cardan's solution to the reduced cubic equation (Korn & Korn, 1968, p. 23), one finds

$$\frac{2b}{\sqrt{k}} \cos\frac{\alpha}{3}, \text{ with } \cos\alpha = -\sqrt{k}. \qquad (12)$$

Because of the constraint $z \leq S$, the value z_1^* is:

$$z_1^* = \min[u^* - b, S]. \qquad (13)$$

The optimum Bayes risk may be calculated by integrating the last expression of Eq. 6

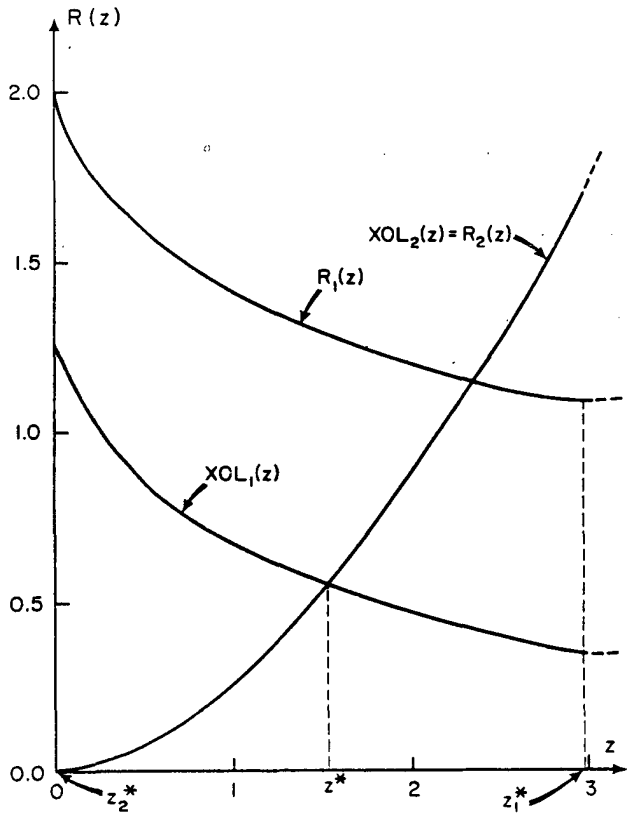


Fig. 1. Bayes Risk R(z) for the case b=1, k=0.16, s=4.

under the prior given in Eq. 9 and letting $z=z_1^*$. Using Eq. 12, $R_1(z_1^*)$ may thus be evaluated as

$$R_1(z_1^*) = \frac{2b^2}{u^*}\left(\frac{z_1^*}{u^*}\right) + \frac{z_1^*}{u^*} + 1). \qquad (14)$$

In this way, the variation of the optimum Bayes risk as a function of b has been established. Climatological data may be used to estimate a and b in the prior of Eq. 7; an examination of standard gamma distribution tables shows that this prior is quite sensitive to the values of a and b, which both increase as more data become available. Furthermore, note that $R_1(Z_1^*)$ is proportional to b^2 in Eq. 14.

2.3 Water Supply Objective

Consider next the control of the same reservoir for the single objective of providing water supply at a later date. The goal is thus to keep the reservoir full, disregarding flood damage, and the corresponding objective is:

$$\text{minimize } \theta_2(z,Q) = \begin{cases} z-S-Q & \text{if } Q \leq z \\ -S & \text{if } z \leq Q \end{cases} \qquad (15)$$

with $0 \leq z \leq S$.

The goal function is now found to be, using the flood DF of Eq. 6:

$$g_2(z,\theta) = z[1-(1+\theta z)e^{-\theta z}]$$

$$+ \frac{1}{\theta}[(\theta z)^2 + 2\theta z + 2]e^{-\theta z} - \frac{2}{\theta} . \qquad (16)$$

The Bayes risk, under the prior of Eq. 9, is evaluated as:

$$R_2(z) = z + \frac{2b^2}{b+z} + \frac{b^2 z}{(b+z)^2} - 2b$$

$$= R_1(z) + (1-k)z - 2b. \qquad (17)$$

This function is represented in Figure 1. As it can be verified using Eq. 12 with k=1, the minimum of this function for $0 \leq z \leq S$ occurs at the origin: $z = z_2^* = 0$. This result may be expected since, for strictly water supply purpose, the optimum policy is not to release any water at all.

2.4 Trade Off Between Water Supply and Flood Control

As shown in Figure 1 (for a case when $z_1^* < S$), the objectives $R_1(z)$ and $R_2(z)$ are in conflict in the range $z_2^* = 0 \leq z \leq z_1^*$. The corresponding set of nondominated objectives is shown in Figure 2.

At this point a set of multiobjective techniques is available to find a "satisficing" solution (Gershon and Duckstein, 1983). For example, multi-attribute utility theory has been applied to a similar reservoir control problem in

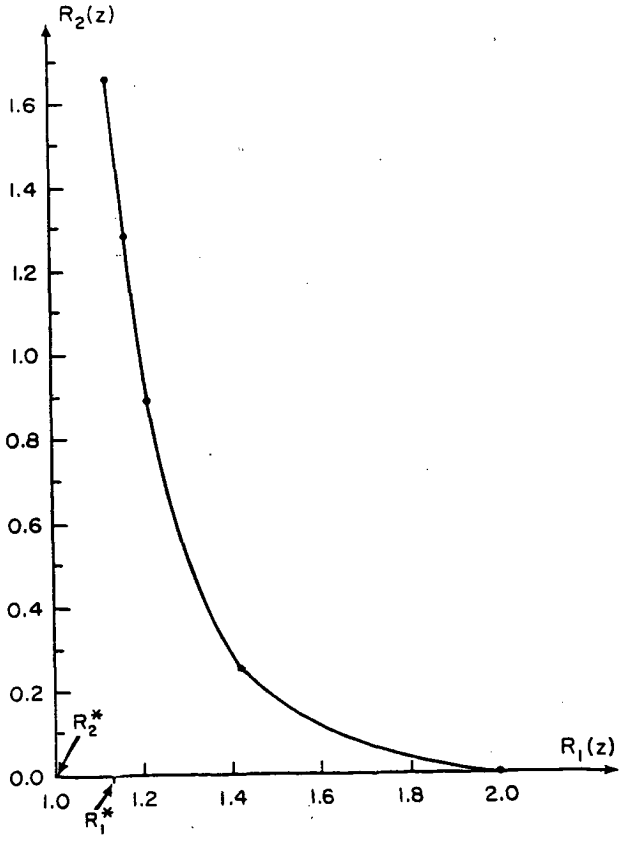


Fig. 2. Set of Nondominated Objectives.

Krzysztofowicz and Duckstein (1979). In the present study, let a distance-based technique be chosen, as in Davis et al (1983); to be specific, compromise programming (Zeleny 1973, 1982) is selected. The objective of this technique is to minimize the weighted sum of the pth power of the normalized objectives:

$$\text{minimize } \ell_p(z) = [w_1^p\, d_1^p(z) + w_2^p\, d_2^p(z)]^{1/p} \qquad (18)$$

$$\text{with} \quad d_i(z) = \frac{R_i(z) - R_i^*}{R_{i*} - R_i^*} , \; i = 1,2 , \qquad (19)$$

and R_i^* = minimum risk, R_{i*} = maximum risk. The result of the transformation (19) is represented graphically in Figure 3. In the numerical case used earlier the ideal point (R_1^*, R_2^*) corresponds to the origin in the (d_1, d_2) plane.

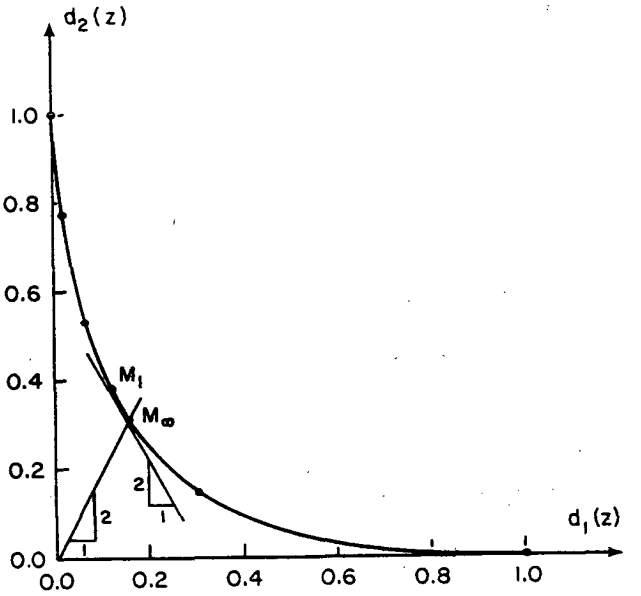


Fig. 3. Normalized Payoff Set and Compromise Programming Solutions for p=1, p=∞.

Let p = 1, then the compromise release z_0^* is found to be given by the equation:

$$u^3 - \frac{3b^2}{k_1+k} + \frac{2b^3}{k_1+k} = 0, \; u = z+b \qquad (20)$$

Using the same formula as in Eq. 12, with $w_1=2$, $w_2=1$, the numerical example leads to the following results:

$$k_1 = 0.1693,\; z_0^* = 1.604,\; R_1(z_0^*) = 1.261,$$

$$R_2(z_0^*) = 0.6088,\; d_1(z_0^*) = 0.119,\; d_2(z_0^*) = 0.3663.$$

This solution is represented by point M in Figure 3, at which the tangent to the non-dominated set has a slope of $-(w_1/w_2) = -2$. The compromise solution for p=2 would have to be found numerically.

For $p = \infty$, which is the minimax case, the algebraic solution is found from the equation:

$$w_1 \frac{R_1(z) - R_1^*}{R_{1*} - R_1^*} = w_2 \frac{R_1(z) + (1-k)z - 2b}{R_{2*}} \qquad (21)$$

Geometrically, one finds the point M_∞ at the intersection of the line $w_1 d_1 = w_2 d_2^\infty$ and the set of non-dominated solutions.

2.5 Value of Perfect Information and Satisficing Decision

As it is known in statistical decision theory and discussed within a hydrological context in earlier papers by the same authors, an opportunity loss is incurred because of uncertain knowledge on the state of nature θ. This uncertainty is then compounded in multiobjective problems (Davis et al, 1983) because of the non-unique way of choosing a "satisficing" solution. Even when a technique such as compromise programming is selected, there still is a choice in the value of parameter p. Thus, using an ℓ_p-norm, the expected opportunity loss EOL is, by definition:

$$EOL = \min_z E_\theta[\ell_p(g_1(z,\theta), g_2(z,\theta)] - E_\theta\{\min_z \ell_p(g_1(z,\theta), g_2(z,\theta))\}.$$

This non-negative expression may be evaluated in closed form for p=1, p=∞, or numerically for p=2, but the interpretation of EOL depends upon the meaning of ℓ_p to the decision-maker.

To keep in line with the multiobjective nature of the release decision, let a vector-valued EOL be defined and called, as in Davis et al (1983), the Expected Multiobjective Opportunity Loss, EMOL:

$$EMOL(z) = \{XOL_1(z), XOL_2(z)\} \qquad (22)$$

with

$$XOL_i(z) = R_i(z) - \int_\Theta \min_z g_i(z,\theta)\pi(\theta)d\theta, \quad i=1,2. \qquad (23)$$

To evaluate the right-hand side integral, one calculates:

$$\min_z g_1(z,\theta) = \frac{1}{\theta}[v^2 e^{-v} + 2k], \text{ with } (1+v)e^{-v} = k$$

$$\min_z g_2(z,\theta) = 0.$$

Using the numerical example, v=0.29, $g_1 = 0.735/\theta$ and Eq. 22 becomes

$$EMOL(z) = [R_1(z) - 0.735, R_2(z)].$$

These two components of EMOL(z) are represented in Figure 1. Again, compromise programming may be used to obtain a satisficing release. More simply, one may choose the release z* which corresponds to equating the XOL's and whose value is: z* = 1.51. Such a procedure has two advantages: first it provides a balance between the expected values of perfect forecast for the elements of EMOL and, second, it facilitates the choice of a satisficing solution. The release z* = 1.51 is quite different from the value of 2.37 obtained by equating the Bayes risk.

Climatological information changes the curves shown in Figures 1, 2 and 3. In particular, more accurate forecasts lower the $XOL_i(z)$ curves in Figure 1, hence the elements of the EMOL(z) vector (Eq. 22). At the same time, the satisficing release, say z*, also changes. A systematic framework for investigating the effect of climatological information on reservoir operation has thus been given. Next, the case of water resources system design is examined.

3. CASE STUDY OF THE SANTA CRUZ RIVER BASIN

3.1 Problem Description

In the Santa Cruz River Basin, which includes the Tucson metropolitan area (550,000 inhabitants), as in many semi-arid communities, water supplied by imports and natural recharge fall short of demand, while flooding remains an impending threat. In 1975, the depletion-supply ration was 3.3 to 1, the three primary users being agriculture (41%), municipalities (29%), and mining (27%) (Duckstein et al, 1981). Originally, the regional water resources problem had been formulated through interaction with Tucson Urban Study - Corps of Engineers group (Corps of Engineers, 1978). Twenty-five alternative plans had been generated by pairing five flood-control actions with each of five water supply actions. These alternatives were then evaluated against thirteen qualitative criteria and then ranked by the multicriterion technique ELECTRE in Gershon et al (1982). However, as pointed out by Alley (1982) in a discussion of the cited paper, 6 of the 25 alternatives were found to be dominated right at the beginning of the ELECTRE I analysis. Thus, to examine the effect of climatological forecasting on the ranking of alternative 50-year development plans, only 19 non-dominated alternatives are considered and renumbered as indicated in Table 1, in which the originally agreed-upon plans are numbered in parentheses. The triplet objectives-specifications-criteria may be found in Gershon et al (1982).

3.2 Use of Climatological Forecast

If climatological forecasts call for increased precipitation over the planning horizon, the elements of the original 13 x 19 payoff matrix found in Duckstein et al (1983) are modified. Thus, the drops in aquifer levels and urban water quality decrease for the "no action" (on flood control) alternatives. The flood losses and frequency increase for levee and floodplain management alternatives. Implementation costs increase for river training alternatives, operation and maintenance costs increase for levee and flood management alternatives. Implementation costs increase for river training measures and the Central Arizona Project (CAP). Indirect costs increase for the CAP because of wider demand fluctuations. Natural resources increase for

Table 1. Non-dominated alternative plans.
Numbering in parenthesis indicates the original definition (Gershon et al, 1982).

ALTERNATIVES	LEVEE CONSTRUCTION	FLOOD CONTROL CHANNELIZATION	RESERVOIRS AND DAMS	FLOOD PLAIN MANAGEMENT	NO ACTION
Waste water reclamation	1(1)	2(2)	3 (3)	4(4)	5 (5)
Groundwater development	(6)	(7)	(8)	6(9)	(10)
Central Arizona Project	7(11)	8(12)	9(13)	10(14)	11(15)
Conservation and education	(16)	(17)	12(18)	13(19)	14(20)
No action	15(21)	16(22)	17(23)	18(24)	19(25)

conservation and education schemes. The preservation of present facilities will decrease if levees and dams are built, so will be the new recreational opportunities in the case of floodplain management alternatives. These changes are indicated in the non zero entries of Table 2.

3.3 Results of Multicriterion Analysis ELECTRE II

The multiple criteria technique ELECTRE II (Roy and Bertier, 1971; Gershon and Duckstein, 1983) applied to the original pay-off matrix and

Table 2. Changes in original table with increased precipitation.

		SYSTEMS																		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1.	Aquifer level					3.0	1.6					3.0			3.0					2.0
2.	Urban W. Qual.					d						d								d
3.	Agr. W. Qual.																			
4.	Exp. Flood L.	15.44						15.44								15.44				
5.	Exp. Freq.	0.04	0.02			0.08		.04	.02			.08				.04	.02			.08
6	Pres. Areas			b						b			b					b		
7	Wild. & Veg.																			
8.	Impl. Costs	14.7	18.8					30.	34.2							14.5	18.6			
9.	O & M Costs	39					5.6	5.2			4.9			4.9		3.2			1.1	
10.	Indir. Costs							c	c	b	d	c								
11.	Nat. Res.													b	b					
12.	Exist. Fac.			d						d			d					d		
13.	New Opp.				c		d				d								d	

The results of a similar analysis performed for the case of a drier climatological forecast lead to Table 3. For example, flood losses and frequency are bound to decrease and new recreational opportunities are enhanced by floodplain management, since less flooding is expected to occur.

then to Tables 1 and 2 yields, for the first seven preferred alternatives, the rankings shown in Figure 4 with a geometric coding for the types of alternatives. It appears that the channelization alternatives (2) (16) are preferred for the expected and more humid cases, while for the drier cases, good compromise are reservoirs and dams (17) followed by combinations of floodplain management with groundwater development (16) (including recharge enhancement) or with conservation and education (13).

Table 3. Changes in original table with decreased precipitation.

		SYSTEMS																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
CRITERIA	1. Aquifer level						3.0	3.5	2.4	2.8	3.0	4.0	2.8	3.24		2.8	2.8	2.4	3.0	3.2
	2. Urban W. Qual.		b	3	b	3														
	3. Agr. W. Qual.						c													
	4. Exp. Flood L.				15.44	19.45	15.44				15.44	19.45		15.44	19.45				15.44	19.45
	5. Exp. Freq.	.008	.008	.002	.015	.03	.015	.008	.008	.002	.015	.03	.002	.016	.03	.008	.008	.002	.015	.03
	6. Pres. Areas																			
	7 Wild. & Veg.		e						d								e			
	8. Impl. Costs																			
	9. O & M Costs																			
	10. Indir. Costs		c						d						b		d			
	11. Nat. Res.	b	b	b	a		e													
	12. Exist. Fac.		a	b					a	b			b				a	b		
	13. New Opp.						b							a					b	

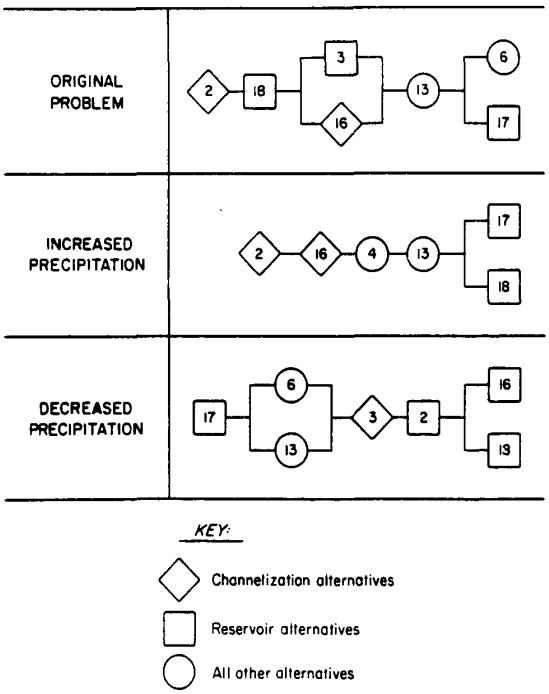


Fig. 4. Preferred Systems without and with Climatological Information for the Tucson Basin Case Study.

It is thus seen that climatological forecasting may cause significant changes in the ranking of alternative systems at least for this semi-arid zone case study. Furthermore, certain alternatives appear to be more robust than others with respect to climatological changes (Fiering, 1982); for example, alternatives (2) and (13) are always ranked among the first four (counting ties as one alternative). The next section deals with a smaller basin, with more temperate climate and higher population density.

4. MARCHFELD (AUSTRIA) CASE STUDY

4.1 Problem Description

The Marchfeld is a 950 Km^2 area situated in East Austria at the north of the Viennese Basin. The region has 57,000 inhabitants. The borders are the Danube in the south and west, the river March in the east, and a hilly region in the north. Agricultural production is highly developed and is an important part of the regional economy. Water requirements for irrigation are covered by exploiting a regional aquifer which also serves to supply industrial and domestic water users. As a result of increasing water requirements, overdraft already occurs and will become worse in the future if no remedial action is taken. In Nachtnebel et al (1982a), the four objectives and twelve criteria given in Table 4 are defined.

Next, 20 alternative systems are developed by combining four fundamental actions at two or three levels of implementation:

- Action A_1 consists of a gravity canal transferring water from the Danube into the Marchfeld region.
- Action A_2 is aimed at increasing water use efficiency, especially in industrial processes (reuse and recycling).
- Action A_3 is concerned with the planting of wind breakers to reduce wind erosion and improve the microclimate.
- Action A_4 consists of modifying land use techniques to decrease agricultural water consumption.

Table 4. Objectives, criteria and units or ratings.

OBJECTIVES	NUMBER	CRITERIA	UNIT OR RATING
National Objective	1	Equity in regional agricultural income	A B C
	2	National Supply Rate	A B C
Regional Objective	3	Manpower Requirement	A B C D
	4	Required Area of Farmland	hectares
	5	Robustness+	A B C
	6	Independence ++	BOD(mg/ℓ) 0-2 A, 2-6 B, 6-13 C, 13- D
Environmental Objective	7	Surface Water Quality	Nitrate Concentration 0-15 A, 15-30 B, 30-50 C, 50-mg/ℓ D
	8	Groundwater Quality	meter/year
	9	Decline of Groundwater Table	A B C D
	10	Preservation of Landscape	A B C D
Satisfaction of Water Requirements	11	Costs of Supply plus Losses of Shortage	Monetary (Austrian Schillings)
	12	Water use Factor	Ratio of annual groundwater depletion to total requirement 0=best 1=worst

+ robustness to internal modifications, uncertainties and violation of assumptions on variables such as increase of population, agricultural production, etc.
++independence of future unforeseeable external projects such as an international navigation canal, additional power stations.

4.2 Multicriterion Analysis by Compromise Programming

Table 5 gives the rating of 20 alternative systems numbered as in Nachnebel et al (1982a), as well as, in the last two rows, the results of a compromise programming analysis (Zeleny, 1973, 1982; Duckstein and Opricovic, 1980). The principle of this technique is to minimize a measure of distance between an "ideal" solution $\{Z(i)^*\}$ and the alternative j under consideration, with ratings $\{Z(i,j)\}$:

$$\ell_p(j) = \left\{ \sum_{i=1}^{I} \alpha(i)^p \left| \frac{Z(i)^* - Z(i,j)}{Z(i)^* - Z(i)_*} \right|^p \right\}^{\frac{1}{p}} \quad (25)$$

where $Z(i)_*$ is the worst value of criterion i, and $\alpha(i)$ is the weight of that criterion, as shown in Table 5. Application of compromise programming yields the following ranking of the first seven preferred systems:

$$\begin{array}{ll} p = 1: & 14, 16, 10, 13, 15, 8, 12 \\ p = 2: & 14, 16, 10, 13, 12, 15, 8. \end{array} \quad (26)$$

essentially, then, ℓ_1 and ℓ_2 yield very similar rankings in which alternatives including importation of water from the Danube combined with one or two of the other fundamental actions are ranked high. Note that, as shown in Nachtnebel et al (1982a, 1982b) the rankings obtained by ELECTRE and compromise programming are quite comparable.

Next, it is assumed that increased precipitation is forecasted, so that changes in Table 5 are made as indicated in Table 6. The main changes affect the uncertainty-related criteria, as noticed earlier. Under these new conditions, the application of compromise programming leads to the results shown in the last two rows of Table 6, and the following rankings are obtained:

$$\begin{array}{ll} p = 1: & 14, 16, 10, 8, 13, 15, 12 \\ p = 2: & 10, 14, 16, 8, 13, 11, 4. \end{array} \quad (27)$$

Here the ranking for the two values of p are somewhat different. In particular, for p = 2, which provide more "balance" between various risks than p = 1 (Freimer and Yu, 1976), alternative 10, corresponding to a smaller Danube canal combined

Table 5. Systems versus criteria for the original Marchfeld problem and results and analysis.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Weights
Equity in income	B	C	A	A	B	C	A	A	A	A	B	B	A	A	B	B	A	C	B	C	2
National Supp. Rate	B	A	B	A	B	A	B	A	C	A	C	A	C	A	C	A	C	B	C	B	2
Manpower	A	A	B	B	A	A	C	B	B	B	C	C	C	B	D	D	C	C	D	D	1
Required Farm Land	0	0	850	850	0	0	850	850	160	160	1010	1010	160	160	1010	1010	220	220	1070	1070	1
Robustness	C	C	C	B	C	B	C	B	C	B	B	A	B	A	A	A	C	C	B	B	1
Independence	C	C	C	C	C	C	B	C	C	B	A	B	A	B	A	B	A	B	A	B	2
Surface Water Qual.	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	II	II	II	II	III	III	III	III	2
Groundwater Quality	D	D	D	D	C	C	C	B	B	C	D	C	C	B	C	B	D	C	C	C	2
Change in Water table	-.29	-.20	-.26	-.17	-.23	-.13	-.20	-.11	.04	.07	.05	.07	.10	.13	.11	.13	.22	.22	.21	.21	2
Pres. of Landscape	D	C	C	A	D	C	C	A	C	B	C	B	C	C	A	A	D	C	C	A	2
Cost 10^6 S/A	147	174	150	160	181	208	184	194	176	188	175	188	210	222	221	223	204	210	212	218	2
Water Use Factor	1	1	1	1	1	1	1	1	.44	.39	.47	.43	.31	.23	.33	.25	.28	.28	.30	.67	1
p = 1	14.0	13.6	13.4	10.5	13.6	12.7	12.3	9.04	10.5	7.42	12.1	9.10	7.72	4.95	8.51	6.13	10.3	10.8	11.5	10.8	
p = 2	4.90	4.79	4.80	4.05	4.56	4.47	3.92	3.59	3.99	2.86	4.04	3.19	3.00	2.49	3.48	2.70	4.04	3.68	4.05	3.67	

Table 6. Modifications in Table 5 in case of increased precipitation and results of analysis.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Weights
Equity in income	C	C	B	B	C	C	B	B													
National Supp. Rate																		A		A	
Manpower																					
Required Farm Land																					
Robustness			B	A			B	A			A						B		A	A	2
Independence																					
Surface Water Qual.			III	III			III	III			III										1
Groundwater Quality									C			D		C							
Change in Water table				-.11				-.05				.09		.16							1
Pres. of Landscape			B				B				B										1
Cost 10^6 S/A		183		172		219		199		197		197		233		234		220		229	
Water Use Factor				.95				.95	.40	.35	.42	.39	.28	.21	.29	.22	.23	.25	.27	.60	
p = 1	13.0	12.4	11.03	13.4	12.7	11.1	10.0	7.41	10.1	6.62	8.88	8.77	7.50	5.47	7.52	6.35	7.02	10.0	10.1	9.2	
p = 2	4.58	4.56	3.75	3.43	4.30	3.85	3.99	2.95	3.89	2.41	3.39	1.66	3.06	2.57	3.48	2.84	3.56	3.75	3.76	3.48	

with land use control, becomes ranked first. On the other hand, for p = 1, which is the linear - hence robust - case, the ranking is almost the same as in the original study (Eq. 26). Overall,

changes in the ranking of the first seven preferred alternatives are rather minor.

5. DISCUSSION AND CONCLUSIONS

Medium range climatological forecasting, say over one season (Baumgartner et al, 1982), can certainly be used to improve hydrologic operation. In particular, such forecasts can be used to reduce the vector-valued risk corresponding to flood control and water supply, respectively. In the example given in the first part of the paper, climatological forecasting could lead to finding a better prior pdf $\pi(\theta)$ on the parameters of the DF of inflow Q, $F_{\underset{\sim}{Q}}(q|\theta)$.

Long range climatological forecasting (several year-long trends) may also influence the multicriterion ranking of alternative development plans. This effect appears to be more pronounced in the semi-arid setting of the Santa Cruz River Basin than in the more densely populated Marchfeld area. Overall, as pointed out in Matalas and Fiering (1977), there seems to be enough flexibility, robustness or resilience (Fiering, 1982) in a comprehensive planning process to absorb long-term climatological changes over the project horizon. However, the two planning examples presented in this study are at a prefeasibility level. As the preferred alternatives are studied in greater detail and the various risk-related criteria are evaluated quantitatively, it is quite possible that more sensitivity to climatological factors appears. For example, it may be necessary to model local environmental problems (Plate, 1982) to assess the consequence of an action such as A_3 (planting windbreakers) on the microclimate.

This study leads to the following concluding points:

1. A dual objective operation of a reservoir for flood control and water supply, under natural and sample uncertainty on inflow Q, leads to a trade-off between either two Bayes risks, or two expected opportunity losses.
2. The compromise release may be chosen so as to equalize the components of the expected multiobjective opportunity loss EMOL, rather than the risks themselves, since this procedure equalizes the worth of perfect forecasts.
3. The compromise release is quite sensitive to the prior pdf the estimation of which, in turn, can be improved by climatological forecasting.
4. The ranking of 19 alternatives with respect to 13 criteria in the case of the Tucson basin, using the ELECTRE technique, is fairly sensitive to changes in climatological or long-range precipitation forecasts.
5. A ranking of 20 alternatives with respect to 12 criteria in the case of the Marchfeld area, using compromise programming, appears to be only moderately sensitive to a long-range forecase announcing increased precipitation.
6. The approach, applied so far to operation and and prefeasibility studies only, can be applied to detailed hydrologic design.
7. Quantitative techniques are available to assess the contribution of climatological forecasting to the expected reduction of hydrologic risk.

6. REFERENCES

Alley, William M., 1983. "Comment on 'Multiobjective River Basin Planning With Qualitative Criteria' by M. Gershon, L. Duckstein, and R. McAniff", Water Resources Research, Vol. 19, No. 1, pp. 293-294.

Baumgartner, A., G. Enders, M. Kerchner and H. Mayer, 1982. "Global Climatology", Ch. 4, Engineering Meteorology, (E. J. Plate, Ed,), pp. 125-276.

Davis, D., F. Szidarovszky, and L. Duckstein, 1983. "A Distance-based Measure of the Effect of Uncertainty on Multiobjective Decision Problems", 11th IFIP Conference on System Modelling and Optimization, Copenhagen, Denmark, July 25-29.

Duckstein, L., D. R. Davis and E. J. Plate, 1983. "Sensitivity of River Basin Planning and Management under Risk to Climatological Information", Working Paper 83-010, Systems and Industrial Engineering Dept., University of Arizona, Tucson, AZ 85721.

Duckstein, L., M. Gershon, and R. McAniff, 1981. "Development of the Santa Cruz River Basin: A Comparison of Multicriterion Approaches", Report #51, Engineering Experiment Station, University of Arizona, Tucson.

Duckstein, L., and S. Opricovic, 1980. "Multiobjective Optimization on River Basin Development", Water Resources Research, Vol. 16, No. 1, pp. 14-20.

Fiering, M. B., 1982. "Alternative Indices of Resilience", Water Resources Research, Vol. 18, No. 1, pp. 33-39, February.

Freimer, M., and P. L. Yu, "Some New Results on Compromise Solutions for Group Decision Problems", Manage. Sci., 22(6), 688-693, 1976.

Gershon, M. and L. Duckstein, "Multiobjective Approaches to River Basin Planning", ASCE J. Water Resources Planning & Management, January, Vol. 109, No. 1, pp. 13-28.

Korn, G. and T. Korn, 1968. Mathematical Handbook for Scientists and Engineers, McGraw-Hill Book Company, 2nd Edition.

Krzysztofowicz, R. and L. Duckstein, 1979. "Preference Criterion for Flood Control Under Uncertainty", Water Resources Research, Vol. 15, No. 3, pp. 513-520.

Matalas, N. and M. Fiering, 1977. Water-Resource Systems Planning, Climate, Climatic Change and Water Supply, National Academy of Sciences, Washington, D.C.

Nachtnebel, H. P., I. Bogardi, and L. Duckstein, 1982a. "Multicriterion Analysis for Regional Water Resources Development. Part I: Cost-Effectiveness Approach" in Environmental Systems Analysis and Management, S. Rinaldi, Editor, North Holland Publishing Co., New York, pp. 71-81.

Nachtnebel, H. P., L. Duckstein, and I. Bogardi, 1982b. "Evaluation of Conflicting Regional Water Requirements: An Austrian Case Study", Optimal Allocation of Water Resources, Proceedings of the Exeter Symposium, July 1982, IAHS Publ. No. 135.

Plate, E. J., 1982. "Wind Tunnel Modelling of Wind Effects in Engineering", Ch. 13, Engineering Meteorology, (E. J. Plate, Ed.), pp. 573-640.

Roy, B., and P. Bertier, 1971. "La Méthode ELECTRE II", Direction Scientifique, Note de Travail No. 142, Groupe, Metra, Paris, France, April.

Russell, S. O., 1974. "Use of Decision Theory in Reservoir Operation", J. Hydraul. Div. Amer. Soc. Civil Eng., 100(HY6), pp. 809-817.

U.S. Army Corps of Engineers, 1978. Urban Study for the Metropolitan Area of Tucson, Arizona, Rep. 12192, Tucson, Ariz., December.

Zeleny, M., 1973. Compromise Programming, in Multiple Criteria Decision Making, J. L. Cochrane and M. Zeleny (Eds.), University of South Carolina Press, Columbia, pp. 262-301.

Zeleny, M., 1982. Multiple Criteria Decision Making, McGraw Hill Book Company, New York, N.Y., 563 p.

USE OF DECISION ANALYSIS TO ASSESS THE ACTUAL AND POTENTIAL VALUE OF SHORT-RANGE CLIMATE FORECASTS

Allan H. Murphy
Department of Atmospheric Sciences
Oregon State University
Corvallis, OR 97331, U.S.A.

Robert L. Winkler
Graduate School of Business
Indiana University
Bloomington, IN 47405, U.S.A.

Richard W. Katz
Department of Atmospheric Sciences
Oregon State University
Corvallis, OR 97331, U.S.A.

1. INTRODUCTION

Studies of the value of climate information can have important implications for decisions made by a variety of individuals, including program managers, researchers, weather service personnel, and users of such information. Therefore, it is surprising to discover that relatively few investigations of the value of climate information have been conducted. Moreover, the information-value studies undertaken in this context generally have been based upon informal procedures involving "back-of-the-envelope" calculations and/or surveys of actual and potential users. It is only recently that formal decision-making models, other than the simple cost-loss ratio model, have been applied to the problem of quantifying the value of climate information (e.g., see Agnew and Anderson, 1977; Phillips et al., 1978).

The purposes of this paper are to describe briefly the decision-analytic approach to information-value problems and to present an application of this methodology to a decision-making situation involving the use of climate forecasts. A more complete treatment of this problem is contained in a paper to appear shortly in the Journal of Climatology (Winkler et al., 1983). In addition, the authors recently have applied this methodology to a decision-making problem involving the use of short-range weather forecasts (Katz and Murphy, 1979; Katz et al., 1982). Decision analysis is discussed in Section 2, with emphasis on its informational aspects. An example in which the methodology of decision analysis is used to assess the value of climate forecasts is described in Section 3. Section 4 contains a brief discussion of possible extensions of this work and of some issues related to information-value problems in meteorology.

2. DECISION ANALYSIS AND THE VALUE OF INFORMATION

Decision analysis is a procedure for analyzing and modeling complex, dynamic decisions in an uncertain environment. The basic principle underlying this procedure is "divide and conquer," which involves dividing the problem of concern into its basic elements, making certain quantitative assessments, and then combining these assessments in a logical and rational manner. Keeney (1982) provides a recent review of decision analysis and defines it as "a philosophy, articulated by a set of logical axioms, and a methodology and collection of systematic procedures, based upon those axioms, for responsibly analyzing the complexities inherent in decision problems." For a recent introduction to decision analysis with a meteorological flavor, see Winkler and Murphy (1983).

In the decision-analytic approach, a decision-making problem is divided into sets of basic elements: actions, events, consequences or outcomes, probabilities (of events), and utilities (of consequences). The actions represent the alternatives available to the decision maker or user. In a meteorological context, the events generally represent values or ranges of values of climate or weather variables, although other factors also may be considered in defining the set of events. Each possible combination of actions and events leads to a consequence. The utilities are defined on these consequences and express quantitatively the decision maker's preferences for the outcomes. The probabilities represent the decision maker's subjective degrees of belief regarding the likelihood of occurrence of the relevant events. In meteorological applications, these probabilities usually are based on, or derived from, information provided by a meteorologist. Finally, a decision criterion must be adopted to determine the recommended course of action. In decision analysis, the expected-utility criterion is employed, according to which the decision maker selects the action associated with the largest expected utility. The expected utility of an action is simply the probability-weighted average of the utilities of the consequences associated with that action.

Structuring the decision-making problem and assessing the utilities and probabilities are key aspects of decision analysis. In brief, the structuring or modeling process involves the identification of potential actions and relevant events, the consideration of logical relationships among actions and events (e.g., the construction of a decision tree with various actions and events in proper sequence), and the determination of consequences for the various combinations of actions and events. Once the structuring phase is completed, it is necessary to assess a utility function over the consequences and a probability distribution over the events. An important feature of decision analysis is that it can incorporate both objective information about the events (e.g., historical data) and subjective information that may be available only in the form of human judgments.

Since we are concerned here primarily with the value of information, special emphasis will be placed on the informational aspects of decision

analysis. This phase of the decision-analysis process involves several steps and types of information. First, it is assumed that the decision maker's beliefs regarding the occurrence of the relevant events, before receiving any information from the meteorologist, are expressed in terms of prior probabilities. Then the potential pieces of information that can be provided by the meteorologist must be identified. That is, what are the available items of climate data? What are the possible forecasts that can be made? Next, the likelihoods of the various pieces of information must be determined. For example, how likely is it that a particular item of climate data or a specific forecast will be provided, given the subsequent occurrence of each of the possible events. Finally, the likelihoods are combined with the decision maker's prior probabilities of the relevant events, using Bayes' theorem, to obtain revised or posterior probabilities of the events.

In order to calculate the value of information, it is first necessary to determine the action providing the highest expected utility under the decision maker's prior probabilities. The expected utility associated with this action is the prior expected utility. Then, for each possible piece of information, the action must be found with the highest expected utility under the posterior probabilities following receipt of that piece of information. Next, these "highest expected utilities" are averaged over the possible pieces of information, using the marginal (or predictive) probabilities of the pieces of information as weights. In essence, these predictive probabilities represent the relative frequency with which the various pieces of information (or possible forecast values) are used by the meteorologist. The expected value of information is the cost of information that would make the prior expected utility and the "average" posterior expected utility (i.e., the average of the above-mentioned highest expected utilities) exactly equal. In other words, it is the most that the decision maker should be willing to pay for the information. Note that this analysis is conducted in an ex ante or expected-value sense. That is, it is an analysis of the expected value of the information before the data or forecasts (and observations) are actually available.

Several points related to the value of information should be mentioned. First, sensitivity analysis is a vital phase of the decision-analysis process, because it is important in any application of this methodology to investigate the sensitivity of the estimates of the value of information to changes in the various inputs (e.g., the probabilities, the utilities). Second, a strong motivation for obtaining estimates of the value of information is the use of these estimates as a guide to the development of improved information. Third, it should be noted that information is of value only if it changes the decision; that is, information itself has no intrinsic value. For example, if the decision made on the basis of the information received from the meteorologist is always the same as the decision made on the basis of the user's prior probabilities, then the information provided has no value. Finally, although perfect information (i.e., forecasts that are always correct) is seldom available, the expected value of such perfect information provides a useful upper bound for the expected value of any information.

3. VALUE OF CLIMATE INFORMATION: AN EXAMPLE

3.1 The Decision-Making Problem

We will focus here on a situation in which a farmer in the Midwest or on the Great Plains of the U.S. must decide whether to plant wheat (W) or barley ($\overline{W}$). The meteorological event of concern is the amount of precipitation during the growing season, which will be characterized as either above (A) or below ($\overline{A}$) average. This situation is depicted in a decision tableau in Figure 1. The entries r, s, t, and u represent (net) payoffs or

		Events: Above average precipitation (A)	Events: Below average precipitation ($\overline{A}$)	Expected returns ER
Actions	Plant wheat (W)	r	u	$ER(W) = p_o r + (1-p_o)u$
	Plant barley ($\overline{W}$)	s	t	$ER(\overline{W}) = p_o s + (1-p_o)t$
	Prior probabilities	p_o	$1-p_o$	

Figure 1. The decision tableau, including the expected returns associated with the actions under prior information, for the agricultural decision-making situation presented in the text.

returns to the farmer associated with the respective consequences. That is, the farmer receives r $/acre if he plants wheat and precipitation is above average; s $/acre if he plants barley and precipitation is above average; t $/acre if he plants barley and precipitation is below average; and u $/acre if he plants wheat and precipitation is below average. The quantity p_o in Figure 1 is the farmer's initial or prior probability of the occurrence of above average precipitation during the growing season.

We assume that the farmer wants to maximize his expected utility and that his utility function for money is linear in the relevant range (i.e., in the range between the largest and smallest payoffs, inclusive). Thus, maximizing expected return is equivalent to maximizing expected utility. In Figure 1, ER(W) and ER($\overline{W}$) denote the expected returns (ERs) associated with planting wheat and barley, respectively. The expected return corresponding to each action is just the probability-weighted average of the two possible payoffs (r and u for wheat, s and t for barley).

For the purposes of this example, we make the realistic assumption that $r > s > t > u$ (>0). That is, if the farmer knew ahead of time that precipitation would be above average he would plant wheat ($r > s$), whereas if he knew that it would be below average he would plant barley

$(t > u)$. In addition, the return under above average conditions is greater than that under below average conditions whether wheat or barley is planted. For convenience, we also assume that the farmer's initial probability of above average precipitation p_o is 0.5. In this case, the expected returns to the farmer are

$$ER(W) = 0.5r + 0.5u = (r + u)/2 \qquad (1)$$

and

$$ER(\overline{W}) = 0.5s + 0.5t = (s + t)/2. \qquad (2)$$

The farmer should choose the larger of the two expected returns in (1) and (2). Thus, the best that the farmer can do is obtain ER*, where

$$ER^* = \max\{(r + u)/2,\ (s + t)/2\}. \qquad (3)$$

We shall refer to the quantity ER* in (3) as the farmer's prior expected return (i.e., the expected return prior to the receipt of information from the meteorologist).

3.2 Value of Forecasts: General Approach

How should the farmer react to new information? In particular, suppose that a meteorologist provides the farmer with a climate forecast concerning the likelihood of occurrence of above average precipitation during the growing season. Let p denote the probability of above average precipitation according to this information (i.e., p is the forecast probability of above average precipitation). We assume that the forecast provided by the meteorologist is reliable in a probabilistic sense, in that the probability of above average precipitation given a specific value of p is equal to p (i.e., it is assumed that above average precipitation occurs on 30% of the occasions on which the meteorologist assigns a 0.3 probability to this event and that this correspondence holds for all values of p). This assumption may not be too unrealistic in view of the reliability of short-range probabilistic weather forecasts (e.g., see Murphy and Winkler, 1977; Murphy, 1981). Finally, it is assumed that the farmer employs the probability p in making his decision as to whether to plant wheat or barley (i.e., he adopts p as his own posterior probability, thereby shifting from his prior probability of $p_o = 0.5$).

After the farmer has adopted the probability p, his expected returns are

$$ER(W|p) = pr + (1 - p)u \qquad (4)$$

for planting wheat and

$$ER(\overline{W}|p) = ps + (1 - p)t \qquad (5)$$

for planting barley. Now the best that the farmer can do is to obtain ER*(p), where

$$ER^*(p) = \max\{(pr + (1 - p)u,\ ps + (1 - p)t\}, \qquad (6)$$

which obviously depends on p. In order to obtain an expected return over all possible values of p, it is necessary to take the expectation (with respect to p) of ER*(p) in (6), yielding

$$ER^{**} = E[\max\{pr + (1 - p)u,\ ps + (1 - p)t\}]. \qquad (7)$$

This expression gives us the expected value, before seeing p, of ER*(p), and it is denoted by ER**.

To calculate ER**, we need to know the probability distribution of p. This distribution, denoted by f(p), indicates the relative likelihood that the farmer will receive information from the meteorologist specifying particular values of p. For example, the distribution of p might be discrete - a 10% chance that p will be 0.3, a 20% chance that p will be 0.4, a 40% chance that p will be 0.5, a 20% chance that p will be 0.6, and a 10% chance that p will be 0.7; or it might be continuous - a normal distribution with mean 0.5 and standard deviation 0.1. For consistency with the prior probability $p_o = 0.5$, the mean of the distribution of p must be 0.5. As the variance of f(p) increases, the accuracy of the forecasts increases. In particular, the lower limit of accuracy is represented by the prior information [i.e., f(0.5) = 1], whereas the upper limit of accuracy - namely, perfect information - is represented by f(0) = 0.5 and f(1) = 0.5.

To complete the operation in (7), we must find the "breakeven point" (or probability) p_b for which the farmer is indifferent between planting wheat or barley. This value is obtained by equating the expected returns associated with the two actions [i.e., Eqs. (4) and (5)], which reveals that

$$p_b = (t - u)/(r - s + t - u). \qquad (8)$$

When $p > p_b$, planting wheat maximizes ER, whereas when $p < p_b$, planting barley maximizes ER. Therefore, ER*(p) in (6) can be expressed in the form

$$ER^*(p) = \begin{cases} pr + (1 - p)u & \text{if } p \geq p_b, \\ ps + (1 - p)t & \text{if } p < p_b, \end{cases} \qquad (9)$$

with p_b given by (8). Thus, in taking the expected value in (7), we will be taking the expected value of pr + (1 - p)u over the values of p such that $p \geq p_b$ and the expected values of ps + (1 - p)t over values of p such that $p < p_b$. Mathematically, this expectation can be written as follows:

$$ER^{**} = \sum_{p < p_b} [ps + (1 - p)t]f(p) + \sum_{p \geq p_b} [pr + (1 - p)u]f(p) \qquad (10)$$

when f(p) is discrete and

$$ER^{**} = \int_0^{p_b} [ps + (1 - p)t]f(p)dp + \int_{p_b}^1 [pr + (1 - p)u]f(p)dp \qquad (11)$$

when f(p) is continuous.

The quantity ER** in (10) or (11) is the farmer's "preposterior" expected return (i.e., the expected value, as calculated before receiving the

information from the meteorologist, of the expected return after receipt of the information). The expected value of climate forecasts (denoted by V) provided by the meteorologist is simply the difference between the posterior and prior expected returns:

$$V = ER^{**} - ER^{*}. \qquad (12)$$

We will denote the value of perfect information by V*, where ER** (Perfect Information) = (r + t)/2 [see Eq. (10) and the definition of perfect information]. As indicated in Section 2, the value of climate information cannot exceed the value of perfect information (i.e., $V \leq V^{*}$).

3.3 Value of Forecasts: A Symmetric Discrete Model

What are the possible values of the climate forecasts provided by the meteorologist? While this information could take a variety of forms, we consider here climate forecasts for growing season precipitation that are similar in format to current long-range forecasts issued by the U.S. National Weather Service. Specifically, suppose that the forecast p only takes on the values 0.5, 0.5 - c, and 0.5 + c. That is, either p will be the same as the initial probability $p_o = 0.5$, p will be less than 0.5 by the amount c (indicating a smaller chance of above average precipitation), or p will be greater than 0.5 by the amount c (indicating a greater chance of above average precipitation). Furthermore, before the forecast is received, the probability of a forecast of 0.5 - c is q and the probability of a forecast of 0.5 + c is also q. Thus, the probability is 1 - 2q for a forecast of 0.5. As a result, the probability distribution of p is discrete and can be expressed in the form

$$f(p) = \begin{cases} q & \text{if } p = 0.5 - c, \\ 1 - 2q & \text{if } p = 0.5, \\ q & \text{if } p = 0.5 + c, \end{cases} \qquad (13)$$

where $0 < q \leq 0.5$ and $0 < c \leq 0.5$ (note that the probabilities sum to one, as required by probability theory). The values of the parameters q and c determine the exact form of this discrete distribution. Two possible distributions of forecasts are presented in Figure 2 for illustrative purposes. In Figure 2(a), q = 0.4 and c = 0.1; whereas in Figure 2(b), q = 0.2 and c = 0.2. In the first distribution, forecasts of p = 0.4 and p = 0.6 would each be given 40% of the time and forecasts of p = 0.5 would be given the remaining 20% of the time. In the second distribution, forecasts of p = 0.3 and p = 0.7 would each be given 20% of the time and forecasts of p = 0.5 would be given 60% of the time. Thus, the second distribution involves larger departures from the initial probability of 0.5 than does the first distribution, but these more extreme values of p are not used very often. The quantities q and c represent two different aspects of the quality of the climate forecasts, with larger values of both parameters indicating forecasts of higher quality (q = c = 0.5 corresponds to perfect forecasts).

When f(p) in (13) represents the distribution of p, the expected value of the climate forecasts,

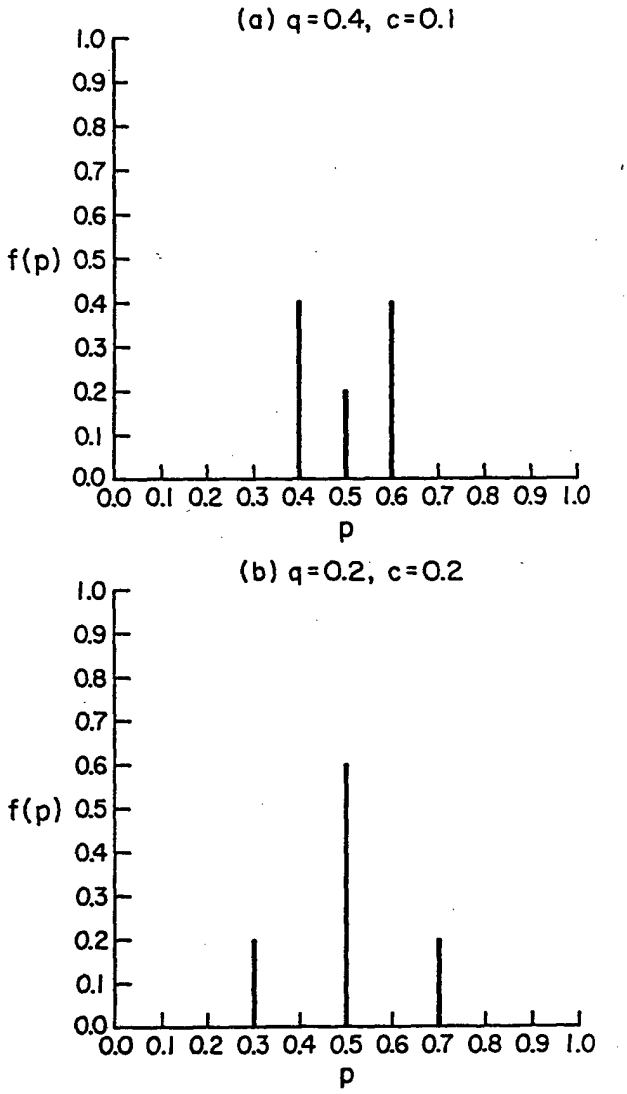


Figure 2. Examples of the form of the distribution of climate forecasts provided by the meteorologist in the case of a symmetric discrete distribution.

as determined from (3), (10), and (12), can be expressed in the following form:

$$V = \begin{cases} q[0.5(s - r + t - u) \\ \quad + c(r - s + t - u)] & \text{if } c > 0.5 - p_b > 0, \\ 0 & \text{if } c \leq |0.5 - p_b|, \\ q[0.5(r - s + u - t) \\ \quad + c(r - s + t - u)] & \text{if } c > p_b - 0.5 > 0. \end{cases} \qquad (14)$$

The expressions for V in (14) indicate that the expected value of climate forecasts is zero if the value of the parameter c is less than the difference between 0.5 and the breakeven point p_b. In this case, the information provided by the meteorologist would always lead to the same course of action that the farmer would have taken had he based his decision only on the initial or prior probability. If c exceeds this difference, then

the expected value of climate forecasts is positive.

It is of interest to note that if we differentiate V in (14) with respect to either q or c, then the derivatives are positive (for $q > 0$, $c > 0$). That is, for $c > |0.5 - p_b|$,

$$\partial V/\partial q = V/q > 0 \qquad (15)$$

and

$$\partial V/\partial c = q(r - s + t - u) > 0. \qquad (16)$$

Thus, as we would expect, the value of forecasts increases as q and c increase (i.e., as the quality of the forecasts increases).

To illustrate these results, we will take some realistic values for the payoffs: r = \$200/acre, s = \$160/acre, t = \$140/acre, and u = \$110/acre. Then p_b in (8) is 0.429 and ER* in (3) is \$155/acre. Moreover, V is zero if $c \leq 0.071$, and V is q(70c - 5) if $c > 0.071$. The value of the forecasts is depicted in Figure 3 as a function of c for selected values of q. Note that the expected value of climate information in this case ranges from zero to \$15/acre, with the latter

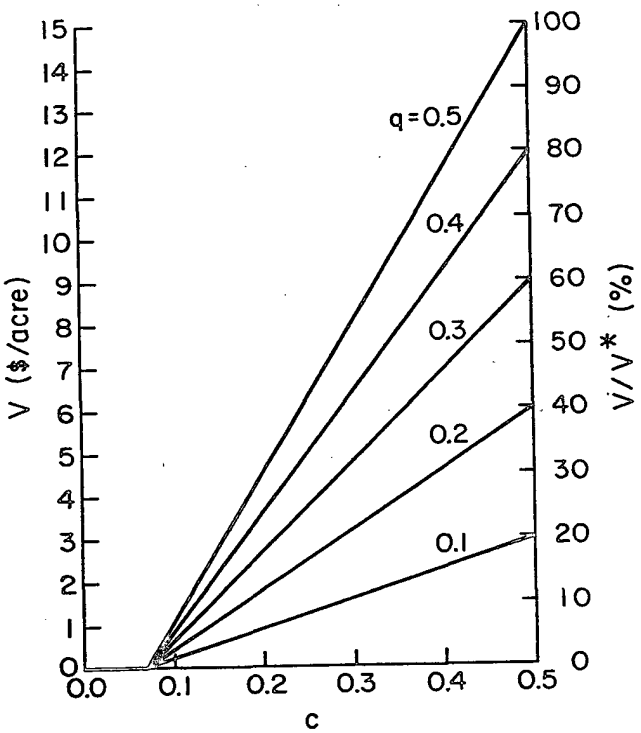


Figure 3. The expected value of climate information V (\$/acre) and the ratio V/V* (in percent) as a function of c for selected values of q for the symmetric discrete distribution.

representing the value of perfect information (V*). To provide a standard of comparison against which to judge the value of climate forecasts, the ratio of V to V* also is indicated in Figure 3. Moreover, the ratio of V to ER*, the prior expected return, can be obtained from Figure 3 by multiplying the values of the ratio of V to V* by 0.0968 (= 15/155). It is of interest to observe that the value of the forecasts is less than 10% of the expected return even for perfect forecasts and that V is less than one-third of V* when q and c are less than or equal to 0.3.

For comparison, we also present the results of an analysis of the situation in which the return t is reduced from \$140/acre to \$130/acre. In this case, $p_b = 0.333$, V is zero if $c \leq 0.167$, and V is q(60c - 10) if $c > 0.167$. The values of V and the ratio V/V* for this situation are depicted in Figure 4 [V/ER* can be obtained from V/V* here by multiplying the latter by 0.0645 (= 10/155)].

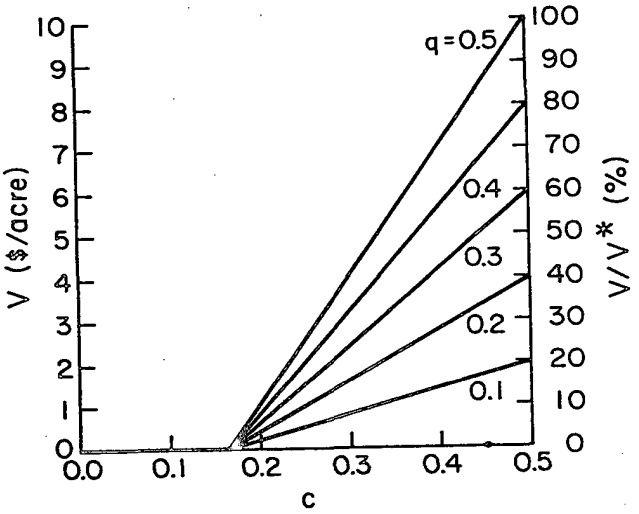


Figure 4. The expected value of climate information V (\$/acre) and the ratio V/V* (in percent) when t = \$130/acre (instead of t = \$140/acre, as in Figure 3) as a function of c for selected values of q for the symmetric discrete distribution.

Note that larger departures from the initial probability of 0.5 are now required for the expected value of climate forecasts to be positive.

4. DISCUSSION AND CONCLUSION

Many decisions are made - and will continue to be made in the future - in decision-making situations in which climate information could (and should) play a vital role. Thus, the value of such information and its appropriate use in these situations are important issues. This paper has presented a decision-analytic approach to the use and value of climate information.

The example presented in Section 3 gives the flavor of the assessment of the value of climate information from a decision-analytic viewpoint. Moreover, the results provide a prescription of the way in which a farmer could use a seasonal climate forecast in an optimal manner to make a

decision regarding which of two crops to plant. This simple example demonstrates that even quite uncertain climate forecasts similar in format to current U.S. National Weather Service long-range forecasts, may be of some value to specific decision makers. The analysis also reveals that climate forecasts may have to attain a minimal level of skill in order to be of positive value in some decision-making situations and suggests that moderate improvements in skill beyond this minimal level can lead to rapid increases in the value of the forecasts.

In order to investigate further the value and use of climate information, a variety of extensions of the analysis presented in this paper could be considered. Such extensions might include modifications of the inputs to the agricultural decision-making problem examined in Section 3, as well as refinements in the basic structure of the problem. With regard to the former, alternative models of the climate forecasts provided to the farmer by the meteorologist could be employed. The results of using two such models - an asymmetric discrete model and a normal (or Gaussian) model - are reported by Winkler et al. (1983). In a related vein, we also could introduce a utility function for (net) monetary payoffs and investigate the effects of this function on the farmer's decisions and on the value-of-information estimates. With regard to refinements in the structure of the problem itself, additional actions (e.g., plant a combination of wheat and barley, plant one or more other crops) and/or additional events (e.g., near normal precipitation, much above and much below normal precipitation) could be included. Of course, decision-analytic studies of other relatively simple problems could be undertaken; such studies might include different decision-making situations in the agricultural sector or climate-information-sensitive problems in other sectors.

Decision analysis also is applicable to much more complex situations (e.g., see Keeney, 1982). Problems with sequences of decisions and events over time, or problems with interrelated long-range and short-range decisions, illustrate the types of complexities that are often found in problems that may be sensitive to climate information. Some recent studies of the value and use of meteorological information have attempted to model such decision-making situations (e.g., see Howe and Cochrane, 1976; Katz et al., 1982). Multi-attribute problems with consequences that cannot all be reduced to a single dimension such as money represent yet another type of complexity. Studies of the value of climate information in the context of both simple and complex decision-making problems should be of interest to a wide spectrum of individuals in the meteorological community.

ACKNOWLEDGMENT

This research was supported in part by the National Science Foundation (Division of Atmospheric Sciences) under grants ATM-8004680 and ATM-8209713 and by the National Oceanic and Atmospheric Administration (Climate Analysis Center) under grant NA82AA-D-0042.

REFERENCES

Agnew, C.E., and R.J. Anderson, 1977: The economic benefits of improved climate forecasting. Princeton, NJ, Mathematica, Inc., Report, 159 pp.

Howe, C.W., and H.C. Cochrane, 1976: A decision model for adjusting to natural hazard events with application to urban snow storms. The Review of Economics and Statistics, 58, 50-58.

Katz, R.W., and A.H. Murphy, 1979: Assessing the value of frost forecasts to orchardists: a decision-analytic approach. Preprints, 14th Conference on Agriculture and Forest Meteorology (Minneapolis). Boston, MA, American Meteorological Society, pp. 34-37.

Katz, R.W., A.H. Murphy, and R.L. Winkler, 1982: Assessing the value of frost forecasts to orchardists: a dynamic decision-making approach. Journal of Applied Meteorology, 21, 518-531.

Keeney, R.L., 1982: Decision analysis: an overview. Operations Research, 30, 803-838.

Murphy, A.H., 1981: Subjective quantification of uncertainty in weather forecasts in the United States. Meteorologische Rundschau, 34, 65-76.

Murphy, A.H., and R.L. Winkler, 1977: Reliability of subjective probability forecasts of precipitation and temperature. Applied Statistics, 26, 41-47.

Phillips, L.D., C.R. Peterson, and T.W. Keelin, 1978: The value of improved forecasts of climate for agricultural decision making. McLean, VA, Decisions and Designs, Inc., Final Report PR 77-7-44, 148 pp.

Winkler, R.L., and A.H. Murphy, 1983: Decision analysis. In Probability, Statistics and Decision Making in the Atmospheric Sciences (A.H. Murphy and R.W. Katz, Editors). Boulder, CO, Westview Press, in preparation.

Winkler, R.L., A.H. Murphy, and R.W. Katz, 1983: The value of climate information: a decision-analytic approach. Journal of Climatology, 3, in press.

ON THE PRESENT AND POTENTIAL USE OF CLIMATE INFORMATION BY THE U. S. PRIVATE AGRICULTURAL SECTOR

Peter J. Lamb, Steven T. Sonka, and Stanley A. Changnon, Jr.

Illinois State Water Survey
Champaign, Illinois 61820

1. INTRODUCTION

One of the more striking science policy developments of the past decade has been the formulation of large, ambitious, multifaceted "climate programs" at both the national and international levels. The U. S. National Climate Program (USNCP), for instance, was established by an Act of Congress (September 1978, PL 95-367) to "assist the Nation and the world to understand and respond to natural and man-induced climate processes and their implications" (Section 3), following legislative deliberations during 1975-78. It has three components (NOAA, 1980). A *Climate Impact Assessment* effort is seeking to identify "procedures to evaluate climate's effects on society, the economy, and the environment in order to develop responses and strategies for dealing with climate fluctuations" (NOAA, 1980, p. E-4). *Climate System Research* will attempt to increase the knowledge of global and regional climate and its variation by means of a range of empirical studies and analyses of the climate record, the development of climate simulation and prediction models, and the investigation of climate system processes (e.g., solar and terrestrial radiation, ocean heat storage and transport). The *Data, Information, and Services* component is designed to provide accurate and timely data and information products, and be responsive to Government and private sector needs. This USNCP is just one of several emerging national programs that are intended to be consistent with a larger World Climate Programme (WCP). The latter was formally established during 1979 by agreement among the World Meteorological Organization, the U. N. Environment Program, the International Council of Scientific Unions, and the Intergovernmental Oceanographic Commission. It will span the two decades 1980-2000, contains subprograms that parallel the components of the USNCP, and is intended to accelerate progress by serving as a catalyst rather than by providing direct support (WMO, 1979, pp. vii, 709-758). The latter role will involve assisting developing countries to build modern data acquisition and application systems, fostering international cooperation when it is a necessary prerequisite for research progress (e.g., on the CO_2 question and ocean heat storage/transport), and other similar activities.

The relatively sudden emergence of these programs from the situation in the 1950s and 1960s where there was little interest in climate, its vagaries, or their effects, has been in response to both the climate system providing an abundance of striking weather extremes and climatic fluctuations during the last 15 years and the wide publicity given to the adverse socioeconomic effects of those episodes by the ever-increasing capabilities of the news media. In particular, the Sahel drought and consequent famine of the early 1970s forced governmental and scientific communities, on an international level, to recognize that climate does vary on short timescales and that such variations can have disastrous human consequences. This new awareness has been increasingly reinforced as the 1970s and 1980s have progressed by other pronounced climatic fluctuations and their adverse impacts — the 1976 heatwave, drought, and water shortages in Western Europe; pronounced extremes in Indian monsoon rainfall and their associated flooding and famine; four recent very severe U. S. winters, including one whose excessive snowfall crippled the Chicago transportation system for many weeks; recurrent poor growing seasons in the Soviet Union; the 1980 central U. S. heatwave and drought that greatly reduced crop production; and so on. For the U. S., the appreciation of climate's central role in human affairs was hastened by the serious economic repercussions of the 1972-73 and 1975 grain sales to the Soviet Union that were occasioned by that nation's aforementioned climate-induced crop failures (NOAA, 1980, p. E1).

It is thus very clear that developing climate programs such as the USNCP and the WCP have been conceived and designed to broadly benefit mankind by reducing the adverse socioeconomic consequences of climatic variability, rather than to foster narrow basic research. This is why they are dominated by the likes of climate impact assessment, data acquisition and applications, and the provision of information and services; the climate system research they do include would very probably have been supported and pursued, for reasons of scientific curiosity, without the creation of the programs. In fact, it seems fair to assert that the existence of these elaborate, ambitious, and expensive programs is based on the assumption that a substantial reduction of the unfavorable effects of climatic variability is attainable. Whether this is the case has never really been demonstrated, as is clearly acknowledged by the twin program objectives for the USNCP Climate Impact Assessment effort that were quoted above. There is no doubt that the second of these goals (the development of "responses and strategies for dealing with climatic fluctuations") will be much more difficult to achieve than its necessary forerunner (the identification of "procedures to evaluate climate's effects on society, the economy, and the environment").

Our contention is thus that the management strategies necessary to substantially reduce the unfavorable socioeconomic consequences of climatic variation are presently largely unknown. If this assessment is correct, the aforementioned climate programs constitute something of a risk. However, we believe that this risk is justified, and indeed that the emergence of the climate programs from the public interest generated by the recent climatic vagaries and their impacts is a desirable development that offers considerable potential and challenge for both atmospheric and social scientists. But the other side of the coin is that an inadequate response to this situation will be to the considerable detriment of the atmospheric sciences' reputation among the wider scientific and governmental communities for its ability to "deliver". The fact that early optimism relating to weather modification and numerical weather prediction has so far not been anywhere near matched by actual achievement underlines the atmospheric sciences' need for the climate programs to be at least modestly successful. In contrast, since social scientists did not initiate the climate programs, they presumably have little to lose (and much to gain) from their participation.

2. THE U. S. PRIVATE AGRICULTURAL SECTOR: BACKGROUND AND PRESENT INVESTIGATION

One of the better starting points for the investigation of the extent to which the adverse socioeconomic consequences of climatic variability could be reduced is the consideration of the world's most productive agricultural system — that of the U. S., and particularly its midwestern heart. This overwhelmingly private enterprise endeavor (save some Federal price-support and acreage-restriction programs) is endowed with highly fertile soil, generally abundant moisture, the finest available scientific and technological support in the fields of plant breeding, chemical development, pest management, and machinery design, and educated operators who function within the motivating (or perish!) environment of the "farm firm". If the present level of use of climate data and information by this highly developed sector could be clearly established, and the benefits ascertained, this knowledge would provide incentives and guidelines for the adoption or increased utilization of such practices by less developed agricultural systems. The latter is a goal of the WCP. In addition, the aforementioned attributes of the U. S. private agricultural sector suggest that it may possess the considerable structural and human flexibility that is necessary to provide an agricultural demonstration of the ultimate potential for improved management strategies to reduce the unfavorable socioeconomic effects of climatic variation.

It is therefore very surprising that this sector has been rather neglected in the development of the USNCP. For example, it was not represented at either of the following meetings that were an important part of this process: (1) the April 1980 "Workshop on the Methodology of Economic Impact Analysis for Climatic Changes" that was sponsored by Resources for the Future (RFF) and the National Climate Program Office (NCPO) and had 43 participants, and (2) the June 1981 "Climate Users' Conference" of the Climate Analysis Center (CAC) that was part of the USNCP (NOAA, 1980, p. 25) and had 50 attendees. The agricultural perspectives/positions/interests at these gatherings were instead taken care of solely by government (Federal and State) and academic economists and scientists, i.e., by people with no practical involvement in the U. S. agricultural system. This constitutes an undesirable situation.

The research project reported here was conceived as an attempt to commence the very necessary redressing of the above deficiency. It has four objectives: (1) to determine the present level of use of climate data and information by the U. S. private agricultural sector, (2) to identify the potential for a fuller use of such material in the future, (3) to assess the uncertainties involved, and (4) to use the results of (1)-(3) to specify the research needed before the level of present use can be increased to the maximum that would seem to be possible. The ultimate goal is thus to develop a prioritized research strategy that would make this important sector less vulnerable to climatic fluctuations. The following components of the private agricultural sector are being considered — rural insurance, integrated pest management, agricultural chemicals, food canning, seed production, farming and farm management, agricultural lending, the brokerage industry, and the grain trade.

Our data have been acquired in two distinct phases. In the first (Spring, 1982), responses to a detailed 10 page questionnaire survey were obtained by mail from 107 people involved in private agricultural concerns (generally 10-15 from each of the above categories) with locations as disparate as San Francisco (CA), Minneapolis (MN), St. Louis (MO), Greensboro (NC), and Midland (TX). The questionnaire focused strongly on the present use of climate data and information, with historical data, "year-to-date" accumulations, "now only" values, and climate prediction being treated separately. The assessment of the potential for a fuller utilization of such material in the future, along with the identification of the research needed before this potential can be realized, took place at a two-day Workshop held in Door County (WI) during August 1982. The primary Workshop participants were 14 of the aforementioned questionnaire survey respondents (1-2 from each category); they were chosen from the 107 original respondents because the insight apparent in their replies suggested they were the best equipped to anticipate possible future developments. Several major agribusiness corporations were represented — Monsanto, Pioneer Seed, Control Data, E. F. Hutton, Continental Grain, Crop-Hail Insurance, and Stokely-Van Camp. Other Workshop attendees included four representatives from Federal agencies (USDA, NSF, NCPO), and one from the Illinois Agricultural Association (i.e., Farm Bureau).

The remainder of this paper provides an *early and rather abbreviated summary* of some of the apparently major findings of the foregoing research effort. We wish to stress that our data analysis had not been completed by the time of writing (December 1982), and that the following results should therefore be regarded as tentative. More detailed, quantitative, supported, and analytical treatments will become available in due course.

3. PRESENT USES OF CLIMATE DATA AND INFORMATION

Several major and somewhat overlapping categories of application of climate data and information (CDAI) have emerged.

One is in the design and planning of ongoing and future operations. Here CDAI are used in the scheduling of field efforts (e.g., tillage, fertilizer application, spraying, planting, etc) by farmers, seed producers, chemical companies, and food processing organizations, and also in the financial decisions made by many of the above activities and by loan companies. In the cases of the seed and food processing firms, the planning includes the selection of sites for contract production, while for the chemical companies CDAI are utilized to assess the likely behavior of their products in the actual environment.

A second type of use of CDAI is as input to the crop yield models that are run routinely during the growing season by grain and brokerage companies and their consultants. These models cover a wide range, and include those of the physiological/daily variety as well as the more traditional statistical/monthly ones. However, it appears that the meteorological input utilized by the former type for the growing period up to the time of the model run is interpolated to a much finer spatial resolution (e.g., down to the county-scale) than characterizes the actual data to which the modelers have access. This activity also makes some assumption about the climatic character of the rest of the growing season beyond the time of a given model run. The alternatives presently in use include the standard 30-year normal, some shorter-period normal (e.g., Lamb and Changnon, 1981), conditional probability predictions derived from the historical records (e.g., there is X% chance August will be Y because July was Z), and the more physically-based monthly and seasonal outlooks of the National Weather Service.

CDAI are also extensively used for the monitoring of in-season conditions by all of the components of the private agricultural sector considered. This monitoring permits the timely and productive adjustments to operating practices that are needed because of prior climatic developments, and also leads to revised estimations of both the procedures that should be used during the rest of the season and their likely outcomes (including yields). Particularly prominent in this regard are decisions relating to seed and chemical selection, integrated pest management, and herbicide application. Historical climatic data also yield a range of probability estimates (e.g., fall and spring frost dates, planting dates, yields) that are frequently used as background-type information.

As the foregoing discussion implies, the present application of CDAI within the private agricultural sector involves a relatively wide range of meteorological parameters. In the case of temperature, the use includes the entire historical data bank on both seasonal and monthly time-scales, daily values for the present season, temporal integrations of interpretive quantities derived from these daily data (e.g., growing degree days, heat units), and information on runs of daily extremes. Precipitation data are also utilized in broadly similar forms. Information on cloud amount/sunshine duration/solar irradiance is considered very useful for photosynthetic and soil moisture considerations, especially when extensive cloudiness persists during important crop growth periods. The latter conditions characterized the middle third of the 1982 growing season in the upper Midwest, and caused plant development there to lag considerably behind the stage that was implied by the accumulated growing degree days. However, the much needed cloud/sunshine/radiation data are not readily available (see Section 4). The other parameters for which CDAI are presently being used include wind (relevant to insect pest problems), soil temperature (planting), soil moisture (crop maturation and Nitrogen application), and frost occurrence (variety selection and overall scheduling).

The most surprising of the foregoing results was the very high level of present use of CDAI by this important economic sector, and especially its grain and brokerage companies and the strongly "controlled" seed and food canning industries. This particularly applies to the "in season" utilization of "now only" (e.g., present soil moisture content) and "year/season to date" (e.g., accumulated growing degree days) information, which evidently constitutes a large part of the total use. The latter is probably little recognized by the atmospheric science community. Furthermore, it appears that there has been a rapid growth of this type of use (and also others) in recent years. The probable contributing factors include the increased financial pressures on agribusiness, a perception that such use provides a company with an economic advantage over its competitors, the dramatic improvement in the sector's modeling and data/information management capabilities that has resulted from the greatly enhanced computer technology, and the financial consequences of the aforementioned 1972-73 and 1975 grain sales to the Soviet Union.

4. MAJOR IMPEDIMENTS TO A FULLER USE OF CLIMATE DATA AND INFORMATION

There is a wide range of reasons why CDAI are presently not used more extensively by the private agricultural sector.

In some cases the non-use is due to the lack of an appropriate delivery system for data that exists, is known to exist, and is desired. This particularly applies to the "now only"/"season to date" type information for which Section 3 noted a substantial need. In the absence of such material, the agribusiness community is forced to use field reports of climatic indicators, various other estimates, experience, and instinctive reactions. A second major impediment to a fuller use of CDAI is the perceived complexity of the problem of which climate is but one part. There is wide recognition that the complex decision making and modeling processes characteristic of this sector have other equally or more important inputs (e.g., economic, social, and political considerations) that are not easily quantified or whose dimensions are imperfectly known. In the face of this situation, there has been a tendency to see little dividend in the sophisticated use of climate data. This may be a product of the lack of understanding of the management techniques needed to ameliorate adverse climatic impacts, and the fact that the benefits to be obtained have not been adequately demonstrated.

There is also deliberate non-use of CDAI known to be available. Such material is either perceived to be of little use, or else its utility is thought to have not yet been demonstrated. In this regard, the monthly and seasonal climate predictions available at present from the National Weather Service and various consultants are not seen to be credible by a large majority of the people in this sector. The zero lead time, coarse spatial resolution, and open distribution (which gives no individual or company an "edge" over competitors) of the Weather Service's predictions further militates against their use.

The capability to fully exploit the existing CDAI is sometimes deficient. Such limitations can be either conceptual (the appropriate models do not exist or are thought not to exist) or physical (lack of the requisite organizational support, computational facilities, trained staff, financial resources). However, the growing trend towards the provision of electronically generated and transmitted information by some large organizations (e.g., grain and brokerage companies, and Farm Bureaus) should help overcome this impediment.

Other reasons for the present use of CDAI not equalling the maximum possible include simple unawareness of the material that is available, the nonexistence or inaccessibility of some highly desirable data (e.g., cloud/sunshine/radiation, soil temperature and moisture, wind, and growing degree days, as was discussed in Section 3), communication problems between scientists and lay users (e.g., the question of what probability predictions mean), the inappropriate formats of some of the present publications and data tapes, and the notion that the cost incurred in acquiring and processing data was not worth the resulting benefits (real or perceived).

5. CLIMATE PREDICTION NEEDS FOR THE FUTURE

One of the Workshop sessions focused exclusively on this topic because the achievement of a really *substantial* reduction of the adverse socioeconomic consequences of climatic variation would seem to require an effective use of skillful climate predictions. The potential benefits to be derived from a fuller utilization of other forms of CDAI are by comparison inherently more modest. The participants were charged with discussing climate prediction needs with respect to the following parameters — applications, lead times, desired length of forecast periods, weather elements needed, resolution (e.g., "above normal" or something more precise), accuracy, and skill. In general, they found this a rather difficult assignment, primarily because they had never before given the subject serious and rigorous consideration. This supports Lamb's (1981) contention that considerable research is needed to assess whether, where, how, and what type of climate predictions cloud/should be used. The full dimensions of this problem are outlined in that publication.

However, three important (if rather general) conclusions did emerge from the Workshop discussions. The first was that, for many agribusiness applications, a prediction of the likely general character of the late spring and summer conditions would be useful if it was made available during the preceding January-March period (certainly no later than April 10). For example, the forecasting of the late May and June climate with this lead time could potentially influence winter decisions on fertilizer use, pesticide and herbicide choices, and production/sales questions. The important meteorological parameters appear to be temperature, sunshine, and rainfall. Since the early-July through mid-August period is the most critical one for crop growth, a demonstrated capability to successfully anticipate its climatic character six months ahead would affect *all* decisions made during the intervening time. A particularly important issue in this regard is the likelihood of a July-August climatic extreme (e.g., the 1980 midwestern heatwave and drought).

The second important conclusion to emerge from this Workshop session was that the private agricultural sector would welcome attempts to predict September and early October conditions with some lead time. It is particularly interested in the likelihood of the early frost that would damage crops, and also the extended wet and cool period that delays harvesting and thereby exposes the crop to a range of yield-reducing threats. It appears that "first frost" predictions would be needed by August 15 to influence late-season decisions; much longer lead times would be required before forecasts could be incorporated into the variety selection process that determines maturation time. Atmospheric scientists have probably underestimated this interest in the predictability of September-October conditions.

Finally, it appears that climate predictions may have to be highly skillful — like "correct" 70-80% of the time — before they are taken seriously by this sector. However, this rather negative attitude could well be the product of the lack of prior deep and rigorous consideration of the subject that was mentioned above. The Workshop participants readily agreed that considerable research and user education will be necessary before an individual company or farmer can properly assess the risks involved in using climate predictions, including those expressed in probabilistic terms. This sentiment is also consistent with the earlier ideas of Lamb (1981).

6. ACKNOWLEDGMENTS

This research is being supported by NSF Grant ATM 81-16615, the USDA, the Crop-Hail Insurance Actuarial Association, Growmark, the Country Companies, and the State of Illinois. We thank Steve Hilberg and Phyllis Stone for logistical support, and Phil Garcia and Wayne Wendland for rapporteuring the Workshop discussions.

7. REFERENCES

Lamb, P. J., 1981: Do we know what we should be trying to forecast — climatically? Bull. Am. Meteorol. Soc., 62, 1000-1001.

Lamb, P. J., and S. A. Changnon, Jr., 1980: On the 'best' temperature and precipitation normals: the Illinois situation. J. Appl. Meteorol., 20, 1383-1390.

NOAA, 1980: National Climate Program: Five-Year Plan. National Oceanic and Atmospheric Administration, Washington, D.C., 101 pp.

WMO, 1979: Proceedings of the World Climate Conference. World Meteorological Organization Publication No. 537, Geneva, 791 pp.

EXTREME VALUES AND METEOROLOGY

J. Tiago de Oliveira

Academy of Sciences of Lisbon
Department of Statistics
Faculty of Sciences of Lisbon
Center for Statistics and Applications (I.N.I.C.)

INTRODUCTION

The purpose of statistical extremes theory is to describe the random behaviour of the m-th largest or smallest values of a sample under very general conditions; subsequently, those results can be used to describe concrete situations (floods, droughts, gusts of wind, large claims in insurance, breakdowns of capacitors, failures, extreme meteorological data, largest waves, etc.) and, in consequence, for decision and design problems, especially when risks of (random) failure or disaster are to be taken in consideration.

Those concrete situations, to be considered in a full-fledged way, need the study of extremes of stochastic processes and, particularly, stochastic processes of extremes. But, a first, and pratically only, step is the study of the asymptotic behaviour of extremes of univariate and multivariate samples because in general, the underlying distribution functions of the sample (vector) variables are completely unknown or, if known, depend on unknown parameters. As usual we will begin by the case of univariate independent samples (random variables independent and identically distributed), dealing previously with the question of exceedances.

After, we will deal with the problems of statistical decision for (eventually small) samples of extremes.

Finally, we will consider briefly multivariate extremes and extremes and stochastic processes.

The basic reference is, evidently, Gumbel (1958); an useful extension is Gumbel (1962).

Although with some precursors, Dodd (1923) seems to be the first one to study sistematically the behaviour of maximum and minimum values of samples of independent and identically distributed random variables.

Fréchet (1927) and Fisher and Tippett (1928) were the first ones to deal with the asymptotic behaviour of extremes using the stability equation; Gumbel (1935) and von Mises (1936) dealt with the same question but from a different perspective (study of the mortality force or hazard rate).

The (almost) complete theory was achieved with the paper by Gnedenko (1943).

In what regards meteorological and oceanographic applications we can cite Jenkinson (1955, 1969), Thom (1968, 1971), Alexander (1970) and Sneyers (1977, 1979).

THE QUESTION OF EXCEEDANCES

Let $(X_1,\ldots X_n)$ be a sample of independent and identically distributed observable (random) variables with distribution function $F(x)$. Let $X'_1 \leq X'_2 \leq \ldots \leq X'_k \leq \ldots \leq X'_n$ be the ordered sample, X'_k being called the k-th order statistics; X'_1 is the minimum or smallest value, X'_n is the maximum or largest value, X'_k is the k-th smallest value (counted from below) or the (n-k+1)-th largest value (counted from above) of the ordered sample.

As it is well known, the density of X'_k is given by $(f(x)=F'(x))$ $\dfrac{n!}{(k-1)!\,(n-k)!}$ $F(x)^{k-1}(1-F(x))^{n-k}\,f(x)$.

Thus the probability of obtaining, in a second sample of size m, p observations smaller or equal than $X'_k(=x)$, is the binomial

$$\binom{m}{p}F(x)^p\,(1-F(x))^{m-p}.$$

So the (advance) probability of obtaining in a second sample p observations smaller than X'_k (whatever may be its value) is

$$\int_{-\infty}^{+\infty}\frac{n!}{(k-1)!\,(n-k)!}\,F(x)^{k-1}\,(1-F(x))^{n-k}f(x)\binom{m}{p}F(x)^p\,(1-F(x))^{m-p}dx=\frac{k\binom{n}{k}\binom{m}{p}}{(k+p)\binom{n+m}{k+p}}.$$

The number of exceedances is, then, a random variable whose mean value is $km/(n+1)$ and variance $mk(n-k+1)\,(m+n+1)/(n+1)^2\,(n+2)$.

Symmetrically, the probability that the k-th largest value of a first sample of n is exceeded exactly p times in a second sample of m is given by the same expression.

Contrary to standard ideas it should be noted that the variance is largest when $k = \frac{n+1}{2}$ (the median) and smallest for k=1 or k=n (extremes); roughly, the variance of the exceedances at the extremes is about 4/n times the variance of the exceedances at the median; this shows that in two sequenced samples, the median (a central value) is less reliable than the extremes - maximum or minimum - in what regards overpassing or downpassing, in the second sample, of observed values.

THE ASYMPTOTIC DISTRIBUTION OF EXTREMES

For the sample $(X_1, \ldots, X_n)$ with random observations (independent and identically distributed), with distribution function F(x), the distribution function of the k-th maximum is

Prob $\{ X'_{n-k+1} \leq x \}$ = Prob [at least (n-k-1) of the $X_j \leq x$] =

$$\sum_{o}^{k-1} \binom{n}{j} F(x)^{n-j} (1-F(x))^{j} .$$

Thus the distribution function of the maximum (k=n) is $F^n(x)$ and of the minimum (k=n) is $\sum_{o}^{n-1} \binom{n}{j} F(x)^{n-j} (1-F(x'))^{j} = 1-(1-F(x))^{n}$.

We can show that for a large family of distribution functions exist attraction coefficients λ_n and $\delta_n (>o)$ such that

$$F^n (\lambda_n + \delta_n z) \to L(z);$$

L(z) must be of one of the three reduced forms, i.e., avoiding location and dispersion parameters:

$\Phi_\alpha(z) = o$ if $z < o$: Fréchet distribution

$= \exp(z^{-\alpha})$ if $z \geq o$, $\alpha > o$;

$\Lambda(z) = \exp(-e^{-z})$, $-\infty < z < +\infty$: Gumbel distribution ;

$\Psi_\alpha(z) = \exp(-(-z)^{\alpha})$ if $z < o$: Weibull distribution

$= 1$ if $z \geq o$, $\alpha \geq o$.

The relation $F^n(\lambda_n + \delta_n z) \to L(z)$ $(o < L(z) < 1)$ is equivalent to $n(1-F(\lambda_n + \delta_n z)) \to -\ln L(z)$ which is easier to manipulate.

In the applications, as the convergence is uniform and there exist constants such that $L^m(z) = L(\alpha_m z + \beta_m)$ (stability), we can approximate $F^n(x)$ by $L\left(\frac{x-\lambda}{\delta}\right)$ with convenient location (λ) and dispersion (δ) parameters; evidently, in the two last forms, we have yet a shape parameter α.

The graphs of the densities $\Lambda'(z)$, $\Phi'_\alpha(z)$ and $\Psi'_\alpha(z)$ are given in figs. 1, 2 and 3; remark that for $\alpha > o$ the graphs of $\Phi'_\alpha(z)$ and $\Psi'_\alpha(z)$ are similar to $\Lambda'(z)$, apart from location and dispersion parameters. See figs. 1, 2 and 3.

The asymptotic reduced distribution of the k-th maxima, if the one of the maximum is L(z), is given by

$$L_k(z) = L(z) \sum_{o}^{k-1} \frac{(-\log L(z))^{j}}{j!}$$

i.e., $\Phi_\alpha(z) \sum_{o}^{k-1} \frac{z^{-j\alpha}}{j!}$ $(z > o)$, $\Lambda(z) \sum_{o}^{k-1} \frac{e^{-jz}}{j!}$

and $\Psi_\alpha(z) \sum_{o}^{k-1} \frac{(z)^{j\alpha}}{j!}$ $(z < o)$ for each form.

The approximation, for practical uses, is evidently $L_k\left(\frac{x-\lambda}{\delta}\right)$.

It can be shown, directly or using the fact that $\min(X_1, \ldots, X_n) = \max(-X_1, \ldots, -X_n)$, that the corresponding reduced distribution functions for the k-th minima are $1-L_k(-z)$. It should be noted that if the maximum and the minimum have asymptotic distributions they are not necessarily L(z) and 1-L(-z); for instance, for the exponential distribution, the asymptotic reduced distribution of the maximum is the Gumbel one ($\Lambda(z)$) although the asymptotic reduced distribution of the minimum is exactly the exponential distribution $1-\Psi_1(-z)$.

Finally, let us remark that if λ_n and δ_n are associated with the sample size n, the attraction coefficients for the sample size N, in general, verify the relations $\lambda_N = \lambda_n + \delta_n \log \frac{N}{n}$, $\delta_N = \delta_n$ in the case of Gumbel distribution, $\lambda_N = \lambda_n = o$, $\delta_N = \delta_n \left(\frac{N}{n}\right)^{1/\alpha}$ in the case of Fréchet distribution and $\lambda_N = \lambda_n = x_o$, $\delta_N = \delta_n (n/N)^{1/\alpha}$ (where x_o denotes the right end point, $F(x_o)=1$, $F(x) < 1$ if $x < x_o$ in the case of Weibull distribution. Those results are very important for

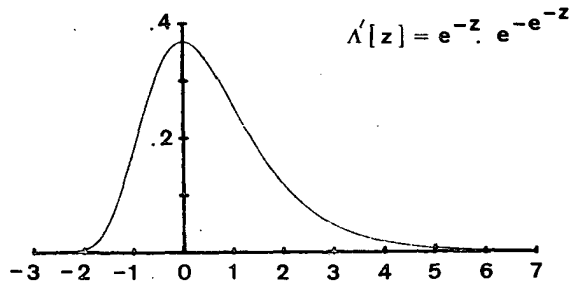


Fig.1 : Gumbel distribution

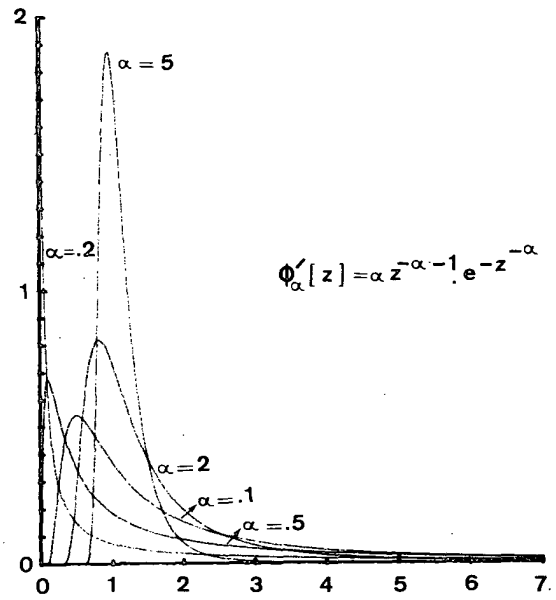


Fig. 2 : Fréchet distribution

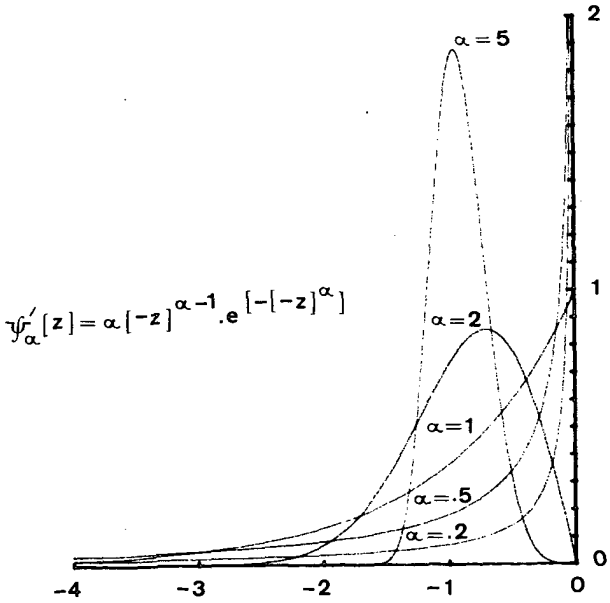


Fig. 3 : Weibull distribution

the pooling of maxima (extremes) of samples of different size. For details see Tiago de Oliveira (1977) and (1977').

It can be shown that the conditions under which the asymptotic distribution were obtained (independence and identical distribution) can be strongly weakened giving the same results. As a consequence, from the practical point of view, we can use the asymptotic distribution Φ_α, Λ and Ψ_α, with location and dispersion parameters as the distributions of the maxima in large samples even when the random observations are not independent (although not strongly dependent) and the marginal distributions are not identical; see Juncosa (1948) for the weakening of "identical distributions" condition, Watson (1954), Newell (1964), Loynes (1965) and others for the weakening of independence condition and Tiago de Oliveira (1979) for the simultaneous weakening of both.

From meteorological and oceanographical point of view the most important result is the one by Watson (1954); in this paper it is shown that if the random sequence $\{X_1, X_2, \ldots, X_n, \ldots\}$ is m-dependent -i.e., observations X_i and X_j such that $|i-j| > m$ are independent (o- dependence is independence) - then, with some mild conditions, have the same asymptotic distributions of maxima (and minima). And this may be applied to meteorological and oceanographic data, as they seem to be pratically independent 5 to 7 days apart.

STATISTICAL DECISION FOR EXTREMES

For simplicity of the exposure of the ideas we will suppose that we are dealing with maxima: as it is the case of floods, largest gusts of wind, maxima temperatures, etc.; minima temperatures, droughts, etc. can be dealt with in the same way (recall the relation $\min(X_i) = -\max(-X_i)$).Analogously can be dealt with the k-th maxima or minima.

The first step is to choose between the three forms and,after do the statistical decision corresponding to the then assumed model.

A first, and quick, approach can be the use of Gumbel probability paper, which in one of the axis (z) is marked in arithmetic scale and in the other (y) the distance y is marked in functional scale p (= F(y)). Then the points (z'_i, p_i), where p_i are the plotting positions, should be close to the 1st diagonal $z=y(=\Lambda^{-1}(p)) = -\log(-\log p_i)$.

As the reduced values $z'_i = (X'_i - \lambda)/\delta$ are not known, because the parameters λ and δ are unknown, we can estimate them by statistics λ^* and δ^* and use estimated reduced values $z'^*_i = (X'_i - \lambda^*)/\delta^*$. Let $Q_n(p)$ be the empirical quantile for probability p, i.e. $Q_n(p) = X'_{[np]+1}$. It is easy to show that for $p_1 = \Lambda(0)$

$= 1/e = 0.367...$ we have $Q_n(p_i) \xrightarrow{p} \lambda$ and so $\lambda^* = Q_n(0.367)$ is an estimator of λ ; analogously for $p_2 = \Lambda(1) = \exp(-e^{-1}) = 0.692...$ We see that $Q_n(p_2) \xrightarrow{p} \lambda + \delta$ so that $\delta^* = Q_n(0.692) - Q_n(0.367)$ is an estimator of δ ; other estimators could be suggested but this choice is simple and straightforward. Thus we will take

$$z'^*_i = \frac{X'_i - Q_n(0.367)}{Q_n(0.692) - Q_n(0.367)}$$ and - let us recall - the plotted points (z'_i, p_i) must be close to the lst. diagonal. For p_i, as a lst. approximation, we can take $p_i = \frac{i - 0.307}{n + 0.386}$ or more simply $p_i = \frac{i - 0.3}{n + 0.4}$.

The use of e^{-1} and $\exp(-e^{-1})$ quantiles give estimators of efficiency of about 15%; the maximum efficiency for estimation using quantiles is about 26%, using the lst. and 3rd. quantile. It seems not, in this stage of analysis, very helpful to use the quartiles; on the other side we get, this way, an elementary and quick estimator of λ and δ.

Consider now the alternative forms Φ_α and Ψ_α . The data, if Φ_α is true shall be close to $y = \alpha \log(1 + (e^{1/\alpha} - 1)z)$ and if Ψ_α is true shall be close to $y = -\alpha \log(1 - e^{-1/\alpha})z)$

Remark that both curves converge to the lst. diagonal when $\alpha \to +\infty$ which explains the remark that for $\alpha > 5$ the densities Φ'_α and Ψ'_α are similar to Λ'. Note also that both curves cross the lst. diagonal at $z = o$ and $z = 1$.
For more details see Tiago de Oliveira (1983).
An approach, more usual, is to plot in this paper, the points $(X'_i, \frac{i-0.3}{n+0.4})$ and if then fall roughly in a straignt line we can assume Gumbel distribution. If this is not the case, as $\log(X - \lambda)$ has a Gumbel distribution with $\log \delta$ and $1/\alpha$ as location and dispersion parameters, when X has the Fréchet distribution and $-\log(X - \lambda)$ has also a Gumbel distribution with $-\log \delta$ and $1/\alpha$ as location and dispersion parameters, when X has the Weibull distribution, we shall, then, try, for a grid of values of λ if with $\log(X - \lambda)$ or $-\log(X - \lambda)$, for some λ, the plotted points fall, approximately, in a straight line in Gumbel probability paper. If one of those plots gives a good fit to the eye, we can assume that this one is the approximate distribution to be used. Numerical computations shall then follow.

At the end of this section it is described an analytical approach which is preferable. The analysis of data contained in Sneyers (1977) is given in fig. 4 (probability paper) and table I (for the analytical test).

Details about statistical decision (esti-

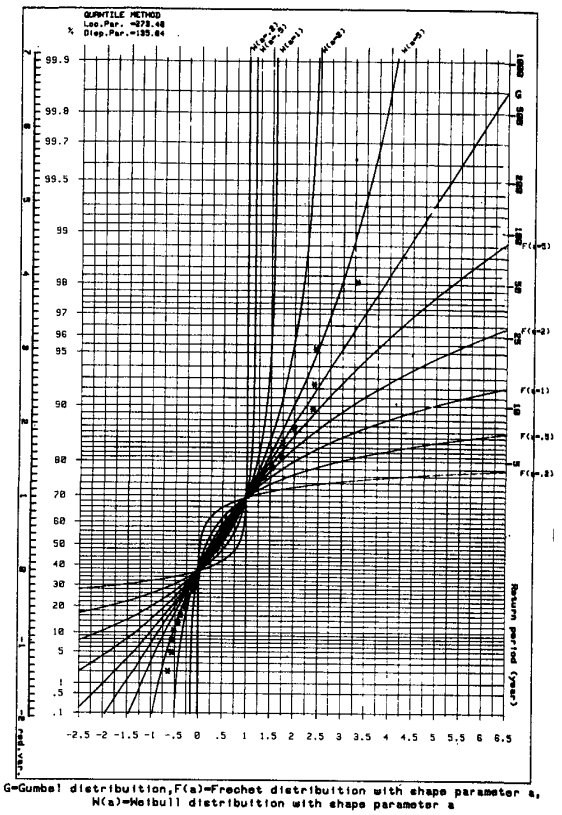


Fig. 4: Annual maxima of rain (Uccle, 24h) — 1938/1972

1938	338
1939	277
1940	600
1941	240
1942	723
1943	507
1944	187
1945	412
1946	266
1947	272
1948	238
1949	198
1950	343
1951	282
1952	511
1953	375
1954	343
1955	222
1956	356
1957	342
1958	243
1959	283
1960	480
1961	324
1962	596
1963	804
1964	270
1965	450
1966	398
1967	216
1968	197
1969	544
1970	291
1971	416
1972	260

$Q(.367) = 273.46$

$Q(.692) = 408.18$

$\lambda^* = 273.46$

$\delta^* = 135.84$

$\hat{\lambda} = 295.75$

$\hat{\delta} = 101.49$

$Vn = 8.42$

$Vn/(2.89797xn)^{1/2} = 0.75$

TABLE I

mation and testing) can be found in Tiago de Oliveira (1975). In the case of Gumbel distribution, sometimes said to be the extreme distribution, the density being, with the location and dispersion parameters, $\Lambda'((x-\lambda)/\delta)/\delta = 1/\delta . \exp(-(x-\lambda)/\delta) . \exp(-\exp(-(x-\lambda)/\delta))$

the likelihood has the expression

$$1/\delta^n . \exp(-\sum_1^n (X_i-\lambda)/\delta) . \exp(-\sum_1^n \exp(-(X_i-\lambda)/\delta))$$

so that the maximum likelihood equations lead to

$$\hat{\delta} = \bar{x} - \frac{\sum_1^n X_i e^{-X_i/\hat{\delta}}}{\sum_1^n e^{-X_i/\hat{\delta}}} \qquad \hat{\lambda} = -\hat{\delta} \log(\sum_1^n e^{-X_i/\hat{\delta}}/n).$$

The second equation is easily solved, in computers, by the Newton-Raphson method, with few iterations; once $\hat{\delta}$ is known, $\hat{\lambda}$ is immediately computed. $\hat{\lambda}$ and $\hat{\delta}$ are asymptotically normal, with mean values λ and δ, variances $(1+\frac{6(1-\gamma)^2}{\pi^2})\frac{\delta^2}{n}$ and $\frac{6}{\pi^2}\frac{\delta^2}{n}$ covariance $\frac{6(1-\gamma)}{\pi^2}\frac{\delta^2}{n}$ and correlation coefficient $\rho = (1+\frac{\pi^2}{6(1-\delta)^2})^{-1/2} = 0.313$

Recall that the mean value of a Gumbel random variable is $\lambda+\gamma\delta$ and the variance is $\pi^2/6.\delta^2$.

Prediction procedures have, also, been dealt with in Tiago de Oliveira (1972) and (1975).

The last paper describes procedures to obtain estimators of the parameters for the three forms. Those estimators have, in general, asymptotically normal behaviour, except in the case $0 < \alpha < 2$ for Weibull distribution, a case which is very important in rupture problems.

A good coordination of statistical decision methods for Gumbel and Weibull (minima) distributions can be found in Johnson and Kotz (1970) (ch. 21 and 20).

The most important quantities to be estimated are the quantiles and the probabilities of overpassing a level a (bound). If the asymptotic distribution is L(z) the p quantile is the (unique) solution of the equation $L\left(\frac{x-\lambda}{\delta}\right) = p$ and has, then, the probability p of not being exceeded and 1-p of being exceeded; in practice, if we are dealing with overpassings we take p close to 1 and if we are dealing with downpassings we choose p close to 0.

The quantiles are, then, $\lambda - \log(-\log p)\delta$ for Gumbel distribution, $\lambda + (-\log p)^{-1/\alpha}\delta$ for Fréchet distributions and $\lambda - (-\log p)^{1/\alpha}\delta$ for Weibull distributions. They are estimated either by order statistics either substituting the parameters by their estimators and, as continuous functions of asymptotically normal variables, they are also asymptotically normal. The last way is in general preferable if p is close to 0 or to 1.

The probability of downpassing a level a, $L(\frac{a-\lambda}{\delta})$ or overpassing a level a, $1-L(\frac{a-\lambda}{\delta})$ is also estimated substituting the parameters by their estimators. In general, the behaviour of this (estimated) probability is also asymptotically normal and can also be computed in the usual way - for details see Fraser (1957) and Tiago de Oliveira (1982').

Connected with those probabilities is the notion of return period. The return period of the overpassing of the level a, i.e., the mean value of the time interval between sucessive overpassings of the level a in a sequence of independent experiences is given as $T(a) = [1-L(\frac{a-\lambda}{\delta})]^{-1}$ and is approximated usually by $\tilde{T}(a) = -1/\log L(\frac{a-\lambda}{\delta}) \leq T(a)$. If we use $\tilde{T}(a)$, instead of T (a), for planning and design we are lead to a larger effective return period which increases the safety of a structure. For downpassings we have $T(a) = L^{-1}(\frac{a-l}{\delta})$.

Let us now discuss, briefly, the statistical choice between the three forms of asymptotic distributions, as in Tiago de Oliveira (1980) and (1982). The three forms can be imbeded in von Mises - Jenkinson form $G(z/k) = \exp\{-(1+kz)^{-1/k}\}$ which for $k=0^+$ or 0^- gives Gumbel distribution, for $k > 0$ gives $G(z/k) = \Phi_{1/k}(1+kz)$ and for $k < 0$ gives $G(z/k) = \Psi_{-1/k}(-1-kz)$. If the observed variables are reduced - i.e., without location and dispersion parameters - a trilemma decision ($k < 0$, $k = 0$ or $k > 0$) can be made according to the value of the statistic $V_n = \sum_1^n v(z_i)$ where $v(z) = \frac{\partial \log G'(z/k)}{\partial k} = \frac{z^2}{2} - z - \frac{z^2}{2}e^{-z}$

if $V_n \leq b_n$ decide for Weibull distribution

if $b_n < V_n \leq a_n$ decide for Gumbel distribution

and if $a_n < V_n$ decide for Fréchet distribution.

As the mean value of v (z) is zero and variance exists if $-1 < k < 1/4$ we can compute asymptotically b_n and a_n, imposing $b_n = -a_n$ for symmetry and a level of significance at $k = 0$.

As in general the data are not without location parameters we estimate λ and δ by the maximum likelihood equations for Gumbel distri-

bution (giving thus $\hat{\lambda}$ and $\hat{\delta}$); use the estimated reduced observations $\hat{z}_i=(x_i-\hat{\lambda})/\hat{\delta}$ and compute the estimated V_n as

$$\hat{V}_1=\sum_1^n V(\hat{z}_i)=\frac{1}{2.\hat{\delta}^2}\{\sum_1^n x^2{}_i-2\hat{\delta}\sum_1^n x_i\ln\frac{\sum_1^n x_i e^{-x_i/\hat{\delta}}}{\sum_1^n e^{-x_i/\hat{\delta}}}\}$$

It was shown, also, that $\hat{V}_n$ is asymptotically normal with mean value zero and variance 2.09797 n. Thus if we impose the significance level w we have Prob {decide k = 0/k = 0} = = Prob $\{|\frac{\hat{V}_n}{\sqrt{2.09797\ n}}| \leq c\} = 1-w$ we must take $N(c) - N(-c)=1-w$ or $N(c) = 1-w/2$ where $N(.)$ denotes the standard normal distribution function. See table I.

A different approach, not overlapping with the one of Tiago de Oliveira (1980) is given in Salambos (1980).

Multivariate extremes and extremes in stochastic processes: Only few hints will be given of those topics, but calling attention to the references below.

Let $F(x,y)$ be the distribution function of a random pair (X,Y). For the independent sample $(X_1, Y_1),\ldots, (X_n, Y_n)$ we have immediately

Prob $\{\max(X_i) \leq x, \max(Y_i) \leq y\} = F^n(x,y)$,

remark that $(\max(X_i), \max(Y_i))$ is not an observed point, in general.

The limiting forms (if they exist) of

Prob $\{\frac{\max(X_i) - \lambda_n}{\delta_n} \leq x, \frac{\max(Y_i) - \lambda'_n}{\delta'_n} \leq y\} =$

$= F^n(\lambda_n + \delta_n x, \lambda'_n + \delta'_n y)$, with Gumbel margins must satisfy the functional equation

$$L^z(x,y) = L(x-\ln z, y-\ln z)$$

for z positive. Its solution is

$$L(x,y) = \exp[-(e^{-x}+e^{-y})\,k(y-x)]$$

where k, the dependence function, must satisfy some conditions; see Tiago de Oliveira (1974) for details. Five models exist until now (logistic, mixed, biextremal, Gumbel, natural) more or less studied. But their different applications, if any, are not yet set up.

For extremes in stochastic processes two lines can be considered:

a) extremes in (some) stochastic processes: see Cramer and Leadbetter (1967) for some basic results;

b) stochastic processes of extremes: the basic papers are the ones of Tiago de Oliveira (1968) and (1973) and of Lamperti (1964) and Dwass (1964).

As a general rule, for large excursions in case a) and for any finite time in case b), the distribution of the maximum of the process is of one of the asymptotic forms of the independent and identically distributed samples.

Applications to meteorological, hydrologic and oceanographic data :

Yearly floods and droughts are the essential problems in what regards to hydrologic data, or as it is known, the first application of statistical extremes. The Rhone river (France) seems to be the 1st. river to be discovered to "obbey" Gumbel law for maxima; experience has shown that not only Gumbel distribution but, also Fréchet seem to fit the distributions of floods. Symmetrically, it seems that the same distributions, for minima, can be applied for droughts.

An apparent paradox can be raised: both Gumbel and Fréchet distributions are infinite to the right and no one can expect infinite floods. In fact, for reduced values, we see that

Prob $\{z\geq 5\}=1-\Lambda(5)=7\times10^{-3}$, Prob$\{z\geq 7\}=1-\Lambda(7)=9\times10^{-4}$ so pratically it is impossible to observe, in 100 or 200 years, floods whose reduced value will be larger than 7. In general words, from a practical point of view, very large (reduced) floods have a very, very small probability and impossible from empirical point of view. The rise of unlimited distributions leads thus to no difficulty in applications for the negligeable probability of very large (small) values, not being needed, then, the rise of an upper (lower) bound.

Some other distributions, as the log-normal, have also been applied to the distribution of floods and droughts: this can be explained because distributions of different expressions can give approximately equal evaluations of probability for intervals of central area, but may be very different in probability evaluations for intervals outside central areas, corresponding to very small probability amounts.

In what regards oceanographic data it seems that Fréchet fitted reasonably well maxima wave heights. But data were obtained through visual observations. Some data, connected with few rigorous observations are analysed in this paper; the regression between visual and rigorous observations suggest that previous use of Fréchet distribution for maxima waves can (not) be extended.

For meteorological data (gusts of wind, maxima rainfalls, etc.) it seems that Gumbel distribution is a good fit.

The study of bivariate (multivariate) extremes and extremes in stochastic processes seems

to be, yet, in the beginning in themselves and also in applications in this area. Suffices, already, to say that for floods of the same river, at different points, the bivariate models have given reasonable fits. But we are, yet, in the first steps.

REFERENCES

Alexander, G. N., 1970: Heavy rainfall as a stochastic process. Rev. Int. Stat. Inst., vol. 38.

Cramer, H. and M. R. Leadbetter, 1967: Stationary and related stochastic processes. Wiley Publ.

Dodd, E. L., 1923: The greatest and least variable under general laws of error. Trans. Amer. Math. Soc., vol. 25.

Dwass, M., 1964: Extremal processes. Ann. Math. Stat., vol. 35.

Fisher, R. A. and L. H. C. Tippett, 1928: Limiting forms of the frequency distribution of the largest and smallest member of a sample. Proc. Camb. Phil. Soc., vol. 24.

Fraser, D. A. C., 1957: Non-parametric methods in Statistics. Wiley Publ.

Fréchet, M., 1927: Sur la loi de probabilité de l'écart maximum. Ann. Soc. Polon. Math. (Krakow), vol. 6.

Galambos, J., 1980: A statistical test for extreme value distributions. In Non-parametric methods of Statistics. Budapest.

Gnedenko, B. V., 1943: Sur la distribution limite du terme maximum d'une serie aleatoire. Ann. Math., vol. 44.

Gumbel, E. J., 1935: Les valeurs extrêmes des distributions statistiques. Ann. Inst. H. Poincaré, vol. 4.

Gumbel, E. J., 1958: Statistics of Extremes, Columbia Univ. Press.

Gumbel, E. J., 1962: Statistical theory of extreme values, in Contributions to order statistics, edit. by Saharan and Greenberg. Wiley Publ.

Jenkinson, A. F., 1955: The frequency distribution of the annual maximum (or minimum) values of meteorological elements. Q. J. Roy. Met. Soc., vol. 81.

Jenkinson, A. F., 1969: Estimation of maximum floods, Tech. Note 98, WMO, Geneva.

Johnson, N. L. and Samuel Kotz, 1970: Continuous univariate distributions-1, Houghton and Mifflin Cy.

Juncosa, M. L., 1948: On the distribution of the minimum of a sequence of mutually independent random variables. Duke Math. J., vol. 16.

Lamperti, J., 1964: On extreme order statistics. Ann. Math. Stat., vol. 35.

Loynes, R. M., 1965: Extreme values in uniformly mixing stochastic processes. Ann. Math. Stat. vol. 36.

Newell, G. F., 1964: Asymptotic extremes for m - dependent random variables. Ann. Math. Stat., vol. 35.

Sneyers, R., 1977: L'intensité maximale des precipitations en Belgique. Publ. Ser.B, nº86, Inst. Rog. Met. Belgique.

Sneyers, R., 1979: L'intensité et la durée maximales des precipitations en Belgique. Publ. Ser. nº 99, Inst. Rog. Met. Belgique.

Thom, H. C. S., 1968: Toward a universal climatological extreme wind distribution, Proc. Sem. Extreme wind Distr., Nat. Res. Comc. Proc.

Thom, H. C. S., 1971: Asymptotic extreme-value distributions of wave heights in the open Ocean. J. Mar. Res., vol. 29.

Tiago de Oliveira, J., 1968: Extremal processes; definition and properties. Publ. Inst. Stat. Univ. Paris, vol. XVII.

Tiago de Oliveira, J., 1972: Statistics for Gumbel and Fréchet distributions, in Structural safety and reliability, ed. by Freudenthal, Pergamon Press.

Tiago de Oliveira, J., 1973: An extreme-markovian-stationary process. Proc. 4th. Conf. Probab. Theory, Romenia.

Tiago de Oliveira, J., 1974: Bivariate and multivariate extreme distributions, in Statistical distributions in scientific work, vol. 1, edit. by Patil, Kotz and Ord, D. Reidel Publ. Cy.

Tiago de Oliveira, J., 1975: Statistical decision for extremes. Trab. Inv. Oper. y Estad. (Madrid), vol. XXVI.

Tiago de Oliveira, J., 1977: Statistical methodology for large claims. The Astin Bull., vol. IX.

Tiago de Oliveira, J., 1977': Asymptotic distributions for univariate and bivariate m-th extremes. In Recent Developments in Statistics, ed. Barra et al.

Tiago de Oliveira, J., 1979: Approximate distributions for sequences of maxima. Metron, vol. XXXVI-1978.

Tiago de Oliveira, J., 1980: Statistical choice of univariate extreme models, Statistical distributions in scientific work, vol. 6, ed. C. Taillie et. al., Reidel.

Tiago de Oliveira, J., 1982: Decision and modelling for extremes, Some Recent Advances in Statistics, ed. J. Tiago de Oliveira and B. Epstein, Academic Press.

Tiago de Oliveira, J., 1982': The δ-method for the obtention of asymptotic distributions; applications. Publ. Inst. Stat. Univ. Paris, vol. XXVII.

Tiago de Oliveira, J., 1983: Probability paper and plotting positions - a new approach, to be published.

von Mises, R., 1936: La distribution de la plus grande de n valeurs. Rev. Math., de l'Union Interbalkanique, vol. 1.

Watson, G. S., 1954: Extreme values in samples for m - dependent stationary stochastic processes. Ann. Math. Stat., vol. 25.

STOCHASTIC PROCESS ANALYSIS OF RAINFALL TOTALS AND EXTREMES

K.J.A. Revfeim

Meteorological Service
P.O. Box 722, Wellington
New Zealand

1. INTRODUCTION

Most statistical analyses of rainfall are based on empirical models of the data e.g. Gamma or Gumbel distributions of maxima, Log-normal distribution of totals, and Markov representations of sequences. Thus long-term average rainfalls are compared on the basis of statistical properties of the mean, with respect to an empirical distribution, or of some higher moment such as variance or skewness. The model chosen for the empirical distribution is probably more dependent on whether it was successfully used by previous investigators than any physical appropriateness for the rainfall process. Access to, or familiarity with, statistical software, may also influence the choice.

To derive a model of typical rainfall measures (e.g. monthly totals or 1-day extremes), based on simplistic but not unreasonable physical assumptions, requires less effort than the subjective justification for preferring a particular empirical model. The results of using such a model with physically meaningful parameters are threefold. First, there is likely to be a near-equivalence between the mathematical form of the physically based model and what has evolved as the best fitting empirical model. This will reinforce the basis of the physical model derivation. Secondly the interpretation of estimated parameters of the physical model is likely to agree with an already accepted understanding, or intuitive notion, of what the parameter should be. Thirdly, each qualifying restriction in the physical model can be separately relaxed leading to more sophisticated models. The tendency to introduce more highly parameterised empirical models is not based on such sure foundations.

With the single exception of a recent unresolved problem the large amount of literature concerned with rainfall analyses using empirical models will not be referred to. The aim of this paper is to clarify the assumptions on which the simplest stochastic model of rainfall are based. Brief mention is made of how this model reflects seasonal patterns and its potential for mapping rainfall objectively. Solutions are given to some practical problems of missing data and a 'proper' definition of maximum rainfall for longer intervals.

The physical basis of more realistic models is indicated.

2. ASSUMPTIONS ABOUT OCCURRENCE AND SIZE

A rainfall 'event' can be simply defined as precipitation from a single shower (cell), or cluster of showers, due to some conjunction of atmospheric processes. Im most rainfall regimes the duration of rainfall events is on average much shorter that the time between events. Thus as a first approximation rainfall events can be treated as a point process of no duration. Figure 1 shows a sequence of rainfall events and a point process of equivalent values.

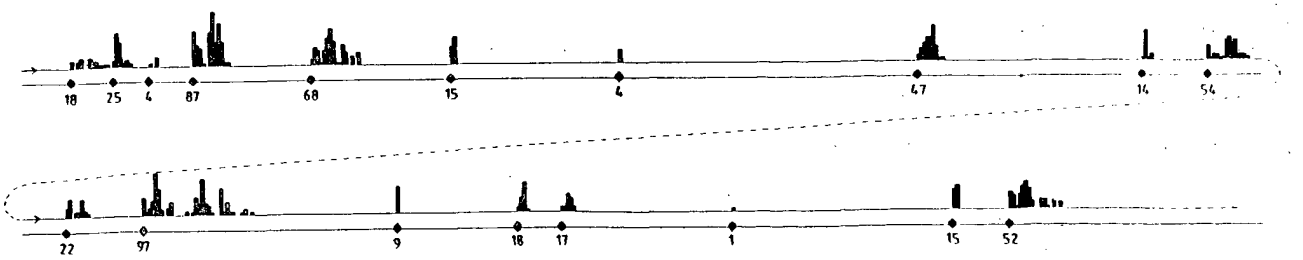


Fig.1. Hourly rainfall time series with an approximating point process representation (size of point events given by numbers underneath).

This approximation is even closer for occurrences underlying maximum rainfalls accumulated in short time intervals, such as 10, 20 or 30 minutes. Several disjoint accumulations may underly the maximum in a single rainfall event. However the largest of these can be taken as the event underlying the maximum.

Obviously if all rainfall events were identifiable then the largest short-interval accumulations might be recorded and a histogram formed. This histogram would be likely to show more than one mode since frontal and convective rainfalls have different short interval intensities. When we are only interested in extremes these will mostly be found in the highest mode. This mode is also likely to show the smallest proportion of the total frequency.

Rainfall totals over periods of months or years are the accumulation of many events. The chance that the duration of a rainfall event overlaps two consecutive months or years will not undermine the point event assumption unless there are on average only one or two events in each period.

Accepting that the rainfall process is satisfactorily approximated by point events it is natural in the first instance to model the occurrence of these events by a Poisson process. This further assumes that the occurrence of rainfall is independent of all past occurrences. The necessity for moist air to transfer by advection or convection into the region where rainfall occurs, following the previous rain event, weakens this assumption. However the random nature of atmospheric processes and analyses such as that of Thom (1959) indicate that the Poisson assumption is reasonable.

Given a Poisson process, of rate ρevents per unit of time, then given that events have occurred, in time T the maxima have density function

$$g_T = e^{-\rho T}\sum_r (\rho T)^r F^{r-1}/(r-1)!(1-e^{-\rho T}) \qquad (1)$$

and the total has density function

$$h_T = e^{-\rho T}\sum_r (\rho T)^r f_r/r!(1-e^{-\rho T}) \qquad (2)$$

where $F = F(x)$ is the probability distribution function of event size (with derivative $f = f(x) = F'(x)$) and f_r is the density function of the total of r events. Obviously frequency distributions of event sizes have the property that there are fewer events as the size increases. The nature of the distribution at the lower end is less certain.

When rainfall events occur the most frequent events will be of some definite size, with the frequency distribution falling back to zero as the size approaches zero. How rapid this tailing off is near the origin would not be easy to determine and may be confounded with the identification of some rainfalls as separate events, or showers within the same event.

In the first instance it is natural to be concerned only about the upper tail of the distribution. This is particularly so for extremes but also for totals since contributions from the lower tail have only a marginal effect. Thus for both maximum rainfalls and total rainfalls the assumption of an underlying negative exponential distribution of sizes is not unreasonable.

Approaches to reality in the time domain include a time varying rate of occurrence of the Poisson process (Zelenhasić, 1970, Revfeim 1982a, b) and for rainfall totals a distribution of rainfall durations which may be either correlated with or independent of event size. More realstic Gamma distributions of event size with a zero left tail, as shown in Figure 2, can also be used. In effect these distributions provide a probabilistic filtering of smaller sizes events which might be compared with threshold censoring.

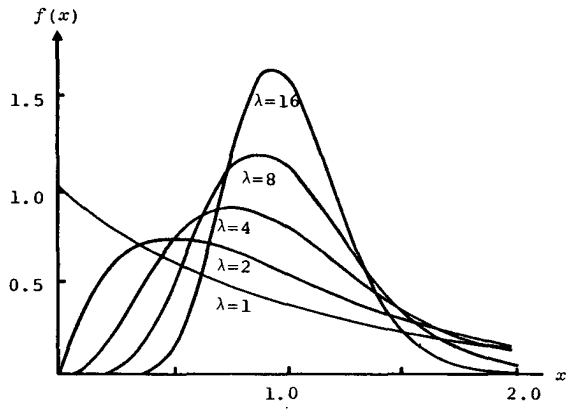


Fig.2. Gamma family of frequency distributions
$f(x) = (\lambda/\mu)^{\lambda} x^{\lambda-1} e^{-\lambda x/\mu}/\lceil\lambda$

These and other considerations underlying the point event assumption and concerning the summability of (1) and (2) are given in Revfeim (1982b).

3. SIMPLE MODELS

Assuming a negative exponential size distribution with

$$F(x) = 1 - e^{-x/\mu}$$

$$f(x) = (1/\mu)\, e^{-x/\mu} \qquad (3)$$

$$f_r(x) = (x/\mu)^{r-1}\, e^{-x/\mu}/\mu\ (r-1)!$$

and a recurrence rate ρ such that $e^{-\rho}$ is negligible gives the maximum value density function

$$g_r(x) \simeq \rho T \exp(-x/\mu - \rho T e^{-x/\mu}), \qquad (4)$$

and the totals density function

$$h_T(x) \simeq \exp(-\rho T - y^2/4\rho) I_1(y), y = 2(\rho T x/\mu)^{\frac{1}{2}} \qquad (5)$$

where I_1 is a modified Bessel function.
An early derivation of (4) in a non-stochastic context given by Epstein (1949) with a more general presentation by Zelenhasic (1970). The

density function (5) was given by Le Cam (1961) for rainfall totals although it had been used earlier in the study of accumulated inputs to storage dams. Versions of (5) have been used by Buishand (1977) and Ozturk (1981) but not emphasizing the simple recurrence rate and mean size interpretation of the parameters.

The recurrence rates of rainfall events are meso-scale parameters with small spatial variation so when mapped should follow the pattern of major topograpical features. The mean size parameter is a more localised property being sensitive to small scale land features. Thus size estimates should only be mapped if the observing site density reflects the local contour variability.

4. ESTIMATION

Where data are observed over unit intervals of T (4) and (5) are functions of ρ and μ. Equating the sample mean $\bar{x}$ and sample variance v to the population value we have for maxima (Revfeim, 1983a)

$$\bar{x} = \mu(\ln\rho+\gamma),\ v = \mu^2\pi^2/6\ ;$$

$$\hat{\mu} = (6v)^{\frac{1}{2}}/\pi\ ,\ \hat{\rho} = \exp\ (\bar{x}/\hat{\mu}-\gamma) \qquad (6)$$

where $\gamma = 0.577...$ is Eulers constant and for totals (Revfeim, 1982b)

$$\bar{x} = \rho\mu,\ v = 2\rho\mu^2\ ;\ \hat{\mu} = v/2\bar{x}\ ,\ \hat{\rho} = \bar{x}/\hat{\mu}. \qquad (7)$$

These initial moment estimators have intuitively reasonable orders of magnitude. The estimate of the mean size of the underlying events is of the same order as the variance of the *totals* (scaled by the mean total) whereas it is of the order of the square-root of the variance of the *extremes*.

For small samples the moments can give anomalous parameter estimates, particularly for small variances which produce large recurrence rates in both cases. In most rainfall régimes rates should be less than the average number of raindays per month (Thompson, 1983) since a single event will sometimes produce rain on two or more consecutive days. If this is not the case then the model may be inappropriate for the events are probably correlated. As events producing traces of rain are of little importance the average number of raindays above nominal 1 mm threshold is a realistic upper bound on estimates of ρ. Similarly large variability in extremes of totals may give small estimated recurrence rates such that $e^{-\rho}$ is no longer negligible. It is therefore sensible to substitute a nominal value of 3 if a lower estimate of ρ is obtained.

Maximum likelihood estimates are not much more difficult to obtain by iterative methods (Revfeim, 1982b, 1983a). The procedures are based on calculations of the sums.

$$S_k = \sum_{j=1}^{n} (x_j/\mu)^k \exp(-x_j/\mu) \qquad (8)$$

for $k = 0, 1, 2$ in the case of maxima and

$$T_k = \sum_{j=1}^{n} \{y_i\ I_0(y_j)/I_1(y_i)\}^k \exp(-x_j/\mu) \qquad (9)$$

for $k = 1, 2$ in the case of totals. In (9) the ratio of modified Bessel functions can be quite accurately approximated by $(1+1/2y)(1+1/2y^2)$ for $y > 1$ and $(1+e^{-y})(1.277+e^{-y}/y)$ for $y < 1$.

It is a simple matter to introduce a time varying rate of occurence with period of 1 year and this leads to simple harmonic expressions for the monthly recurrence rates of a Poisson process (Revfeim, 1982a). Since seasonal rainfall processes are driven by solar heating it is natural to assume a predominant single harmonic at latitude beyond the extremes of solar declination but include a second harmonic at latitudes within 23½° of the equator. The estimation procedures involves computation of the S_k or T_k on a monthly basis and the iterative updating of 6 or 10 parameters for one or two harmonics respectively.

5. DATA BOUNDS

In the estimation of parameters it is necessary to recognise that few data sets are perfect. However, it is not necessary to describe values as 'missing' if observations for a few days in the month are lost, or to use available totals or maxima in the unual analysis and describe the results as 'suspect'. If all observations had been recorded the available maxima and total may have been exceeded so these available data are lower bounds of the true values. Such censored data can easily be included in a likelihood formulation and the maximum likelihood equations differ little from those for the uncensored case (Revfeim, 1983b).

Upper bounds can also occur where the maximum or total of interest can be dominated by some other observation. This is particularly so for maximum rainfalls over increasing time intervals where, for instance, 72 hour rainfalls can be dominated by 48 hour rainfalls. Rainfall observations over intervals on a fixed time scale provide both lower and upper bounds on rainfalls over the same intervals on an arbitrarily located time scale (Revfeim, 1983c).

Upper bounds also occur in the case of classified extremes. It is quite usual for rainfall to occur as a result of frontal movement from a particular direction. It is not usual for this information to be recorded with the magnitudes of maxima although the dates of occurremce are commonly recorded. Hence a histogram of prevailing weather directions associated with maxima can be established from recorded climate observations or synoptic messages. This histogram is likely to show modes for certain prevailing weather direction. For example Table 1 shows the 9 a.m. wind directions in octants on days of months maximum 1 day rainfalls at Wellington, N.Z.

Table 1. Frequencies of 9 a.m. wind directions preceding 1-day maximum rainfalls for Wellington, N.Z. (monthly records 1862-1971)

NW	N	NE	E	SE	S	SW	W	CALM
407	180	52	11	260	178	100	11	70

Thus rainfall maxima can be analysed within direction classes assuming data within a particular class to be valid observations. Data from other classes are upper bounds of maxima for the particular class. This provides estimates of recurrence rates and mean sizes for each direction class which with the harmonic analysis of monthly data may show differing seasonal patterns. Analysis of the 1-day maximum rainfalls within the predominant northerly (NW, N, NE) and southerly (SE, S, SW) classes gave recurrence rates and mean sizes as shown in Table 2.

Table 2. Estimates monthly recurrence rate, r, mean event size, m, (mm) with amplititude, a, and phase angle, p, of seasonal pattern. (Standard errors shown in brackets)

	r	a_r	p_r	m	a_m	p_m
All	5.7 (.2)	1.4 (.3)	135 (12)	13.6 (0.3)	1.6 (0.4)	-87 (15)
Nth	3.6 (.2)	0.8 (.3)	111 (21)	11.5 (0.4)	1.2 (0.5)	-103 (23)
Sth	1.8 (.1)	0.5 (.1)	138 (18)	16.4 (0.6)	2.3 (0.9)	-93 (19)

A more obvious example of this latter situation is to be found in maximum wind gust data which are usually recorded with direction.

In all these cases the maximum likelihood procedure can utilise the information content of data censored at variable bounds together with any uncensored onservations.

6. MIXED PROCESSES

Since rainfall is the result of convective or frontal processes so the frequency distribution of all events could be represented as a mixture of the individual distributions.

$$g(x) = pg^{(1)}(x) + qg^{(2)}(x),\ p+q = 1 \qquad (10)$$

However the fact that one type of rainfall may preclude the other and that size distributions may differ as well as recurrence rates does not lead to a distribution of rainfall totals in this manner. The superposition of two assumed independent Poisson processes gives a distribution of total rainfall in terms of confluent hypergeometric functions. If this model proves to be tractable then it should agree with an intuitive notion of the relative frequency of occurrence and relative magnitude of the rainfall types.

Observed maxima from a mixture of types represent an *exclusive or* situation with probability distribution function

$$G(x) = \prod_k G^{(k)}(x) \qquad (11)$$

and density function

$$g(x) = \sum_k g^{(k)}(x) \prod_{j\neq k} G^{(j)}(x) \qquad (12)$$

This formulation was used by Waylen and Woo (1982) for floods produced by snowmelt and rainfall generated processes. However these authors assumed an ability to classify floods on the basis of 4-day antecedant rainfall above and below a 25 mm threshold: in effect this meant a summer/winter classification. Hence the parameter estimates of each class were estimated separately and the product p.d.f (11) was shown to provide a better fit.

Joint estimation of the four parameters by maximum likelihood methods using the density function (12) should produce comparable results without the need for arbitrary classification. Such an analysis would presume a moderate annual recurrence rate of each process with $e^{-\rho}$ negligible in both cases.

Another application of some interest is the problem of 'outliers' where a few observations stand out from the rest and result in a poor fit to a single extreme value p.d.f. A recent discussion of this problem by Kishihara and Gregory (1982) again uses subjective means for classifying the data types. The observation that individual stations "often change" classes suggests that the problem has not been resolved. An objective analysis is at least conceivable where the 'outliers' are assumed to be relatively rare events with recurrence rates such that $e^{-\rho}$ is not negligible. In this case we cannot assume that the size distribution underlying these outlying extremes is negative exponential. A distribution with non-zero mode is required. Not surprisingly joint parameter estimation is sensitive to the choice of initial parameters in the iterative procedure and a successful fitting has yet to be achieved.

7. EXTENDED MODELS AND OTHER PHENOMENA

The exponential distribution of event sizes seems to provide a good fit in the hydrological situation. However, marginal improvements may be made by fitting a size distribution from the gamma family of order $1 + \varepsilon$ with ε small which has a zero left tail. For small ε the function $\sqrt{1+\varepsilon}$ can be approximated by $1-\varepsilon\gamma$ and the density function for unit intervals equivalent to (4) becomes

$$g(x) \simeq \rho(1+\varepsilon\gamma)(x/\mu)^{\varepsilon}\mu^{-1}\exp(-x/\mu-\rho(1+\varepsilon\gamma)e^{-x/\mu}).$$

The maximum likelihood estimates of ρ and μ for this function can be found for $\varepsilon = 0$, 0.1 and 0.2. For a sample of annual 10 minute rainfall maxima the mean likelihood was still increasing at $\varepsilon = 0.2$ indicating non-zero modality in the size distribution. This is expected from the discussion in section 2. The mean likelihood was greatest for 1-hour maxima at $\varepsilon = 0.1$ and for 24 hour maxima at $\varepsilon = 0$. This trend shows how the lower tail of the size distribution becomes less important as the sample range increases relative to the sample mean.

Traditionally wind gusts have also been analysed using the Gumbel distribution. Considering the physical interpretation of an exponential gust size distribution this implies that most frequent gust size is infinitesimally small which defies intuition. Not surprisingly moment estimators of parameters of (4) for gust data give a recurrence rate of the order of 10^5 events per month. That this model is inappro-

priate for gusts is also shown by the mean size underlying the gusts being less than the mean wind speed!

A recent analysis of gusts using higher order gamma distributions for the size of event (Revfeim and Hessell, 1983) indicates that improved fits can be obtained. Recurrence rates fall to a two-digit number or less for gamma functions of order 8 which also increases the mean gust speed to well above the mean wind speed.

It is necessary to assume that gusts underlying the maxima are independent which conflicts with the clustered nature of gusts. Similar clustering applies to intense rainfalls over short intervals of less than one minute as well as the non-zero modality discussed earlier. Again the negative exponential distribution becomes inappropriate for size of events and further research is required on the clustering problem.

In contrast to the discrete nature of the rainfall process and semi-discrete nature of clusters of wind gusts the maxima of continuous variables such as pressure and temperature have also been fitted by the Gumbel method. Although there are seasonal and diurnal tides of atmospheric pressure the maxima are dominated by the fairly random passage of ridges and troughs. Thus the recurrence model may still be a reasonable approximation but the size distribution is certainly modal being positively skewed for maxima and negatively skewed for minima.

For temperature the dominant seasonal and diurnal cycles cannot be ignored in comparison with the effects of warm and cold fronts. In terms of physically meaningful parametric models the maxima of continuous processes in the atmosphere are a fruitful field for study.

8. SUMMARY

The simplest stochastic models have been described and qualified in their application to imperfect data sets. Mapping of estimated rainfall recurrence rates is more informative and interpretable than isohyets.

The next stage of realism for rainfall totals is to include a distribution of event durations with corrections for the possibility of partitioning rainfall events at observation times. For maximum rainfalls a characterisation of clustering within events is required for short time intervals. Such a model is certainly necessary for wind gusts.

The algorithm for processing monthly rainfall extremes or rainfall totals is given in Wang 2200 BASIC as an Appendix.

9. ACKNOWLEDGEMENTS

The figures were drawn by Mrs C. Kreft and the type script prepared by Mrs R. Kraal of the N.Z. Meteorological Service.

10. REFERENCES

Buishand, T.A., 1977: Stochastic modelling of daily rainfall sequences. Meded. Lansbouw-hogesch. Wageningen, 77/3.

Epstein, B., 1949: A modified extreme value problem. Ann.Math.Stat., 20, 99-103.

Kishihara, N., and Gregory, S., 1982: Probable rainfall estimates and the problems of outliers. J. Hydrol., 58, 341-356.

Le Cam, L., 1961: A stochastic description of precipitation. In J. Neyman (Ed.), Proc. 4th Berkley Symp. on Mathematical Statistics and Probability. Univ. of California Press.

Öztürk, A., 1981: On the study of a probability distribution for precipitation totals. J. Appl. Meteor., 20, 1499-1505.

Revfeim, K.J.A., 1982a: Seasonal patterns in extreme 1-hr rainfalls. Water Resour. Res., 18, 1741-1744.

————————, 1982b: Comments "On the study of a probability distribution for rainfall totals". J.Appl.Meteor., 21, 1942-1945.

————————, 1983a: On the analysis of extreme rainfalls. J. Hydrol., 62, 107-117.

————————, 1983b: Generating mechanisms of, and parameter estimators for, the extreme value distribution. Aust. J. Statistics.

————————, 1983c: Identification and interpretation of three data bounds occurring in extreme rainfall records. J. Appl. Meteor.

————————, and Hessell, J.W.D., 1983: More realistic distributions for extreme wind gusts. Unpublished paper.

Thompson, C.S., 1983: Homogeneity analysis of rainfall series. Unpublished paper.

Waylen, P., and Woo, M-K., 1982: Prediction of annual floods generated by mixed processes. Water Resour. Res., 18, 1283-1286.

```
10 REM PROGRAM FOR MONTHLY RAINFALLS
20 INPUT "ENTER TYPE; EXTREMES=0, TOTALS=1",P:Q=1-P:SELECT D
30 DIM L(6),M(6),P(6),Q(6),T(12),C(12),D(6,6),U(6,6),V(6,6),W(6,6)
40 DIM X(100,12):REM X() IS MATRIX OF MONTHLY DATA
50 REM P() IS VECTOR OF PARAMETERS (R0,R1,R2,M0,M1,M2)
60 REM Q() "    "    "  PARAMETER INCREMENTS
70 REM T() "    "    "  MONTHLY TOTALS
80 REM L() "    "    "  MAX-LIKELIHOOD EVALUATED AT P()
90 REM V() IS THE INFORMATION MATRIX      "     "   "
100 INPUT "NO.YRS",N:PRINT "ENTER MTHLY DATA":MAT INPUT X
110 MAT D=IDN:MAT P=ZER:C,L,Y,W=0:FOR J=1 TO 12:Z,K,X,V=0
120 FOR I=1 TO N:D=X(I,J):IF D*(999-D)>0 THEN 140
130 IF D=999 THEN 150:Z=Z+1:GOTO 150
140 D=D-X:K=K+1:X=X+D/K:V=V+(1-1/K)*D*D
150 NEXT I:D=X-Y:L=L+K:Y=Y+K*D/L:W=W+V+K*(1-K/L)*D*D
160 C=K+Z:T(J)=K*X:C(J)=K+Z:GOSUB 380
170 FOR K=1 TO 3:I=K-INT(K/3):L(K)=I*R:L(K+3)=I*M:NEXT K
180 GOSUB 370:MAT Q=D/L:MAT P=P+Q
190 NEXT J:X=Y:V=W:K=L:T=K*X:GOSUB 380
200 ON P+1 GOSUB 410,430:IF ABS(D)>.001 THEN 200
210 MAT P=(1/12)*P:PRINT "MLE",R,M
220 MAT V=ZER:MAT L=ZER:FOR J=1 TO 12:GOSUB 370
230 R=P(1)+P(2)*C+P(3)*S:M=P(4)+P(5)*C+P(6)*S:GOSUB 450
240 FOR I=1 TO 3:FOR K=1 TO 3:U(I,K)=D:U(I+3,K),U(I,K+3)=E
250 U(I+3,K+3)=F:NEXT K:Q(I)=A:Q(I+3)=B:NEXT I
260 MAT M=D*Q:MAT W=U*D:MAT U=D*W:MAT V=V+U:MAT L=L+M:NEXT J
270 MAT V=INV(V):MAT Q=V*L:MAT P=P+Q:E=0
280 FOR K=1 TO 6:E=E+ABS(Q(K)/P(K)):NEXT K:PRINT E:IF E>.01 THEN 220
290 MAT U=ZER:FOR I=1 TO 6:Q(I)=SQR(V(I,I)):FOR J=1 TO I
300 U(I,J)=V(I,J)/Q(I)/Q(J):NEXT J:NEXT I:MAT PRINT U
310 MAT PRINT P:MAT PRINT Q:J=2:GOSUB 320:J=5:GOSUB 320:STOP
320 X=P(J):Y=P(J+1):A=SQR(X*X+Y*Y):P=ARCSIN(Y/A)
330 V=X*X*V(J,J)+2*X*Y*V(J,J+1)+Y*Y*V(J+1,J+1):B=SQR(V/N)/A
340 W=V(J,J)*Y*Y-2*V(J,J+1)*X*Y+V(J+1,J+1)*X*X
350 Q=SQR(W/N)*180/#PI/A/A:PRINTUSING 360,A,B,P,Q:RETURN
360 %AMPLITUDE=##.## (#.##)    PHASE=###.## (##.##)
370 D(2,2),D(5,5),C=COS(J*30):D(3,3),D(6,6),S=SIN(J*30):RETURN
380 M=P*V/K/2/X+Q*SQR(6*V/K)/#PI:R=P*X/M+Q*EXP(X/M-.5772)
390 D=2:IF R<D THEN 400:D=20:IF R>D THEN 400:D=R
400 PRINT R,M:R=D:RETURN
410 Y=X/M:S0,S1,S2=0:FOR J=1 TO 12:GOSUB 470:NEXT J
420 D=(S0+S1-Y*S0)/(S2+Y*(S0-S1)):M=M/(1+D):R=L/S0:RETURN
430 S1,S2=0:FOR J=1 TO 12:GOSUB 520:NEXT J:GOSUB 560:K=D*F-E*E
440 F=(F*A-E*B)/K:R=R+F:E=(D*B-E*A)/K:M=M+E:D=E*F/R/M:RETURN
450 S0,S1,S2=0:ON P+1 GOSUB 470,520
460 T=T(J):C=C(J):ON P+1 GOSUB 500,560:RETURN
470 FOR I=1 TO N:D=X(I,J):IF D=999 THEN 490:Z=D/M:E=EXP(-Z)
480 S0=S0+E:S1=S1+Z*E:S2=S2+Z*Z*E
490 NEXT I:RETURN
500 Y=T/M:E=S1/M:A=C/R-S0:B=(Y-C)/M-R*E
510 D=C/R/R:F=(2*B+(C+R*S2)/M)/M:RETURN
520 FOR I=1 TO N:D=X(I,J):IF D*(D-999)=0 THEN 550
530 Y=2*SQR(R*D/M):U=1/2/Y:IF Y>1 THEN 540:U=EXP(-Y):Y=1.277*Y
540 Z=(1+U)*(1+U/Y):U=Y*Z:S1=S1+U:S2=S2+U*U
550 NEXT I:RETURN
560 Y=T/M:Z=S2/4-R*Y:E=S1/2:A=E/R-C:B=(Y-E)/M
570 D=Z/R/R:E=(E-Z)/R/M:F=(2*B+Z/M)/M:RETURN
```

On the estimation in the case of an exponential distribution with two parameters with an application to the problem of large values (exceedances).

R. Sneyers and M. Vandiepenbeeck
Institut royal météorologique de Belgique

1. Introduction.

The problem of large values may be considered in two different ways according as they are defined as the portion of the whole population of observations exceeding a given percentile or as being the k largest values accompanying the yearly extreme.

Following Langbein [1] and Chow [2], Berger et al [3] have shown that the distribution of exceedances follow an exponential distribution with two parameters when the extreme distribution is the first asymptote of Fisher-Tippett, the variate x being replaced by $\log x$ (naperian logarithm) when the extreme distribution is the second asymptote of Fisher-Tippett.

On the other hand it is easy to establish [4] that the distribution of large values (k largest values accompanying the yearly extremes) follow an asymptote of Fisher-Tippett of the same kind as the extreme.

The purpose of this paper is to show with an example that the two solutions are not equivalent. The problem of estimation in the case of an exponential distribution with two parameters will however be first considered.

2. Adjustment of an exponential distribution with two parameters.

The exponential distribution with two parameters may be defined by the distribution function :

$$F(x) = 1 - \exp\left[-\frac{x-\mu}{\sigma}\right] \quad \text{with } x \geqslant \mu > 0 \qquad (1)$$

and if $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n$ is the ordered sample to be adjusted with this distribution, the method used is generally based on the fact that (1) leads to :

$$x = \mu - \sigma \log [1 - F(x)]. \qquad (2)$$

μ and σ are then estimated from the equations obtained by putting in (2) : $x = x_i$ and $F(x_i) = i/(n+1)$.

This problem of estimation is however a particular case of a more general problem.

Let

$$F(x, \theta) \quad \text{with} \quad x \geqslant 0 \qquad (3)$$

be a distribution function of a non negative variate, θ being uni- or multidimensional.

In certain cases it may be that the lower bound of the variate is $\mu > 0$, the distribution function to be considered being than :

$$F(x - \mu, \theta) \quad \text{with} \quad x \geqslant \mu > 0 \qquad (4)$$

and being derived from (3) by introducing in (3) the location parameter μ.

Examples of possible such distributions are the exponential distribution, the lognormal distribution, the gamma distribution and the logarithmic double exponential (Weibull) distribution.

The maximum likelihood solution leads to the estimation :

$$\hat{\mu} = x_1 \qquad (5)$$

but owing to the fact that the distribution function of the lowest value is given by :

$$\varphi(x, \mu, \theta) = 1 - [1 - F(x - \mu, \theta)]^n \qquad (6)$$

a more realistic solution is given by putting :

$$\varphi(x_1, \hat{\mu}, \hat{\theta}) = 1/2 \qquad (7)$$

To find this solution one may consider for μ a series of increasing values $0 < \mu_1 < \mu_2 \ldots < x_1$, estimate for each of these values the corresponding value θ_j of θ and proceed by successive approximations to find the solution of (7). The condition for the existence of a solution is of course that we have :

$$\varphi(x_1, \hat{\theta}) > 1/2 \quad \text{with} \quad \mu = 0 \qquad (8)$$

The goodness of fit of the adjustment may then be tested by calculating the value of the distribution function of the largest value for x_n, that is :

$$\phi(x_n, \hat{\mu}, \hat{\theta}) = [F(x_n, \hat{\mu}, \hat{\theta})]^n \qquad (9)$$

the critical values being at the 0,05 level of significance : 0,025 and 0,975.

In the case of the exponential distribution (1), with $\bar{x} = \Sigma x_i/n$, equation (7) becomes :

$$\left[\exp -\frac{x_1 - \mu}{\bar{x} - \mu}\right]^n = 1/2 \qquad (10)$$

that is :

$$\hat{\mu} = (x_1 + \frac{\log 0{,}5}{n} \cdot \bar{x})/(1 + \frac{\log 0{,}5}{n}) \qquad (11)$$

3. Two kinds of large values. An example.

If we have a sample of N years of observations, the subsample of k N exceedances considered by Langbein [1], Chow [2] and Berger [3] is given by the k N largest values of the whole sample. On the contrary, a different subsample of k N

elements is obtained if we consider for each year the k largest values [4]. If the distribution function of the whole sample is of the exponential type, the first case will lead for the large values to an exponential distribution with two parameters and the second case to a double exponential distribution.

As an example we have considered four years of atmospheric SO_2 - concentrations at Gent (Belgium), station 701 from 1977 to 1980. The subsamples are obtained by extracting :

(1) the 4 x 10 = 40 largest values of the whole sample (table 1) and

(2) the 10 largest values for each of the four years (table 2).

Table 1. Atmospheric SO_2 - concentrations at Gent (701) in $\mu g/m^3$ from 1977 to 1980. Ordered largest values of the whole sample (40 values).

135	135	135	135	139	139	139	146	146	154
158	158	161	161	180	180	184	184	191	191
195	195	195	203	206	210	210	218	221	221
240	255	266	270	285	297	297	304	353	360

Table 2. Atmospheric SO_2 - concentrations at Gent (701) in $\mu g/m^3$. 10 largest values for each year from 1977 to 1980.

1977	135	139	146	180	180	195	195	203	206	210
1978	109	116	116	135	135	135	161	191	218	221
1979	135	146	184	240	266	270	285	297	304	360
1980	158	161	184	191	195	210	221	255	297	353

3.1. Adjustment of the exponential distribution.

a) The 40 largest values of table 1 give $\bar{x} = 203{,}8$. With equation 2(11) we have thus :

$$\hat{\mu} = 133{,}79$$

The distribution function of the 40 large values is thus :

$$F(x) = 1 - \exp\left(-\frac{x - 133{,}79}{70{,}01}\right) \qquad (1)$$

and the probability associated to the largest value 360 is :

$$[F(360)]^{40} = 0{,}20 \qquad (2)$$

result which shows a good fit of the adjusted distribution.

b) The 10 largest values for each year 1977 to 1980 of table 2 give : $\bar{x} = 200{,}95$. With equation 2(11) we have thus :

$$\hat{\mu} = 107{,}38$$

The distribution function becomes thus :

$$F(x) = 1 - \exp\left(-\frac{x - 107{,}38}{93{,}57}\right) \qquad (1)$$

and the probability associated to the largest value is this time :

$$[F(360)]^{40} = 0{,}06 \qquad (2)$$

result which shows a bad fit.

3.2. Adjustment of the double exponential distribution.

The distribution function of the double exponential distribution is :

$$F(x) = \exp\left[-e^{-\frac{x - \mu}{\sigma}}\right] \qquad (1)$$

Using the approximate maximum likelihood solution given by Kimball (cf [5] p. 54), we have with the data of table 2 :

$$\hat{\sigma} = 51{,}08 \quad \text{and} \quad \hat{\mu} = 171{,}5 \qquad (2)$$

The probabilities P_1 and P_2 associated to the lowest and to the largest value of the sample are thus :

$$P_1 = 1 - [1 - F(109)]^{40} = 0{,}74$$

$$P_2 = [F(360)]^{40} = 0{,}37 \qquad (3)$$

results which show this time a good fit.

4. Conclusion.

As was already suggested in (4) the two kinds of large values have different distributions.

They have also different uses. In this connection, it should be reminded that the second kind of large values is especially useful for estimating extreme distributions when the series of observations are relatively short.

References

[1] Langbein, W.B. (1949). Annual floods and the partial duration flood series. Trans. Am. Geophys. Un. 30, 879-881.

[2] Chow, V.T. (1950). Discussion on "Annual floods and the partial duration flood series" by W.B. Langbein, Trans. Am. Geophys. Un. 31, 939-941.

[3] Berger, A., Melice, J.L. and Demuth, Cl. Statistical distributions of daily and high atmospheric SO_2 - concentrations, Atmospheric Environment, to be published.

[4] Sneyers, R. and Vandiepenbeeck, M. On the use of large values for the determination of the distribution of maximum values, Arch. Met. Geoph. u. Bioclim. B., to be published.

[5] Sneyers, R. (1975). Sur l'analyse statistique des séries d'observations. Organisation météorologique mondiale. Note technique n° 143.

THE EFFECT OF SEASONAL VARIATION AND SERIAL CORRELATION ON THE EXTREME VALUE DISTRIBUTION OF RAINFALL DATA

T.A. Buishand

Royal Netherlands Meteorological Institute (KNMI)
De Bilt, The Netherlands

1. INTRODUCTION

Statistical information about the occurrence of heavy rainfall can be given in different ways. For hydrological applications the data are often presented in the form of extreme value statistics (usually annual maxima). But this is not for every user the most suitable form of presentation. For instance, in radio engineering the cumulative distribution of the ordinary 1-minute or 5-minute intensities is preferred to the distribution of their annual maxima.

There is of course a relation between the parent and extreme value distribution. For most applied scientists this relation is well-known in the situation of independent random variables with the same distribution. Unfortunately rainfall data over short intervals are neither independent nor identically distributed.

In most parts of the world there is an obvious seasonal variation in the occurrence of high rain rates. For instance, in De Bilt (The Netherlands) 5-minute intensities larger than 30 mm/hour are only found during the warm summer period May-September. Often there is not only an annual cycle but also a daily cycle.

Fig. 1 gives four situations in which there is at least one 5-minute interval with an average intensity greater than 10 mm/hour. In case A there is a single 5-minute interval with a rain rate greater than 10 mm/hour, whereas in the other cases there is some clustering of high rain rates as a consequence of dependence between the rainfall amounts of neighbouring 5-minute intervals. In case B there is a run of two adjacent intervals in which the rain rate exceeds the threshold of 10 mm/hour. The cases C and D are more complicated since there are two runs in which the threshold is exceeded. In case C these runs are situated in the same rainy spell, whereas in case D the two runs are separated by a short dry period. Such situations are not very likely unless the data have a complex correlation structure.

A good insight into the relation between the parent and extreme value distribution of such data requires some knowledge of modern extreme value theory. In this paper a number of mathematical results about the distribution of maxima are discussed which are useful for the applied scientist to evaluate the effect of seasonal variation and dependence.

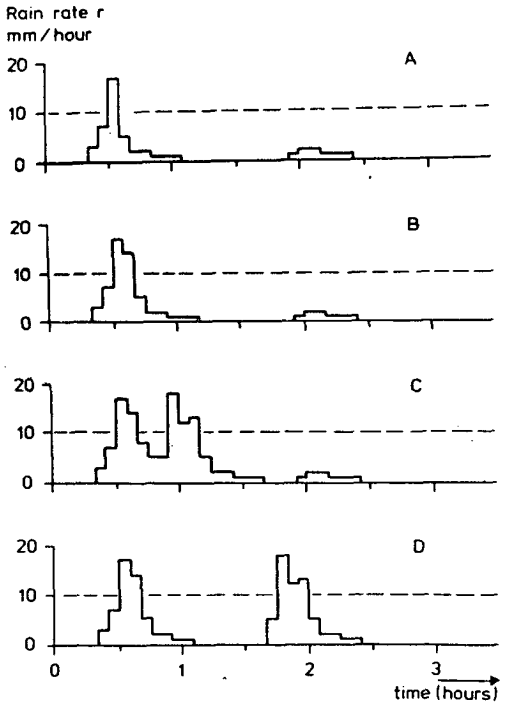


Fig. 1. Rainfall events with high 5-minute rain rates.

In Section 2 some notes are given on the extreme value distribution of independent random variables. Then maxima of sequences of dependent random variables are dealt with in Section 3. Finally, an application of the theory to 5-minute rain rates at De Bilt is discussed in Section 4.

2. THE DISTRIBUTION OF MAXIMA OF INDEPENDENT RANDOM VARIABLES

Let $R_1, R_2, \ldots$ be a sequence of random variables and write $Z_n = \max(R_1, R_2, \ldots, R_n)$. In our application to the rainfall data the R_i's will refer to 5-minute rain rates in a particular year ($n = 105,120$).

In the first instance we assume that the R_i's are independent with common distribution function:

$$F(r) = Pr(R_i < r),\ i = 1, \ldots, n. \qquad (1)$$

Then for the distribution of the maximum Z_n we obtain:

$$Pr(Z_n < r) = Pr(R_1 < r, \ldots, R_n < r) = F^n(r). \qquad (2)$$

For large n the distribution of Z_n is determined by the shape of the right tail of the parent distribution $F(r)$. Using the approximation $\ln F(r) \sim -\{1-F(r)\}$ for large r we obtain for the distribution of Z_n:

$$Pr(Z_n < r) \sim e^{-n\{1 - F(r)\}} \qquad (3)$$

The quantity $n\{1 - F(r)\}$ gives the expected number of the R_i's that exceed the threshold r. Denoting this quantity by $a(r)$ the approximation (3) becomes:

$$Pr(Z_n < r) \sim e^{-a(r)} \qquad (4)$$

The number of exceedances $S_n(r)$ of the level r has a binomial distribution which for large r can be approximated by a Poisson distribution with parameter $a(r)$. The approximation (4) follows then immediately from the fact that $Pr(Z_n < r) = Pr\{S_n(r) = 0\}$.

The theory can easily be extended to independent sequences with seasonal variation. In general $F_i(r) = Pr(R_i < r)$ slowly varies with i and for large r the process of exceedances can be regarded as a non-homogeneous Poisson process. The quantity $S_n(r)$ has still approximately a Poisson distribution and (4) can be applied with:

$$a(r) = \sum_{i=1}^{n} \{1 - F_i(r)\} \qquad (5)$$

being the expected number of exceedances in a sequence of length n.

Up to now we did not make any assumption about the shape of the tail of the parent distribution. In many applications the tail can be approximated by an exponential distribution. This implies that for large r:

$$a(r) \sim e^{-\sigma_a(r - u_a)} \qquad (6)$$

Substitution of (6) in (4) gives:

$$Pr(Z_n < r) \sim \exp\{-e^{-\sigma_a(r - u_a)}\} \qquad (7)$$

which is a Gumbel distribution with location parameter u_a and scale parameter σ_a. Note from (6) that u_a is the value of r for which $a(r) = 1$, i.e. u_a has expectation of being exceeded just once in a sequence of length n.

3. EXTREME VALUE THEORY FOR SEQUENCES OF DEPENDENT RANDOM VARIABLES

Correlation between successive values in a time series does not always put limitations on the applicability of the approximations given in the previous section. For large n eq. (4) can still be used if the following conditions are satisfied:

(i) The sequence $R_1, R_2, \ldots$ is a mixing sequence. Roughly this condition requires that the various terms in the sequence may only be weakly dependent when their separation in time is large. For instance, in a mixing sequence $Pr(R_1 < r, R_2 < r, R_k < r)$ tends to $Pr(R_1 < r, R_2 < r)\ Pr(R_k < r)$ as $k \to \infty$.

(ii) The random variables R_i and R_{i+k} are pairwise independent in the right tail for every $k \neq 0$, i.e. $Pr(R_{i+k} \geq r | R_i \geq r)$ tends to zero as $r \to \infty$. In particular if we consider exceedances of a high level r there will be no local clustering of such events (case A in Fig. 1).

It goes beyond the scope of this paper to consider these conditions in detail. The interested reader is referred to the review paper of Leadbetter (1978) or to Ch. 3 of Galambos (1978).

Since rainfall sequences are characterized by a relative short memory, we may safely assume that the mixing condition (i) is satisfied. The main problem causes condition (ii). Departures from this condition may have a serious impact on the distribution of maxima as will be demonstrated in the next subsections.

3.1 Sequences of 1-dependent random variables

A sequence of 1-dependent random variables is characterized by the property that the events $\{R_1 < r_1, \ldots, R_j < r_j\}$ and $\{R_{j+k} < r_{j+k}, \ldots, R_n < r_n\}$ are dependent if $k = 1$, but they are independent for $k > 1$. Such a sequence can be generated for instance by a normal first-order moving average process:

$$R_t = A_t + \theta A_{t-1},\ t = 1, 2, \ldots \qquad (8)$$

Here $A_0, A_1, \ldots$ is a sequence of independent normal variates with mean zero and variance σ_A^2. The variables R_t and R_{t+1} are dependent because both are linearly dependent on A_t. There is no correlation between R_t and R_{t+2} because they have no terms of the sequence $A_0, A_1, \ldots$ in common.

For a stationary 1-dependent process it can be shown that for large n the following approximation holds for the distribution of the maximum Z_n (Newell, 1964):

$$Pr(Z_n < r) \sim e^{-n\ Pr(R_i \geq r,\ R_{i+1} < r)} \qquad (9)$$

If the event $\{R_i \geq r, R_{i+1} < r\}$ occurs then there is a down-crossing of the level r at $t = i+1$. The quantity $n\ Pr(R_i \geq r, R_{i+1} < r)$ gives the expected number of down-crossings in a sequence of length n. Since a down-crossing corresponds with the end of a run we may write:

$$Pr(Z_n < r) \sim e^{-b(r)} \qquad (10)$$

wherein b(r) denotes the expected number of runs above the threshold r.

Obviously $b(r) < a(r)$, since there is always at least one event in a run. For the quantity b(r) we can write:

$$b(r) = n\{Pr(R_i \geq r) - Pr(R_i \geq r, R_{i+1} \geq r)\}$$

$$= a(r)\{1 - Pr(R_{i+1} \geq r \mid R_i \geq r\}. \qquad (11)$$

If R_i and R_{i+1} are asymptotically independent in the right tail we have for large r that $a(r) \sim b(r)$ and the eqs. (4) and (10) lead to the same asymptotic distribution of Z_n. However (10) has to be used if $Pr(R_{i+1} \geq r \mid R_i \geq r)$ does not tend to zero as $r \to \infty$. Moreover, for small n this approximation may give better results than (4) in situations that $Pr(R_{i+1} \geq r \mid R_i \geq r)$ slowly converges to zero.

3.2 A normal first-order moving average process

As an example we consider the moving average process (8) with $\theta = 1$ and $\sigma_A^2 = 0.5$. For this process we obtain that var $R_t = 1$ and $\rho = corr(R_t, R_{t+1}) = 0.5$. Further it can be shown that $Pr(R_{t+1} \geq r \mid R_t \geq r)$ tends to zero as $n \to \infty$ (Watson, 1954), so that the ultimate asymptotic distribution of Z_n is given by (4). For small n, however, this approximation gives poor results as is shown in Table 1 for n = 20.

Table 1. $Pr(Z_n < r)$ and some approximations for a stationary normal first-order moving average process (8) with $\theta = 1$ and $\sigma_A^2 = 0.5$ (n = 20).

Threshold r	$Pr(Z_n < r)$ Exact	Eq. (4)	Eq. (10)	Eq. (12)
1.0	0.09	0.04	0.15	0.09
1.5	0.34	0.26	0.38	0.37
2.0	0.69	0.63	0.69	0.72
2.5	0.90	0.88	0.90	0.91
3.0	0.98	0.97	0.98	0.98

When r is not too small (10) gives a much better approximation to the distribution of Z_n. Another approximation to $Pr(Z_n < r)$ can be obtained from (Greig, 1967):

$$Pr(Z_n < r) \approx \Phi(r)\{1 - \sqrt{1-\rho}\,[1-\Phi(r)]\}^{n-1} \qquad (12)$$

which gives also good results for small r. Here $\Phi(.)$ stands for the standard normal distribution function.

A Monte Carlo method was used to obtain the exact probabilities in Table 1 (with 9,999 repetitions).

3.3 Digression on the distribution of run-lengths

To examine the distribution of run-lengths in more detail, let us consider the event that a run with length greater than 2 starts at t = i+1:

$$E = \{R_i < r, R_{i+1} \geq r, R_{i+2} \geq r, R_{i+3} \geq r\}. \qquad (13)$$

For 1-dependent processes we obtain:

$$Pr(E) \leq Pr(R_i < r, R_{i+1} \geq r, R_{i+3} \geq r)$$

$$= Pr(R_i < r, R_{i+1} \geq r)\, Pr(R_{i+3} \geq r) \qquad (14)$$

and thus:

$$\frac{Pr(E)}{Pr(R_i < r, R_{i+1} \geq r)} \leq Pr(R_{i+3} \geq r). \qquad (15)$$

The left-hand side of (15) gives the fraction of runs with a length greater than 2. Since the right-hand side of (15) vanishes as $r \to \infty$ it follows that for large values of the threshold r there will be no or hardly any run with a length greater than 2. So for 1-dependent processes exceedances of a high level r occur either as single peaks (case A in Fig. 1) or paired in a run (case B in Fig. 1). To explain more complex situations like the cases C and D in Fig. 1 we have to consider higher order dependence models. These will be introduced in the next subsection.

3.4 Sequences with higher order dependence

Instead of 1-dependent sequences we may consider more general m-dependent sequences. In an m-dependent sequence events are independent if they are separated by more than m time-units (e.g. an m-th order moving average process). For the distribution of the maximum Z_n of such a sequence it was shown by Newell (1964) that for large n:

$$Pr(Z_n < r) \sim e^{-n\,Pr(D_i)} \qquad (16)$$

where D_i denotes the event:

$$D_i = \{R_i \geq r, R_{i+1} < r, \ldots, R_{i+m} < r\}. \qquad (17)$$

If m = 1 then (16) reduces to (10). The quantity $n\,Pr(D_i)$ gives the expected number of the events $D_1, D_2, \ldots$ that occur in a sequence of length n. If r is large then $Pr(D_i)$ will be small and the occurrence-times of the events $D_1, D_2, \ldots$ form approximately a Poisson process. The right-hand side of (16) gives the probability that none of the events $D_1, D_2, \ldots, D_n$ occur in this limiting Poisson process.

When $Pr(D_i)$ varies with season then in general the limiting process will be a non-homogeneous Poisson process, which leads to the following asymptotic distribution for the maximum:

$$Pr(Z_n < r) \sim e^{-\sum_{i=1}^{n} Pr(D_i)} \qquad (18)$$

Let us now return to case C in Fig. 1. Assume that the rainfall process is a 12-dependent process (i.e. it has a memory of 1 hour). If we choose a threshold of 10 mm/hour then we have two runs, but the event D_i only occurs at the end of the second run. In general, for the distribution of the maximum the number of clusters has to be taken into account rather than the number of runs or individual exceedances. Local clustering of high level exceedances is discussed rigorously in Rootzén (1978) for a class of moving average processes.

For the 1-dependent process exceedances of a high level r may occur in runs, but there is no local clustering of these runs. If we have an m-dependent process with m > 1 then it is possible that runs also occur in bunches. This is not necessary, however. In fact there will be no clustering of runs if for large r:

$$\Pr(D_i) \sim \Pr(R_i \geq r,\ R_{i+1} < r). \tag{19}$$

Then (10) can be used to obtain the asymptotic distribution of the maximum. In contrast with the 1-dependent process runs with a length greater than 2 need not to be rare events.

But if for large r we have that

$$\Pr(D_i) < \Pr(R_i \geq r,\ R_{i+1} < r) \tag{20}$$

then runs above the threshold r occur in bunches and (18) should be used to obtain the distribution of Z_n.

In practical situations it is not always obvious how a cluster of rare events has to be defined. Unlike the situation in the mathematical literature the stochastic structure of the process is seldom known. Some information about clustering of high level exceedances can be obtained from data.

In the cases C and D in Fig. 1 there is strong local clustering of large values. The difference between these cases is that in the former there are only wet intervals between the two runs above the level of 10 mm/hour, whereas in the latter there is a short dry period between the runs. For 5-minute data of De Bilt it turns out that for large r case C is more relevant than case D (see Section 4). Therefore as an alternative to (18) we can try the approximation:

$$\Pr(Z_n < r) \sim e^{-c(r)} \tag{21}$$

where c(r) denotes the expected number of rainy spells in a sequence of length n with at least one 5-minute rain rate greater than or equal to the threshold r. Here a rainy spell is defined as an uninterrupted sequence of wet 5-minute intervals on each side bounded by a dry 5-minute interval. This is in fact a run for which the threshold r is equal to the smallest measurable rainfall amount in the sequence (for the 5-minute data of De Bilt this value of r is 0.16 mm/hour).

Eq. (21) is based on the assumption that for large r the number of clusters in a sequence of length n has a Poisson distribution with parameter c(r). To be more specific about the distribution of Z_n we need to know the form of c(r) for large r. If c(r) is exponential:

$$c(r) \sim e^{-\sigma_c(r - u_c)} \tag{22}$$

then we obtain:

$$\Pr(Z_n < r) \sim \exp\{-e^{-\sigma_c(r - u_c)}\}. \tag{23}$$

Hence Z_n has a Gumbel distribution with location parameter u_c and scale parameter σ_c.

4. APPLICATION TO 5-MINUTE DATA FROM DE BILT

Twelve years of data (1928, 1933, 1951-1960) were digitized from the self-recording rain-gauge in De Bilt. For the 5-minute data estimates of a(r) and c(r) were obtained by counting the number of events and rainy spells with a rainfall intensity of at least r mm/hour. The estimates $\hat{a}(r)$ and $\hat{c}(r)$ are plotted on a logarithmic scale in Fig. 2.

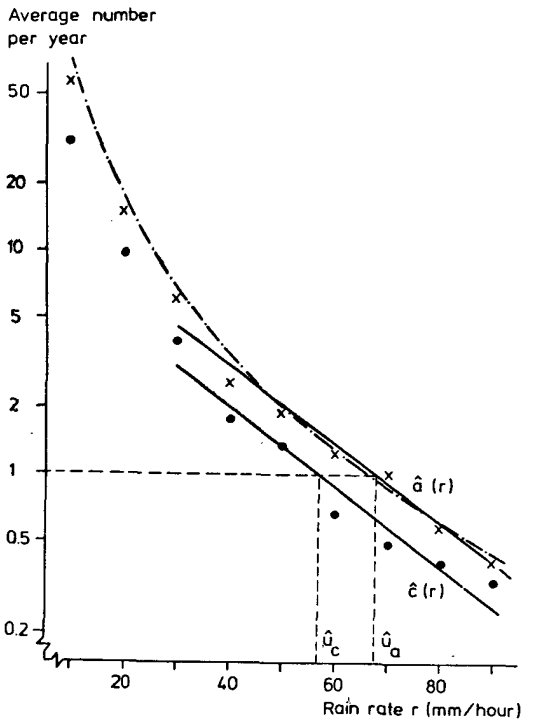


Fig. 2. Estimates of a(r) and c(r) for 5-minute rain rates at De Bilt. The quantities $\hat{a}(r)$ and $\hat{c}(r)$ refer to a sequence with a length of one year. The straight lines represent the exponential approximations (6) and (22), respectively; the curved line is based on the lognormal distribution (24).

From the figure it is seen that even for very high rainfall intensities $\hat{c}(r)$ still differs from $\hat{a}(r)$. The ratio $\hat{a}(r)/\hat{c}(r)$ is roughly about 1.5 which means that there are on average 1.5 exceedances in a rainy spell. A consequence of this local clustering of high rain rates is that the eqs. (4) and (21) do not lead to the same result.

It turns out further that for r ≥ 30 mm/hour

the average number of rainy spells $\hat{c}(r)$ is equal to the average number of days with a rain rate of at least r mm/hour. This means that cases like situation D in Fig. 1 are of minor importance for such high rain rates. Therefore we may expect that (21) gives a reasonable approximation to the distribution of maxima.

For $r \geq 30$ mm/hour straight lines are fitted in Fig. 2. The fit is not too bad and therefore the approximations (6) and (22) can be applied. For $r < 30$ mm/hour the points deviate from the fitted line. However, this has no serious consequences for the distribution of extreme values because the annual maximum is nearly always greater than 30 mm/hour, see Fig. 3.

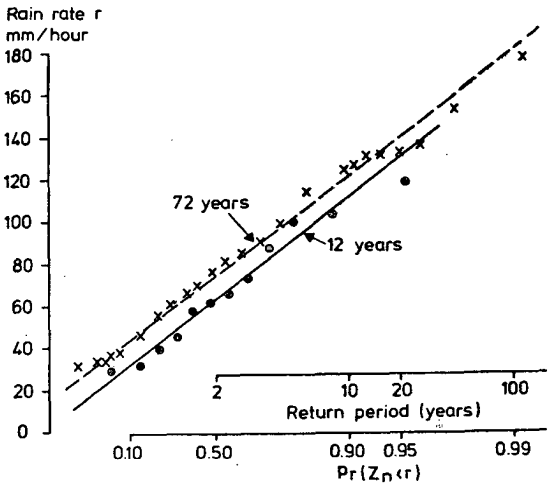


Fig. 3. Gumbel probability plot of annual 5-minute rain rate maxima for De Bilt.

Besides the annual maxima of the 12-year digitized data Fig. 3 gives also the maxima for the 72-year period 1906-1977. These maxima were extracted manually from the pluviograph charts. For both sets of data the Gumbel distribution fits well, but the distributions do not coincide. Except for climatic influences the differences can also be attributed to the fact that the data do not precisely refer to the same quantity. The digitized data are rainfall amounts for fixed 5-minute intervals (e.g. 8.00-8.05, 8.05-8.10, etc.) whereas the annual maxima of the 72-year period refer to variable 5-minute intervals. This leads to the curious fact that the smallest value in the 12-year period is smaller than the smallest value in the 72-year period.

From the straight lines in Fig. 2 we can get an estimate of the parameters u_a, σ_a and u_c, σ_c. The Gumbel lines with these parameters are given in Fig. 4. The upper line with parameters $\hat{u}_a$ and $\hat{\sigma}_a$ should be the correct distribution of the annual maxima if the effect of dependence could be ignored. This is, however, not the case. Although the differences between the two lines are not very large (about 10 mm/hour) the lower line, which is based on a reasonable dependence model, gives a better fit. This line nearly coincides with the solid line in Fig. 3.

To correct for dependence we must have some idea about the average number of exceedances in a

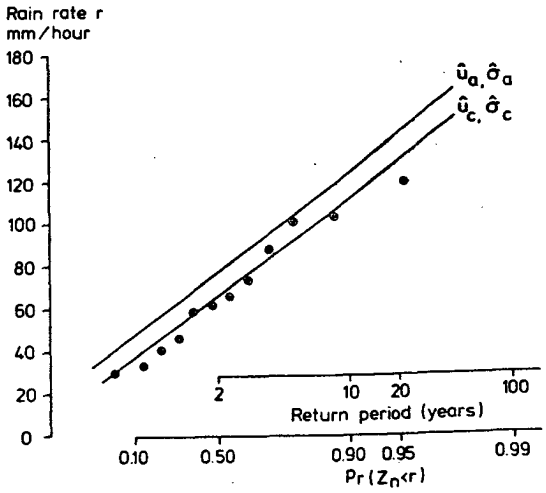


Fig. 4. Gumbel probability plot of annual 5-minute rain rate maxima for De Bilt (12-year period). The Gumbel line with parameters $\hat{u}_a$ and $\hat{\sigma}_a$ is based on the assumption that the data are independent, whereas the Gumbel line with parameters $\hat{u}_c$ and $\hat{\sigma}_c$ is obtained with a model in which dependence is taken into account.

cluster, i.e. the ratio $a(r)/c(r)$. When this quantity is known it is not difficult to derive the extreme value distribution from the parent distribution or the right tail of the parent distribution from the extreme value distribution.

It should be stressed, however, that an extreme value distribution only gives information about the upper tail of the parent distribution. Although we may obtain the exponential tail from the extreme value distribution it is dangerous to extrapolate this result to lower rain rates. To make sensible extrapolations to rain rates in the range of 10 to 30 mm/hour, which are of particular interest in radio engineering, we need some additional information about the shape of $a(r)$ in this range.

From Fig. 2 it is seen that the exponential distribution greatly underestimates the average number of exceedances $a(r)$ for rain rates less than 30 mm/hour. The curved line in this figure is based on a two-parameter lognormal distribution:

$$a(r) = n\{1 - \Phi(\frac{\ln r - \mu}{\sigma})\} \tag{24}$$

where $\Phi(.)$ stands for the standard normal distribution function. The parameters μ and σ can be obtained from u_a and σ_a using the relations (Singpurwalla, 1972):

$$\sigma = \sqrt{2 \ln n} \; / \; (u_a \, \sigma_a) \tag{25}$$

and

$$\mu = \ln u_a - \sigma \, e_n \tag{26}$$

where

$$e_n = \sqrt{2 \ln n} - \frac{\ln \ln n + \ln 4\pi}{2 \sqrt{2 \ln n}} \tag{27}$$

This two-parameter lognormal distribution gives a reasonable fit for values of r between 10 and 90 mm/hour. It is not necessary to include the fraction of wet 5-minute intervals as a third parameter which is usually done in radio engineering (Lin, 1978; Fedi, 1981).

5. SUMMARY AND CONCLUSIONS

The classical theory of extreme values is based on the assumption that the data are independent and identically distributed. Even when these two assumptions are not satisfied the classical theory may give the correct asymptotic distribution of maxima. But the theory should be extended if rare events occur in bunches. Then the distribution of maxima is not directly related to the average number of exceedances of some high level r as in the independent case but to the number of clusters with large values.

It is of course not always obvious how a cluster of rare events has to be defined. But for 5-minute rates at De Bilt it is sensible to look at the number of rainy spells in which the level r is exceeded. This results in a better approximation to the distribution of annual maxima than a model in which the distribution of extremes is related to the average number of exceedances of some threshold r.

6. ACKNOWLEDGEMENT

The author wishes to express his gratitude to Mr. R. Agterberg for his work with the rainfall data.

7. REFERENCES

Fedi, F., 1981: Prediction of attenuation due to rainfall on terrestrial links. Radio Sci., 16, 731-743.

Galambos, J., 1978: The asymptotic theory of extreme order statistics. Wiley, New York, 352 pp.

Greig, M., 1967: Extremes in a random assembly. Biometrika, 54, 273-282.

Leadbetter, M.R., 1978: Extreme value theory under weak mixing conditions. MAA studies in mathematics, Vol. 18: Studies in probability theory (M. Rosenblatt, ed.), 46-110. The Mathematical Association of America.

Lin, S.H., 1978: More on rain rate distributions and extreme value statistics. Bell Syst. Tech. J., 57, 1545-1568.

Newell, G.F., 1964: Asymptotic extremes for m-dependent random variables. Ann. Math. Statist., 35, 1322-1325.

Rootzén, H., 1978: Extremes of moving averages of stable processes. Ann. Probab., 6, 847-869.

Singpurwalla, N.D., 1972: Extreme values from a lognormal law with applications to air pollution problems. Technometrics, 14, 703-711.

Watson, G.S., 1954: Extreme values in samples from m-dependent stationary stochastic processes. Ann. Math. Statist., 25, 798-800.

THE DISTRIBUTION OF DESIGN STORMS FROM EXCEEDANCE SERIES

Walter Zucchini and Peter Adamson

Department of Civil Engineering
University of Stellenbosch
Stellenbosch 7600, South Africa

1. INTRODUCTION

The standard statistical method of estimating design storms from observed annual maximum storms can be described briefly as follows:

Let F_a be the probability distribution function (d.f.) of annual maximum storms. Then under the assumption that these storms are independently and identically distributed, the d.f. of the largest storm in h years is F_a^h. So the storm associated with a design horizon of h years and a risk of occurrence r, is the solution to the equation:

$$F_a^h(x) = 1-r \qquad (1)$$

The design storm, x, is estimated by fitting a model, $\hat{F}_a$, to the d.f. F_a and then solving (1), i.e.

$$\hat{x} = \hat{F}_a^{-1}\{(1-r)^{1/h}\} \qquad (2)$$

In most texts (e.g. Ven Te Chow, 1964) the estimator (2) is derived using additional steps involving the notion of return period.

Often it is of advantage to use exceedance series rather than annual maximum series (e.g. Cunnane, 1973). In this paper we show that, also for this type of data, estimators for design storms can be derived directly. A general method of estimating the standard deviations and confidence limits for the estimates is described.

2. DERIVATION OF ESTIMATORS

Suppose that any rainfall event exceeding a given truncation level x_o in a given time, e.g. a day, is defined to be "storm" of depth x_o+x, $x>=0$. The storms arrive according to some point process at a rate which we assume to be the same for different years, but not necessarily constant within the year. For $j=0,1,2,\ldots$, let $p(j)$ denote the probability of having j storms in a year. We assume that the storm depths are identically and independently distributed with d.f. F and independently distributed of the process of storm arrivals.

Under these conditions the distribution of the biggest storm in h years is the distribution of the maximum of a random number of storms. It can be shown that the d.f., F_h, of the largest exceedance in h years is given by

$$F_h(x) = p_h(0) \qquad \text{for } x = 0$$

$$F_h(x) = \sum_{j=1}^{\infty} p_h(j)F(x)^j \qquad \text{for } x > 0 \qquad (3)$$

where $p_h(j)$ is the probability that there will be j storms, $j=0,1,2,\ldots$, in h years. There is therefore a probability $p_h(0)$ that no storms will occur in the h years, but in the range which is most often of interest this probability is negligible.

Two problems have to be solved in order to exploit (3). Firstly one has to find an expression for p_h in terms of the parameters of p because, particularly if h is large, few if any direct observations on p_h are available. Secondly one has to give a closed expression for the infinite sum. Fortunately it turns out that for the two types of arrival processes of most consequence in applications, namely the Poisson and the negative binomial, both problems can be solved.

Poisson arrivals

Suppose that the number of storms per year varies according to the Poisson distribution with parameter λ, i.e.

$$p(j) = \lambda^j \exp(-\lambda)/j! \qquad ,j=0,1,2,.. \qquad (4)$$

Under the assumption that the numbers of storms in different years are independently distributed the probability of having j storms in h years is given by

$$p_h(j) = (\lambda h)^j \exp(-\lambda h)/j! \qquad j=0,1,2,.. \qquad (5)$$

and (3) can be summed to yield

$$F_h(x) = \exp\{-\lambda h(1-F(x))\} \qquad (6)$$

To estimate the required design storm one has to first estimate λ and F and then compute

$$\hat{x} = x_o + \hat{F}^{-1}\{1+\ell n(1-r)/\hat{\lambda}h\} \qquad (7)$$

For many of the standard distributions $\hat{F}^{-1}$ is not available in closed form and must therefore be computed using either numerical or graphical methods. An exception to this is the exponential distribution:

$$F(x) = 1 - \exp(-x/\theta) \qquad x > 0 \qquad (8)$$

For this case (6) and (7) become

$$F_h(x) = \exp\{-\exp\{-(x-\theta\ell n(\lambda h))/\theta\}\} \quad x > 0 \quad (9)$$

$$\hat{x} = x_o + \hat{\theta}\ell n\{-\hat{\lambda}h/\ell n(1-r)\} \quad (10)$$

Here F_h turns out to be the d.f. of an extreme value (Type 1) random variable with parameters $\theta\ell n(\lambda h)$ and θ. For this particular case it is also possible to derive the approximate standard deviation of $\hat{x}$, namely

$$\text{Std.Dev.}(\hat{x}) \simeq \theta\ell n\{-\lambda h/\ell n(1-r)\}/\sqrt{n\lambda} \quad (11)$$

where n is the length of the historical record in years.

Negative binomial arrivals

The variance of a Poisson distributed random variable is equal to its expectation. In some regions (particularly in arid regions) the sample variance of the annual number of storms is larger than the sample average by an amount which is incompatible with the assumption of a Poisson distribution. For such cases the negative binomial distribution, which is a compounded form of the Poisson, is usually an appropriate alternative:

$$p(j) = \binom{1/\eta+j-1}{j}\left(\frac{1}{1+\eta\lambda}\right)^{1/\eta}\left(\frac{\eta\lambda}{1+\eta\lambda}\right)^{j} \quad j=0,1,2... \quad (12)$$

As in the case of the Poisson distribution the parameter λ is the mean number of storms per year. The parameter η can be regarded as a measure of deviation from the Poisson; it is the difference between the variance and the expectation, divided by the expectation squared. As η tends to zero so the negative binomial tends to the Poisson.

If the number of storms in differnet years are independently negative binomially distributed then the number of storms on h years is also negative binomially distributed. The probability that j storms occur in h years is given by

$$p_h(j) = \binom{h/\eta+j-1}{j}\left(\frac{1}{1+\lambda\eta}\right)^{h/\eta}\left(\frac{\eta\lambda}{1+\lambda\eta}\right)^{j} \quad j=0,1,2.. \quad (13)$$

and (3) can be summed to yield

$$F_h(x) = \{1+\eta\lambda(1-F(x))\}^{-h/\eta} \quad (14)$$

The required design storm is estimated by first estimating η, λ and F and then using

$$\hat{x} = x_o + \hat{F}^{-1}\left(1-\{(1-r)^{-\hat{\eta}/h}-1\}/\hat{\eta}\hat{\lambda}\right) \quad (15)$$

As was the case of Poisson arrivals, for most of the standard distributions $\hat{F}^{-1}$ has to be computed using either numerical or graphical methods.

3. ESTIMATION OF CONFIDENCE LIMITS

To derive exact confidence limits or even standard errors for (7) and (14) is difficult. For some special cases and with considerable effort approximate standard errors can be obtained. One such case is given in (11). Each variation of the following four factors leads to a different distribution of the extimated design storm and consequently to different confidence limits: the distribution of storm arrivals, the method used to estimate its parameters, the distribution of exceedances and the method used to estimate its parameters.

At cost of some additional computing these theoretical difficulties can be circumvented by using the "Bootstrap" method of Efron, (1982). In our context the Bootstrap algorithm to estimate the confidence limits (or the standard errors) of (7) and (15) can be implemented as follows:

Let n be the length of the historical record in years, m_i be the number of storms which occured in year i, i=1,2,...,n, and M the total number of storms. Let x_j, j=1,2,...,M, be the observed exceedances.

Bootstrap Algorithm

Step 1: Select a random sample of size n (with replacement) from the set $\{m_1,m_2,\ldots,m_n\}$ to obtain a Bootstrap sample $m_1^*,m_2^*,\ldots,m_n^*$. Let M* be the sum of the m_i^*.

Step 2: Select a random sample of size M* (with replacement) from the set $\{x_1,x_2,\ldots,x_M\}$ to obtain a Bootstrap sample $x_1^*,x_2^*,\ldots,x_{M^*}^*$.

Step 3: Estimate the design storm $\hat{x}^*$ by the same method as was used to estimate $\hat{x}$, but using the Bootstrap "data" generated in Steps 1 and 2.

Step 4: Repeat Steps 1 to 3 a large number of times keeping a record to the estimates $\hat{x}^*$.

As the number of Bootstrap iterations increases so the sample standard deviation of the Bootstrap estimates $\hat{x}^*$ converges to an estimate of the standard deviation of $\hat{x}$ and the sample percentage points of the Bootstrap estimates converge to estimates of the corresponding percentage points of the distribution of $\hat{x}$. For example, the estimates of the 90% confidence limits for $\hat{x}$ based on 1000 Bootstrap iterations would be the 50-th and 950-th largest values of $\hat{x}^*$.

A number of refinements are possible to improve the performance of the above "uncorrected percentile method" (c.f. Efron, 1982) but to implement these one has to use information about the particular form of the models describing the data, whereas the above method is quite general. Judging by the results of Monte Carlo experiments which we have carried out in this context the above algorithm yields satisfactory estimates using even as few as 100 iterations.

An analogous (and simpler) algorithm can also be used to estimate the confidence limits for (2). However for some of the standard distributions direct estimators are given in Stedinger (1983).

4. EXAMPLES OF APPLICATION

Example 1

The method described above were used to estimate the 1-day design storms at Vryheid, South Africa, for two combinations of design horizon and risk of occurence, viz. h=25, r=1% and h=50, r=10%. A historical record comprising 30 years (1951-1980) of daily rainfall data was used and the truncation level was set to 30 mm. Estimates based on the three 10-year records (1951-1960, 1961-1970, 1971-1980) were also computed as additional indicators of the sampling variation.

The Poisson distribution was fitted to the number of storms per year and a two parameter gamma distribution was fitted to the exceedances. Table 1 gives the (maximum likelihood) estimates for the different records and Tables 2 and 3 the estimated design storms and their corresponding estimated confidence intervals.

Table 1: Maximum likelihood estimates of the Poisson and gamma parameters.

Record	Poisson parameter	gamma scale parameter	gamma shape parameter
1951-1980	7,433	18,55	0,8569
1951-1960	7,600	21,39	0,7980
1961-1970	6,700	17,55	0,8518
1971-1980	8,000	16,69	0,9330

Tabel 2: Estimates and confidence intervals, in mm, (based on 200 Bootstrap iterations) for the design storm with h=25 and r=1%.

Record	design storm	90% confidence interval
1951-1980	204	(175, 238)
1951-1960	227	(171, 279)
1961-1970	193	(142, 247)
1971-1980	192	(149, 224)

Table 3: Estimates and confidence intervals, in mm, (based on 200 Bootstrap iterations) for the design storm with h=50 and r=10%.

Record	design storm	90% confidence interval
1951-1980	174	(156, 194)
1951-1960	193	(149, 229)
1961-1970	164	(113, 195)
1971-1980	164	(136, 189)

The estimated confidence intervals are rather wide, especially for the 10-year records. This is of course a consequence of the inherent variability in the situation relative to the length of each record - it does not reflect on the quality of the estimates for the confidence intervals.

Example 2

To illustrate the performance of the Bootstrap method, estimates for the standard deviation of four design storms were computed with Bootstrap methods and compared with those obtained using (11), a case for which a theoretically based estimator is available.

Daily rainfall data for the period 1861-1960 at the Cape Town Observatory, South Africa, was found to be suitable for this purpose, i.e. the exceedances are approximately exponentially distributed. The extimates which we obtained are given in Table 4.

Table 4: Estimates of the standard deviation of different design storms using 100 Bootstrap iterations and an estimator based on (11).

Design horizon (years)	Risk of occurrence	Bootstrap estimate (mm)	Estimate based on (11) (mm)
20	10%	4,0	4,1
20	1%	5,5	5,7
50	10%	4,7	4,9
50	1%	6,8	6,9
100	10%	5,0	5,2
100	1%	6,4	7,2

Clearly, these results are in good agreement. The Bootstrap estimate is consistently lower than that based on (11). The discrepancy in all but the last case given in Table 3 is mainly due to the approximations made in the derivation of (11). (In fact (11) can be slightly improved by using more terms in the approximation, but this does not seem to be worthwhile.) The last standard error given should be about 7,0. By using an additional 100 Bootstrap iterations the estimate was raised to 6,7. This indicates that for such extreme cases, viz. h=100 and r=1% larger Bootstrap samples should perhaps be used.

5. CONCLUDING REMARKS

The methods discussed in this paper can also be applied to estimate other design events, such as design floods. Naturally the conditions under which the estimators were derived must be met. In particular the truncation level needs to be selected with care because for low truncation levels the assumption that the exceedances are independantly distributed is often violated.

As was pointed out by Efron (1982), Bootstrap methods are "prodigious computational spendthrifts", and such methods would not have been feasible 25 years ago, before the era of cheap and fast computing. The cost of computing confidence limits is a very small fraction of that incurred in, say, measuring and coding meteorological data. For the application discussed in this paper the computational requirements are by no means excessive. The computations for the examples given above were carried out on a typical desk-top micro-computer.

6. REFERENCES

Cunnane, C., 1973: A particular comparison of annual maxima and partial duration series methods of flood frequency prediction. Journal of Hydrology, 18, 257-271.

Efron, B., 1982: The Jackknife, the Bootstrap and Other Resampling Plans. CBMS-NSF Regional Conference Series in Applied Mathematics, Philadelphia.

Stedinger, J. 1983: Confidence intervals for design events. Journal of Hydraulic Engineering, ASCE, 109, 13-27.

Ven Te Chow, 1964: Handbook of Applied Hydrology. McGraw-Hill, New York.

TIME SERIES ANALYSIS IN CLIMATIC SYSTEM MODELS

Masanori B. Okamoto

Section of Statistics, Department of Economics
Hiroshima University
Higashisenda,Naka-ku,Hiroshima,730,Japan

1.INTRODUCTION

Climatic system involves a synthetic system which consist of various physical domains in atmosphere,ocean,cryosphere,lithosphere and biosphere interacting each other with different time scales. Each domain has the proper time scale,which is assumed to have longer scales than life-time of cyclonic disturbances. A good illustration of the climatic system is given in the report of GARP-16 (1975). Various phenomena which is presumed to be related with climatic variations and its characterized time scales are shown as the Table by Kutzbach(1974). From the illustration of climatic system and the Table of time scales,one variation of climatic phenomena with far longer time scales occurs as a result of alteration in the external or boundary condition of the system. Another climatic variation is due to the many interaction within the climatic system itself. Unfortunately we observe different aspect of the same climatic phenomena with different sampling interval and sample size. In addition to this, differences of relaxation time of climatic reaction due to external or boundary change on climatic system make observed results ambiguous. Under such situation it is required to postulate climatic model. For the analysis of climatic time series, it is also required to fix what scale of climatic phenomena to study beforehand.

2. TIME SERIES MODELS FOR CLIMATIC VARIATIONS

After suitable detrending climatic time series, we presume that we are able to have stationary time series. As more general time series model in time domain we have ARMAX model,of which estimation of coefficents are studied by several authors (e.g. Nicoles(1976), Hannan(1976), Hannan(1979)). However, it is not so easy to detect true exogenous variable when treating climatic variation due to interaction whithin system itself. When we include extra-terrestial activities,volcanic activies and artifical activities, it may be possible to use ARMAX which is far from practical use except for model with very low order. Thus,we confine ouselves to the use of AR, ARMA (if necessary,ARIMA). For the practical use of these models, the computer package TIMSAC-78 (include TIMSAC, TIMSAC-74) is very powerful. Details of TIMSAC-78 is repoted by Kitagawa and Akaike(1981).

3. ANALYSIS IN FEEDBACK SYSTEM

In order to construct climatic model, it is reqired to know causal relationship between climatological variables at the first stage. Climatic model is essentially feedback system consist of many variables,say $(x_1(n), x_2(n), \ldots, x_k(n))$ and can be expressed by a discrete time parameter model

$$x_i(n) = \sum_{j=1}^{k} \sum_{m=1}^{M} a_{ij}(m)\, x_j(n-m) + u_i(n)$$

$$i = 1,2,\ldots,k$$

where $[a_{ij}(m)]$ is impulse response function of output $x_i(n)$ for input $x_j(n)$ (Akaike,1967). $a_{ii}(m) = 0$ is assumed here. Unobservable noise $u_i(n)$ cannot be assumed to uncorrelated with or orthogonal to, the input series $x_j(n)$ $(j = 1,2,\ldots,k)$. So ordinary cross-spectral mehtod has no sense for the analysis in feedback system. Akaike's method(1968) for the analysis of linear feedback system is still powerful. Dealing with the probelms in the time domain he get directly estimates of impulse response fuction $[a_{ij}(m)], m = 0, 1,2\ldots$. As the estimation procedure by a physically realizable transformation,without using the future values of the inputs,conceptually whiten the spectrum of the additive disturbance $u_i(n)$ are obtained at first. Then the regression coefficients of the linear predictor of the process are estimated by the least squares method. Finally the estimated predictor are back-transformed into the original form of the system structure. Consistency of estimate by the least squares method bases on the strong assumption of independency of the innovation process of the disturbances(Akaike,1969).

4. CAUSAL RELATIONSHIP BY RPC

The only practical condition to assure the validity of Akaike's estimation procedure is an instantanous return form the output to itself through the feedback loops is prohibited. For the climatic feedback system,this delayed reaction will be more common at least for the monthly averaged data. In Akaike's study (1969) relative power contribution (RPC) $r_{ij}(f)$ of $u_j(n)$ to the

power of $x_i(n)$ at frequency f is given as follows,

$$r_{ij}(f) = \frac{\hat{q}_{ij}(f)}{\hat{p}_{ii}(f)}$$

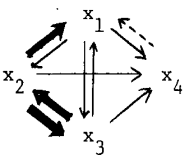
where $\hat{q}_{ij}(f)$ is the power contribution of noise source $u_j(n)$ to the output $x_i(n)$ and $\hat{p}_{ii}(f) = \sum_{j=0}^{k} \hat{q}_{ij}(f)$ is an estimation of the power spectral density at f of $x_i(n)$. Cumulative $R_{ij}(f)$ is also defined by

$$R_{ij}(f) = \sum_{h=1}^{j} r_{ih}(f) \qquad j = 1,2,\ldots,k$$

Feedback occurs if and only if cross-correlation $\rho_{ij}(h)$ between whitened series ε_i and ε_j $(i \neq j)$ is not zero for some $h>0$ and for some $h<0$, when autoregressive form holds among variables $x_i(n)$, $i = 1,2,..,k$ (Pierce and Haugh,1977). As cross-spectrum is Fourier transform of $\rho_{ij}(h)$, RPC can be a good indicator of the existence of feedback arm after suitable whitening noise which drives the feedback system. However, Granger cauality is not synonymous with any of the definitions of causation in the natural science, although many authors (Granger,1969, Pierce and Haugh,1977, Dag TJ STHEIM, 1981, Chamberlain,1982) have been discussed. It is a statement about predictability in the population but contains no reference to the notion of systematic forcing. In a universe governed by causal laws,unidirectional Granger causal ordering may emerge. In a feedback system, if there exists feedback arm acutally, reverse unidirectional causal ordering should appear,because systematic forcing in both direction can produce such orderings. Thus, feedback detected by RPC will have real meaning only when it is supported by some physical law.

5. CLIMATIC CAUSAL RELATIONSHIP:EXAMPLE

We presume a close system which consists of 4 variables,i.e.,monthly means of surface pressure (x_1), surface temperature (x_2), thickness between 850-500mb(x_3), clould amount(x_4). As an example, RPC of noise source $u_2(n)$ of surface temperature to the power of thickness $x_3(n)$ and inversely RPC of noise source $u_3(n)$ of thickness to the power of surface temperature $x_2(n)$ at the stations Hachijojima (33°06'N,139°47'E),and Shionomisaki(33°27'N,135°46' E) are shown in Fig.1. Sampling duration of monthly mean values is AD 1970-1981(N = 144). RPC of $u_2(n)$ to $x_3(n)$ reaches a peak at the periods of about 8 months,while reverse RPC of $u_3(n)$ to $x_2(n)$ extends continously from about 4 months to more longer periods of month. Causal relationship in the presumed system is shown in Fig.2,on which boldfaced arrow indicates distinguished feedback. From Fig.2 we cannot see feedback arm from the variation of



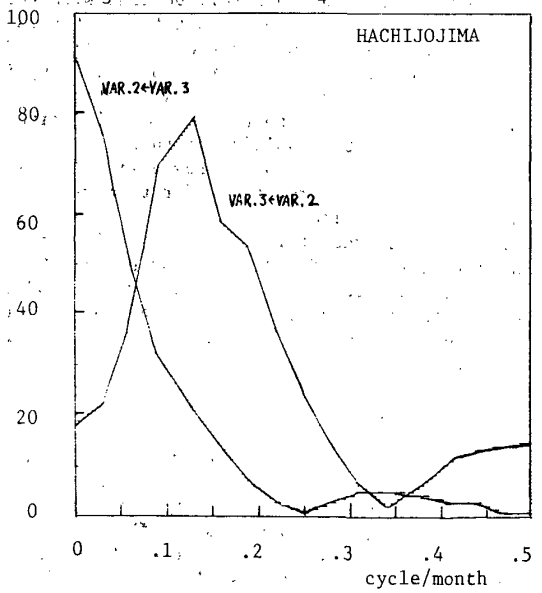
x_1 = surface pressure
x_2 = surface temperature
x_3 = thickness(850-500mb)
x_4 = cloud amount

Fig.2 Causal relationship of monthly means by Relative Power Contribution(RPC).

cloud amount to the variations of the other 3 variables,except for very weak contribution from x_4 to x_1 mainly on shorter periods within 6 months, but can observe feedback loop between variations of variables x_1, x_2 and x_3. Such a feedback loop will be expected by meteorological reasoning. On the other hand, only unidirectional noise contribution from x_2 to x_3 appears at Shionomisaki. The difference between Shionomisaki and Hachijojima suggests that extensive knowledge of RPC on global area are required to get a certain climatological meaning. In fact, Yamamoto et al.(1979) also claimed that it is generally difficult to estimate the global climatic conditions only from the climatic records at stations confined to a limited area.

6. CLIMATIC VARIATION WITH LONGER TIME SCALES

For climatic time series with longer time scale, say 100 to 1000 years,precision of time scale and of sample data are not always sufficiently better.



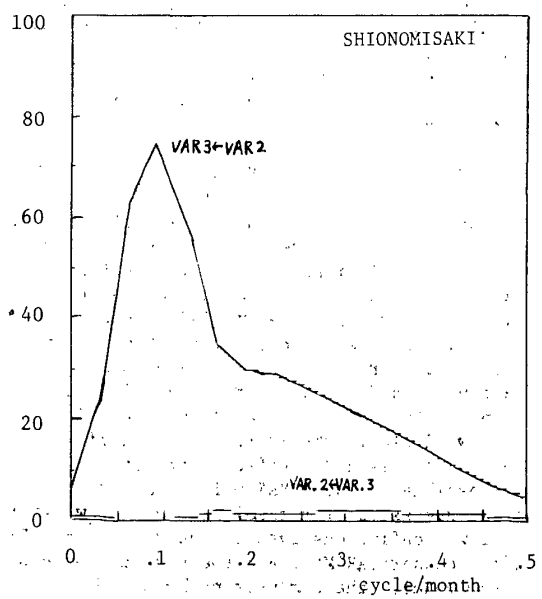


Fig.1 Relative Power Conribution(RPC) at Hachijojima and Shionomisaki (x_2=surf.temp.,x_3=thickness)

For example,determination of time obatained by tree rings is less reliable. In addition to this, some climatic variables are indirectly deduced or estimated by the related other climatic factors. However, air or water temperature, precipitation amount and wind direction has fairly high reliability. Sample size of climatic variable to be study is not always sufficiently large for the fitting time series model such as AR, ARMA. Sometimes we met serious missing observations. Remarkable spectrum estimation can be done with the use of TIMSAC-78. This procedure is due to the method of calculating the exact likelihood function of ARMA based on Akaike's Markovian representation and using Kalman recursion. Jones(1980) extended the recursive calculation procedure to include the cases of missing observations and observational error. The example shows that estimation of spectrum is possible even for 60 missing observation of all 114 observations(Kitagawa and Akaike,1981).

When dealing with multivariate climatic times series, difficulties arise from different sample size among them and different smoothing of them. Fitting ARMA cannot be done straightforward for climatic time series which are often contaminated data.For analysis of climatic data with longer time scale, one of interests is effect of extra-terrestrial activity on climatic variation. Sunspot numbers have been used as an indicator of irradiational activity of the Sun. However the reaction of sunspots to climate is not simple; up to a certain level sunspot level,there may be temperature rise, but then counter oscillation intervenes and reversal occurs (Fairbridge,1961). It is deriable to make a synthetic climatic indicator no matter how reverse oscillation and conatminated nature of climatic time series. Using hard limiting procedure on original vector time series, Okamoto and Iwase (1982) give an synthetic indicator for the representation of gross feature of them.

7. HARD LIMITING PROCEDURE

We consider p-variate climatic time series $Z(t) = [z_1(t), z_2(t),\ldots,z_p(t)]',t = 1,2,\ldots,N$,which is observation of sample size N from stationary Gaussian process $\{Z(t)\}$. In the covariance matrix $\Sigma(h) = [\gamma_{ij}(h)]$; $\gamma_{ij}(h) = E[z_i(t)\, z_j(t+h)]$,only autocovariances $\gamma_{ii}(h)$ is assumed known. By choosing a siutable level u, we get hard limiting series $C_u(z_\alpha(t))$ from the originally obtained time series as follows

$$C_u(z_\alpha(t))=\begin{cases} 1 & \text{for} \quad z_\alpha(t) \geq u \\ 0 & \text{for} \quad |z_\alpha(t)| \leq u \\ -1 & \text{for} \quad z_\alpha(t) < -u \end{cases}$$

for non-negative u and $\alpha = 1,2,\ldots,p,t = 1,2,\ldots,N$. For a fixed α and different two time points of series t and t+h, we can assume that $X=z_\alpha(t)$ and $Y=z_\alpha(t+h)$ follow to a bivariate standard normal distribution $N(0,0,1,1,\rho)$, since the finite value of variance can be reduced to 1 without the loss of generality. By denoting hard limiting series $C_0(z_\alpha(t))$ having the crossing level u=0 by sgn(X),we have

$$C_u(X) = \frac{1}{2}[\mathrm{sgn}(X-u) + \mathrm{sgn}(X+u)].$$

Further, we have the next expression of sgn(X)

$$\mathrm{sgn}(X) = \frac{1}{\pi i}\lim_{\varepsilon\downarrow 0}\lim_{L\to\infty}\int_{L\geq|t|\geq\varepsilon} \frac{1}{t}\exp(itx)dt$$

After some algebra, we have

$$\rho = \frac{\sqrt{2\pi}}{4}\exp(\frac{1}{2}u^2)\{E[X\cdot C_u(Y)]+ E[Y\cdot C_u(X)]\}$$

So we have as estimator $\tilde{\rho}$ of ρ

$$\tilde{\rho} = \frac{\sqrt{2\pi}}{4}\exp(\frac{1}{2}u^2)[X\cdot C_u(Y) + Y\cdot C_u(X)]$$

for a given level u. $\tilde{\rho}$ is unbiased estimator of ρ. $\mathrm{Var}(\tilde{\rho})$ is explicitly given as follows,

$$\mathrm{Var}(\tilde{\rho}) = -\rho^2 + \frac{\pi}{4}\exp(u^2) + \frac{\sqrt{2\pi}}{4}\rho^2 u\exp(\frac{1}{2}u^2)$$
$$-\frac{\pi}{2}\exp(u^2)\int_0^u \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)\,dx$$
$$+\frac{1}{4}\frac{1+\rho^2}{\sqrt{1-\rho^2}}\exp(u^2)\{\exp(-\frac{u^2}{1+\rho})+\exp(-\frac{u^2}{1-\rho})\}$$
$$-\frac{1}{4}\rho\exp(u^2)\int_{-\rho}^{\rho}\frac{(1+x)(1+x^2)-2u^2(1-x)(1+2x)^2}{(1-x^2)^{3/2}(1+x)}$$
$$\times\exp(-\frac{u^2}{1+x})\,dx$$

for $|\rho|<1$, $u\geq 0$ and for the case when $|\rho|=1$ and $u \geq 0$

$$\mathrm{Var}(\tilde{\rho}) = -1 + \sqrt{\frac{\pi}{2}}u\exp(\frac{1}{2}u^2)$$
$$+\pi\exp(u^2)\int_u^\infty \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)\,dx$$

A numerical diagram of $\mathrm{Var}(\tilde{\rho})$ against ρ for given levels u is shown in Fig.3, from which $\mathrm{var}(\tilde{\rho})$ is near constant 1.12 for the level u=5/8 ~2/3,although $\mathrm{var}(\tilde{\rho})$ attains the minimum value of about 0.68 at $\rho = 1.0$ for the level u=0. Thus it is desirable to

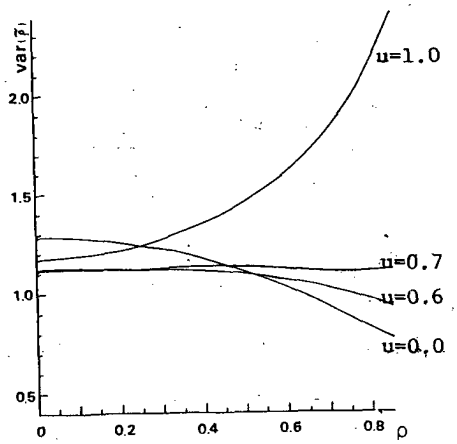


Fig.3 $\mathrm{Var}(\tilde{\rho})$ against ρ for the crossing level u = 0.0, 0.6,0.7 and 1.0

adopt u=2/3 which gives nearly uniform variance of estimator $\tilde{\rho}$ for all ρ if there are no information about ρ.

Further autocovariance $E[C_u(X)C_u(Y)]$ of hard limiting series is given for $-1<\rho<1$ as follows

$$E[C_u(X)\, C_u(Y)] = \frac{1}{\pi}\int_{-\rho}^{\rho}\frac{1}{\sqrt{1-\rho^2}}\exp(-\frac{u^2}{1+x})dx,$$

from which we obtain arcsin-law

$$E[\mathrm{sgn}(X)\cdot\mathrm{sgn}(Y)] = \frac{2}{\pi}\mathrm{Sin}^{-1}\rho$$

as the special case for u = 0.

8. SYNTHETIC CLIMATIC INDEX SERIES

For the expression of gross feature over several climatic time series, we adopt simply the total sum of hard limiting series $T_p(t)$;

$$T_p(t) = \sum_{\alpha=1}^{p} C_u(z_\alpha(t)) \qquad t=1,2,\ldots,N$$

For the analysis of climatic time series we used the collected series ranging over A.D. 1200-1900 given by Kai(1979). Among them,we take seven time series which are suitable to analyze;

(1) 10 years average of Wolf's sunspot numbers (Takahashi, 1975)
(2) Width of tree-rings in U.S.A.(Schove, 1978)
(3) Width of tree-rings in Germany(Schove,1978)
(4) Concentration of 14C in the atmosphere(Lerman,1970)
(5) 20 years average of an index of good harvest estimated by difference of tree-rings from the one of the standard growth rate(Suda, 1951)
(6) Isotope ratio of oxygen in tree-rings of Yakushima,Japan(Pandolfi et al.,1976)
(7) Isotope ratio of oxygen in tree-rings of Yakushima,Japan (Libby et al.,1976)

These time series are reproduced in Fig.3. The other time series had the very short period of observation or resulted in strong smoothing. All values of time series are read off each five years from the printed curves in the paper of Kai(1979) and $(z_\alpha(t))$ (here p=7,N=139) are obtained after substraction of each value from its mean. Index series $T_p(t)$ are computed for the three different crossing levels,u=0,2s/3 and s where s is the standard deviation.

If an absolute value of $T_p(t)$ is equal to the total number of different time series,then there will be a complete concordance between them for each t. However,we admit a weak concordance among $z_\alpha(t)$ for such a case when

$$p \geqq |T_p(t)| \geqq p/2$$

By this criterion we can see that a warm period was around A.D. 1570;a cold period was from A.D. 1680 to 1700. The latter cold period corresponds to the Maunder minimum.

In Fig.4 the index series $(T_p(t))$ for u=2s/3 is compared with the changes in time of eustasy and cosmic radiation(tree-rings 14C) given by Fairbridge(1961),and also sunspot numbers averge for 10 years given by Takahashi(1979),which are not included in the above seven time series. There reasonable match between cosmic radiation and eustasy. The index series is more similar to variation of cosmic radiation, although Fairbridge(1961) indicates a striking similarity between eustasy and sunspot curve smoothed to 100-year mean.

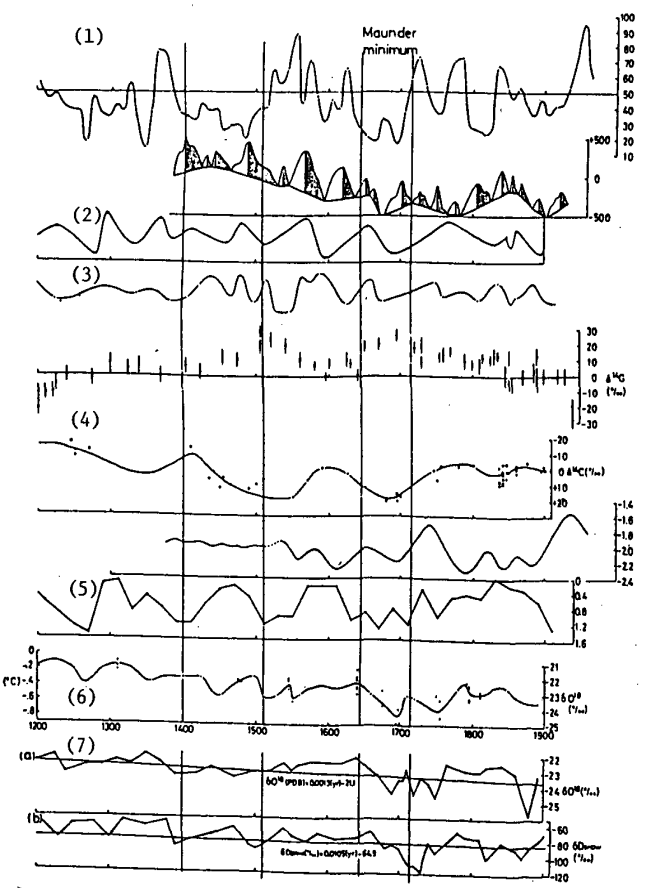


Fig.3 Climatic time series (1)-(7)(After K.Kai 1979). Sources of (1) -(7) are given in in the text.

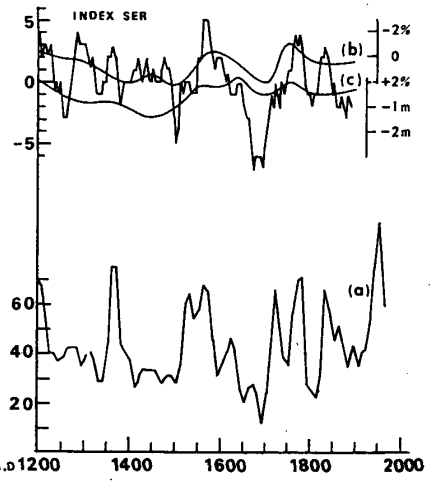


Fig.4 Climatic indes series $(T_p(t))$ from A.D. 1200 to 1900 for the crossing level u=2s/3. (a) 10-yrs average sunspot numbers,Takahashi (1979),(b)cosmic radiation (tree-ring 14C), Fairbridge(1961) (c) Eustasy, Fairbridge(1961).

Autocovariance of index series $(T_p(t))$ will be decomposed into two parts of autocovariances and crosscovariances of clipped series $(C_u(z_\alpha(t))$;

$$E[T_p(t)\ T_p(t+h)] = \sum_{\alpha=1}^{p} E[C_u(z_\alpha(t))\ C_u(z_\alpha\ (t+h))]$$

$$+ \sum_{\alpha\neq\beta} E[C_u(z_\alpha(t))\ C_u(z_\beta\ (t+h))]$$

Under the assumption of bivariate normal distribution $N(0,0,1,1,\rho_{ij})(i,j=\alpha,\beta)$ of a pair of original series, we have

$$E[C_u(z_\alpha(t))C_u(z_\alpha(t+h)] = \Psi(\rho_{\alpha\alpha}(h))$$

and

$$E[C_u(z_\alpha(t))C_u(z_\beta(t+h)] = \Psi(\rho_{\alpha\beta}(h))$$

where

$$\Psi(\rho_{ij}(h;u)) = \frac{1}{\pi}\int_{-\rho}^{\rho} \frac{1}{\sqrt{1-x^2}} \exp\left(-\frac{u^2}{1+x}\right) dx$$

for $i,j = \alpha,\beta$. Thus we have

$$E[T_p(t)\ T_p(t+h)] = \sum_{\alpha} \Psi(\rho_{\alpha\alpha}(h)) + \sum_{\alpha\neq\beta} \Psi(\rho_{\alpha\beta}(h))$$

Estimator of $E[T_p(t)\ T_p(t+h)]$ is simply given as follows,

$$\gamma_T(h;u) = \frac{1}{N}\sum_{t=1}^{N-h} T_p(t)\ T_p(t+h)$$

$$= \gamma_c(h;\alpha) + \gamma_c(h;\alpha,\beta)$$

where

$$\gamma_c(h;\alpha) = \frac{1}{N}\sum_{t=1}^{N-h} C_u(z_\alpha(t))\ C_u(z_\alpha(t+h))$$

and

$$\gamma_c(h;\alpha,\beta) = \frac{1}{N}\sum_{t=1}^{N-h} C_u(z_\alpha(t))C_u(z_\beta(t+h)) \quad .$$

These γ_c's are estimators of $\Psi(\rho_{ij}(h))$,$i,j = \alpha,\beta$. $\Psi(\tilde{\rho}_{ij}(h))$ will be near $\gamma_c(h;\alpha)$ and $\gamma_c(h;\alpha,\beta)$ if estimator $\tilde{\rho}_{ij}$ is a good estimator of ρ_{ij}. Usual estimator $\tilde{\rho}$ of ρ is given by

$$\tilde{\rho} = \frac{1}{N}\sum_{t=1}^{N-h} z_\alpha(t)z_\alpha(t+h) \quad ,$$

but we adopt estimator $\tilde{\rho}_{ij}$ which is given by

$$\tilde{\rho}_{ij}(h;u) = \frac{N}{2(N-h)} \times$$

$$\frac{\sum_{t=1}^{N} [z_i(t)C_u(z_j(t+h)) + z_j(t+h)C_u(z_i(t))]}{\sum_{t=1}^{N} z_i(t)\ C_u(z_i(t))}$$

The latter estimator is not unbised;in fact this estimator includes Huzii's estimator (1963) as the special case for u=0. However, difference of $\gamma_c(h;\alpha) - \Psi(\tilde{\rho}_{\alpha\alpha}(h;u))$ extends from -0.06 to 0.07 as far as the present data are concerned. $\gamma_c(h;\alpha)$ is very near to the theoretical integral value as

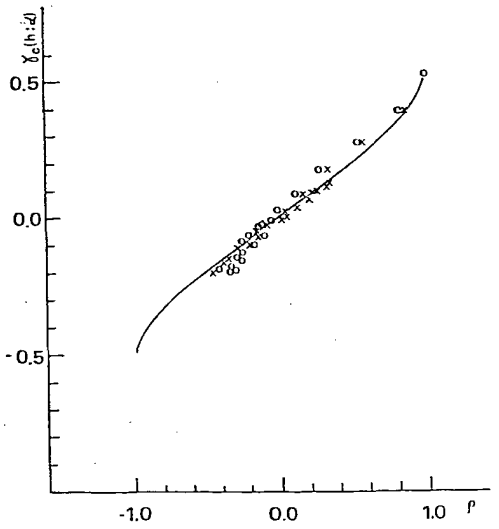


Fig.5 Estimated autocorrelation $\gamma_c(h;\alpha)$ of hard limiting series and the theoretical integral $\Psi(\rho_{\alpha\alpha}(h))$. $\Psi(\tilde{\rho}_{\alpha\alpha}(h;u))$ with × and $\Psi(\tilde{\rho}(h))$ with o are plotted against $\tilde{\rho}$.

a function of ρ and $\Psi(\tilde{\rho}_{\alpha\alpha}(h;u))$or $\Psi(\tilde{\rho}(h))$ are plotted as a function of $\tilde{\rho}$ as shown in Fig.5. $\Psi(\tilde{\rho}_{\alpha\alpha}(h;u)$ is less biased than $\Psi(\tilde{\rho}(h))$ against the theoretical integral value $\Psi(\rho(h))$. Statistical properties of these estimators are still unknown.

9. CONCLUSION

A recent developd technique of time series analysis can be applied for climatic system modelling. However, choice of analysis method depends entirely on the time scale adopted on the climatic study. For the climatic variation due to interaction within climatic system itself, but mainly due to self-adjustment of atmosphere, time scale extends from months to several years. It is still desirable to use Akaike's RPC (relative power contribution) when confirming existence of feedback arm which is presumed by climatological reasoning or physical law,while Okubo(1982) has already applied RPC for the analysis of causality among financial variables. In this procedure it is required to check whether the assumption of randomness among whited noise of the fitted AR model holds or not for the climatic variables which are actually used for the computation of RPC.

For the climatic variation of more longer time scale,say 100 to 1000 years, available climatic time series are less reliable and conatminated in general. However, we have some of them observed in various area in the world. The situation is to deal with multivariate time series, in which each series is often contaminated data. Avoiding fitting the known vector time series model such as AR,ARMA, a synthetic climatic index series are proposed by hard limiting procedure. The index series is defiend as the total sum of each hard limiting series from original climatic time series. The variation in time of this index series seems to reflect change extending from solar activity to eustasy. A weak concordance given by certain creiterion on the index series indicates the distingushed climatological cold periods(the Maunder Minimum). In statistical analysis of index series, we have unbiased estimator of population correlation under the assumption that a pair of sample follows to bivariate normal distribution with known mean and variance. This unbiased estimator of ρ can be computed by the use of a pair of original and its clipped series. An optimum level for the estimation of autocorrelation is given as about 2/3 of the standard deviation by the criterion of minimum variance of estimator for the case of unknown true autocorrelation.

10. REFERENCES

Akaike, H., 1967: Some problem in application of the cross-spectral method , in "Spectral analysis of time series", ed. by B. Harris,John Wiley, New York,81-107.

_____, 1968: On the use of linear model for the identification of feedback systems. Ann. Inst. Statist. Math., 20,425-439.

_____, 1969: A method od statistical identification of discrete time parameter linear system. Ann. Inst. Statist. Math.,21,225-242.

Chamberlain, G., 1982: The general equivalence of Granger and Sims causality, Econometrica, 50, 569-581.

Fairbridge, R.W., 1961: Convergence of evidence of climatic changes and ice ages. Ann. New

York Academy of Sci., 95, Art 1, 543-579.
GARP-16 1975: The physical basis of climate and climatic modeling. WMO-ICSU Joint Organization Comm., GARP Pub. Ser. No. 16, Geneva.
Granger, C.W.J.., 1969: Investigating causal relations by econometric models and cross-spectral methods, Econometrica, 37, 424-438.
Hannan, E.J., 1976: The identification and parametrization of armax and state space models, Econometrica, 44, 713-722.
_____, 1979: The ststistical theory of linear systems ,in "developments in stsatistics", vol. 2, ed. by P.R. Krishnaiah, Academic Press, New York, 83-121.
Huzii, M., 1963: On a simplified method of the estimation of the correlogram for stationary Gaussian process. Ann. Inst. Statist. Math., 14, 259-269.
Jones, R.H., 1980: Maximum likelihood fitting of ARMA models to time series with missing observations, Technometrics, 22, 389-395.
Kai,k., 1979: Collection of curves of climatic variation since A.D.1200 (I),in " Studies on Climatic Variation (in Japanes)". Report of Climatology and Meteorology. Institute of Geoscience, Univ. Tsukuba (4),24-37.
Kitagawa, G. and Akaike, H., 1981: On TIMSAC-78 , in " Applied time series analysis II", ed. by D.F. Findley, Academic Press, 499-547.
Kutzbach, J.E., 1974: Fluctuations of climatic monitoring and modelling. WMO Bulletin 23, 47-54.
Nicoles, E.J., 1976 : The efficient estimation of vector linear time series models, Biometrika, 63, 381-390.
Okamoto, M.B., and Iwase, K., 1982: A statistical index representing a gross feature of several climatic time series by hard limiting. J. Meteor. Soc. Japan, 60, 726-738.
Okubo, T., 1982: Mony supply and causal relationship among economic variables (in Japanese). Finacial Studies, Financial Studies Institute, Bank of Japan.
Pierce, D.A. and Haugh, L.D., 1977: Causality in temporal systems:Characterization and a survey. J. Econometrics, 5, 265-293.
TJØSTHEIM, Dag., 1981: Granger-causality in multiple time series, J. Econometrics, 17, 157-176.
Yamamotō,R., Hoshiai, M., and Iwashima, T., 1977: Change of surface air temperature averaged globally during the years 1957-1972. Arch. Met. Geoph. Biokl., Ser. B, 25, 105-115.

HIGH-RESOLUTION FREQUENCY ANALYSIS AND APPLICATIONS TO THE PREDICTION OF ALMOST-PERIODIC FUNCTIONS

G. W. Brier, M. M. Siddiqui, M. W. Haurwitz and R. Biondini

Colorado State University
Fort Collins, Colorado, U. S. A.

1. INTRODUCTION

This paper reports on the development of two new methods for the statistical analysis of time series showing quasi-periodicities or other characteristics of almost-periodic (AP) functions. Applications to geophysical and climatic time series indicate that these methods can provide a greater understanding of the complex variations in many series and possibly lead to statistical prediction procedures more effective than those in current use. The first method is a High-resolution Frequency Analysis (HRFA) which possesses the two very desirable properties of high-resolution and consistency. It has been successfully used to identify the line frequencies in a sample AP time series representing the continuous solution of a well known set of differential equations. When applied to series composed of known periodic signals with superposed random noise, the technique provides detailed information that would be missed by autocorrelation or Fast-Fourier-Transform methods.

The second technique is concerned with analysis and prediction of time series responding in a non-linear fashion to multiple periodic forcings where the periods $p_1, p_2 \ldots p_m$ are known or specified a priori. The One Parameter Multiplicative Model (OPMM) makes no assumptions regarding the waveform, amplitude or phase of the periodic forcings or responses, and no estimation of the usual Fourier coefficients is required. The statistical forecasting methods in general use are usually based on models that assume the relationship between response variable y and independent variables $x_i (i=1,2,\ldots m)$ is linear in unknown paramaters. The OPMM is a new approach to the problem which considers a general multiplicative form $z=f_1(x_1)\, f_2(x_2) \ldots f_m(x_m) + K + \varepsilon$, where the response z of the system depends upon m variables, ε is a random error and K is some unknown systematic (constant) error, the one parameter to be estimated from the data. Prediction equations are derived for the extrapolation (forward or backward) of cyclical phenomena with non-sinusoidal waveforms and for AP functions where the data are not adequate for the resolution of line frequencies that are close together. When periods $p_1, p_2, \ldots p_m$ cannot be specified a priori from physical considerations, the dominant frequencies determined by HRFA can be used in the prediction equations. Application to real data and computer simulations on artificial series indicate that the methods are more robust than the usual autoregressive schemes, with less danger of over fitting.

2. METHODS OF ANALYSIS

2.1 High-resolution Frequency Analysis (HRFA)

Consider a (complex valued) function of time x(t) which we shall refer to as "signal" at time t. The quantity $|x(t)|^2$ is its instantaneous power and $|x(t)|^2 dt$ its instantaneous energy. If

$$\int_{-\infty}^{\infty} |x(t)|^2 dt < \infty , \qquad (1)$$

the signal has finite energy and hence is a transient signal which fades out. On the other hand, if

$$|x(t)|^2 dt = +\infty \text{ but } \lim_{T\to\infty} \frac{1}{T} \int_{-T}^{T} |x(t)|^2 dt < \infty \qquad (2)$$

the signal has infinite energy but finite average power. It does not fade out. AP functions are shown in the mathematical literature as bounded solutions of linear differential equations and can be approximated by trigonometric polynomials of finite order. Solar activity, whether measured by sunspot numbers or by solar flux in a certain frequency band, belongs to AP signals. Similarly, insolation (the amount of solar radiation reaching at the top of the earth's atmosphere) and its effect on our ocean-air-land system, belongs to AP signals.

In the statistical literature, AP functions are treated under the "hidden periodicities" model

$$x(t) = \mu(t) + \varepsilon(t) \qquad (3)$$

$\mu(t) = \alpha_o + \sum_{j=1}^{k} (\alpha_j \cos 2\pi f_j t + \beta_j \sin 2\pi f_j t)$, $f_1; \ldots f_k$, where k, $f_1, \ldots, f_k, \alpha_o, (\alpha_j, \beta_j)$ j = 1, ... k are all unknown constants and $\varepsilon(t)$ is a second order stationary process.

The problem we are dealing with consists in analyzing a sample from such a process x(t), $-\infty < t < \infty$. Spectral density estimates which are currently used in such a situation are consistent but lack high resolution. A method has been developed (Siddiqui and Izenman, 1981; Siddiqui and Wang, 1983) for scanning the frequency interval

$o<f<1/2$, which possesses the desirable properties of high resolution and consistency. The technique is derived from a theorem (see Doob, 1953) on the inversion of Fourier-Stieltjes transform

$$C(f+1/2\Delta f)-C(f-1/2\Delta f)=\lim_{T\to\infty}\int_{-T}^{T} e^{-2\pi itf}\,\frac{\sin\pi t\Delta f}{\pi t}\,X(t)dt \qquad (4)$$

$$=\lim_{N\to\infty}\sum_{m=-N}^{N} e^{-2\pi imf}\,\frac{\sin\pi m\Delta f}{\pi m}\,X(m).$$

We illustrate the method here by an analysis of the time series of the orbital parameters of the earth's motion and with an application to some long-term climatic records. We gain some insight regarding the validity and stability of the method by applying it to samples of different lengths from time series with known periodic components and with varying amounts of superposed random noise.

2.2 One Parameter Multiplicative Models

Consider the case of m periodic functions $x_1(t),x_2(t),\ldots x_m(t)$ with periods $p_1, p_2 \ldots p_m$, $\bar{x}_i=0$, where an additive signal $y(t)$ is composed of the superposition of the m signals, i.e., $y(t)= x_1(t)+ x_2(t)+\ldots x_m(t) + K$, where K is some unknown constant. The $x_i(t)$ may be sinusoidal, but the exact waveform is not relevant to the arguments here. For the case of two periodic signals, p_1 and p_2, it is easy to show (Brier, 1979) that

$$y(t) = y(t-n_1p_1) + y(t-n_2p_2)$$

$$-y(t-n_1p_1 - n_2p_2) \qquad (5)$$

(where n_1 and n_2 are any integer), regardless of the waveform, amplitude or phase of the components $x_i(t)$. Thus, if p_1 and p_2 are known and an observed time series $y(t)$ is available, then it is possible to obtain estimates $\hat{y}(t)$ at any time providing sufficient data are available. It is not necessary for the periods p_i to be integers, or even commensurable. No estimates of the Fourier coefficients α_i and β_i (see Eq. 3)are needed if the objective is prediction or extrapolation. In passing, we note that (5) also represents the extrapolation of a linear trend, so that no detrending is necessary.

The concepts can be extended to m periods or to the multiplicative model (OPMM) without difficulty. For the multiplicative case the response variable is defined as

$$z(t)=x_1(t)x_2(t)\ldots x_m(t) + K \qquad (6)$$

and the appropriate equations for m=2 are

$$z'(t)=\frac{z'(t-n_1p_1)\;z'(t-n_2p_2)}{z'(t-n_1p_1-n_2p_2)}, \qquad (7)$$

where $z'(t)=z(t)-K$, $z'(t-n_1p_1)=z(t-n_1p_1)-K$, etc.

The importance of estimating the parameter K has been stressed by Brier (1979) and Brier and Kelbe (1980). Fortunately, estimates of K can be obtained easily for m=2, and,with a little more trouble, for additional periods. OPMM cannot be transformed to a linear model by taking logarithms of the data. The additive unknown parameter, K, is an essential part of the model which prevents linearizing via logarithms. In fact, the magnitude of the estimated K can be used to discriminate between an additive and a multiplicative model. If K becomes very large, an additive model is indicated. In the applications presented here we confine our analysis to the additive model for reasons of computational simplicity, although it is possible that the multiplicative model might give better results. In most applications it appears to be desirable to compute both the additive and the multiplicative estimates, since the same terms are involved. Then the user can choose the model that makes the "best" predictions.

The general concept of the multiplicative model is not limited to time series applications but can be extended to the analysis of a number of factors in a design matrix and to some types of multivariate problems. Details are discussed in Brier and Kelbe (1980).

3. APPLICATIONS

3.1 Analysis of signal with known periods

One way of evaluating the techniques developed here is to test them on a data set where the signal is known. An artificial series was constructed where the signal $y(t)$ was a composite of four periodic signals with periods 17, 18, 19 and 22 units. Figure 1 shows a plot of these four cycles. They are not sinusoids and the cycle with the period of 22 units is not symmetrical, having a waveform something like the sunspot curve. If one knew that $y(t)$ was composed of four cycles with periods 17, 18, 19 and 22, it would be possible to extrapolate $y(t)$ indefinitely on the basis of any initial sequence of 76 values of $y(t)$. No knowledge of the phase amplitude, waveform or frequency decomposition of the four component periodic functions would be necessary. If only 76 values of $y(t)$ are available, the simplest prediction equation ($n_1=n_2=n_3=n_4=1$) would be

$$y(t)=y(t-17)+y(t-18)+y(t-19)+y(t-22)-y(t-35)-y(t-36)$$

$$- 2y(t-37)-y(t-39)-y(t-40)-y(t-41) \qquad (8)$$

$$+ y(t-54)+y(t-57)+y(t-58)+y(t-59)-y(t-76)$$

If there is no noise on the system, the predictions will be exact. If there is noise on the system the accuracy of the predictions will, of course, depend upon the signal to noise ratio.

The signal $y(t)$ will be periodic with a cycle of 63,954 units, the least common multiple of 17, 18, 19 and 22. With this many values a Fourier analysis would show all the power concentrated at the four specified frequencies and their harmonics. Spectrum or harmonic analysis on short record spans will show apparent instabilities from one part of the record to another. In this example, we will take the point of view that a typical time series record will have only a few hundred,

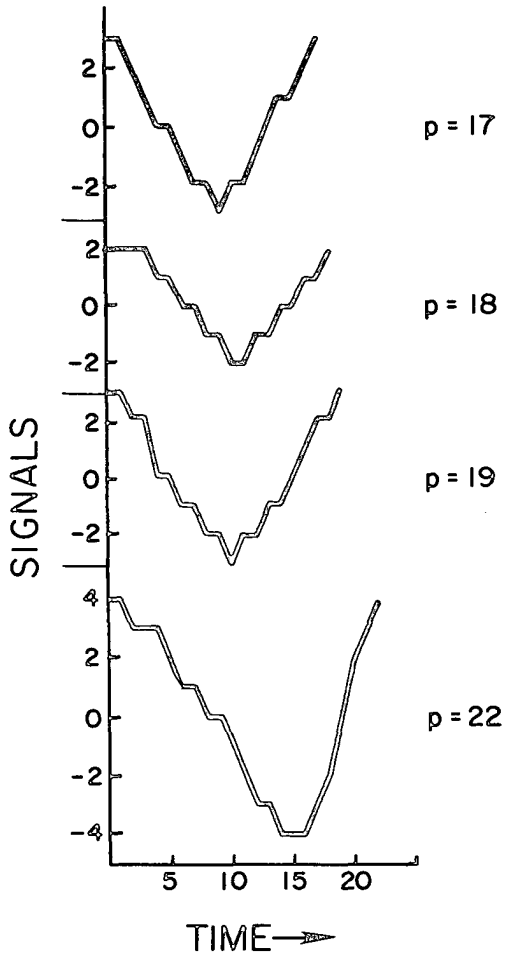


Fig. 1. Waveforms for the periodic signal components of time series y(t).

or a thousand observations at most. Thus, for all practical purposes the y(t) series given here will behave like an AP function.

However, we now ask the question, "Suppose one did not know a priori the underlying generating frequencies, how many observations would be necessary to obtain reasonable estimates of them and how would these be used for prediction purposes?" In an attempt to answer these questions, we have applied HRFA to y(t) for samples of various lengths before time t=0. The results of the four analyses involving 149, 301, 451, and 601 observations are shown in Table 1. Figures 2 and 3 are the graphs for 301 and 601.

One notes that the separate peak around a period of 18 units shows up only when n=601 although some evidence of power at that frequency is indicated when n=301 since the trough between the periods at 17 and 19 does not dip as low as the other troughs which are obviously in the noise level. But even n=601 is a small number compared with the repetition period of 63,954!

Table 1. Periods detected by HRFA in artificial time series with true periods 17, 18, 19 and 22 for different sample sizes. The relative amplitudes are shown in parentheses.

	True Periods			
N	17	18	19	22
149	16.367 (0.1292)		18.832 (0.1175)	22.173 (0.2310)
301	17.036 (0.3212)	17.889 (0.11139) on residual scan	18.692 (0.3584)	21.976 (0.4372)
451	16.920 (0.4152)		18.832 (0.3990)	21.978 (0.6697)
601	17.036 (0.06390)	18.018 (0.4623)	18.975 (0.6201)	21.978 (0.8504)

Our HRFA technique includes an iterative program to examine the residuals of a time series after it has been fitted by the line frequencies detected by the initial fine scan. For the case of 301 observations, the fine scan was made on the residuals after the original series was fitted with cycles corresponding to the three largest peaks shown in Table 1. This resulted in the detection of a major peak at a period of 17.889, very close to the period of 18 which was introduced. Further iterations could be made to determine whether improved estimates could be made of the location of the 4 lines.

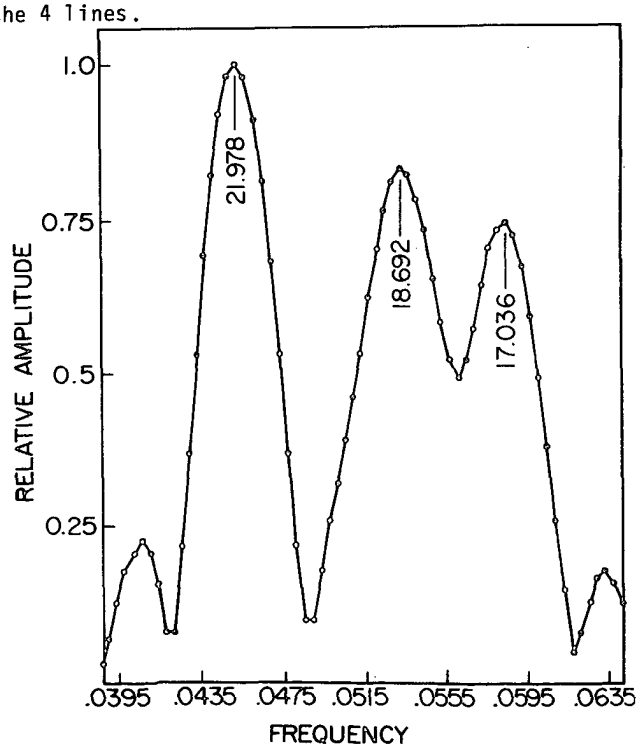


Fig. 2. Result of HRFA scan for 301 observations of y(t).

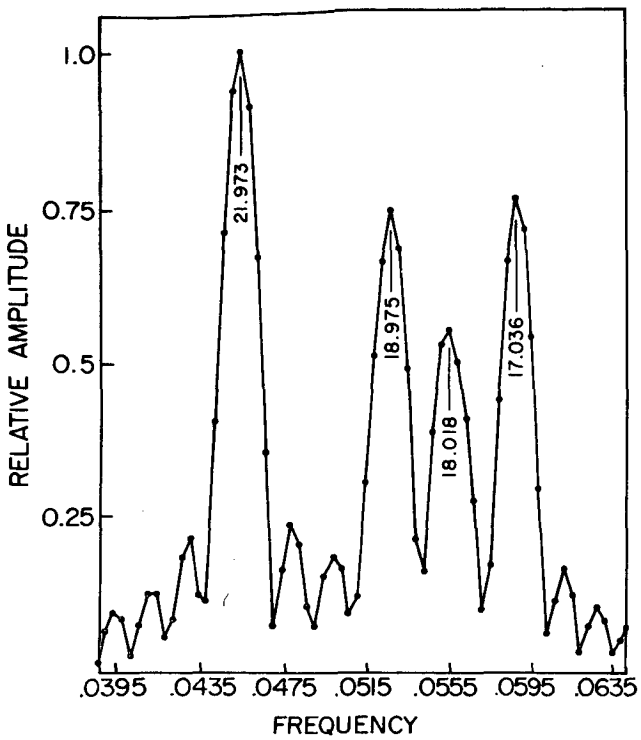


Fig. 3. Result of HRFA scan for 601 observations of $y(t)$.

The next step was to determine the effect of superposing random noise on the signal. Random normal deviates were generated for the case of 301 observations, with $\sigma^2(n)$=1, 4, 9, 36 and 100. The original signal had a standard deviation of 4. The results of using HRFA on these noisy signals is shown in Table 2. Only when $\sigma^2(n)$ approached 100 did there seem to be an appreciable effect on the location of the major peaks. The graph of the fine scan for $\sigma^2(n)$=100 is shown in Figure 4. Three of the 5 highest peaks are still "real" and two are "spurious." The effect of spurious peaks on the accuracy of statistical prediction schemes will be discussed later.

To illustrate the application of OPMM to predictions based on frequencies detected by HRFA we refer to the results from the sample of 301 observations. The peaks around 17 and 19 can be interpreted as evidence of spectral power in that general neighborhood that cannot be resolved with the data sample available. If predictions are to be made for less than one or two cycles in advance it can be shown that an average based on the frequencies indicated by two close peaks will do very well. Therefore, we shall use 21.978 and 17.864 = (18.692 + 17.036)/2 as the basic periods. Also, to illustrate another point, we put in a "spurious" period of 14 which does not bias the OPMM predictions, and would have no undesirable effect when there is no noise on the system. In contrast, extrapolation of data by Fourier analysis would perpetuate indefinitely the errors resulting from inclusion of a spurious period.

Using these periods, the appropriate prediction equation is

Table 2. The effect of noise on the HRFA estimates of frequency for N=301. The relative amplitudes are shown in parentheses.

Noise	True Periods				
$\sigma^2(n)$	17	18	19	22	"SPURIOUS"
0	17.036 (0.3220)	(No peaks showing)	18.692 (0.3584)	21.978 (0.4372)	
1	17.036 (0.3156)		18.692 (0.3504)	21.978 (0.4444)	
4	17.036 (0.3100)		18.692 (0.3427)	21.978 (0.4516)	
9	17.036 (0.3044)		18.692 (0.3354)	26.978 (0.4588)	
36	16.921 (0.2941)		18.692 (0.3158)	21.978 (0.4806)	
100	16.807 (0.2782)		18.553 (0.29739)	21.978 (0.5098)	15.552 (0.2941) 14.641 (0.2870)

$$\hat{y}(t) = y(t-14) + 2(t-17.864) + y(t-21.978) - 2y(t-31.864) - y(t-35.728) - y(t-35.978) - 2y(t-39.842) + y(t-49.728) + 2y(t-53.842) + y(t-57.706) - y(t-71.706). \quad (9)$$

In application of (9) one can interpolate between data points or round the lag values to the nearest integer so that 31.864 becomes 32, e.g.

Trial predictions using (9) were made for 200 data points beyond the initial sample of 301 observations. When no noise was added to the signal, the correlation between observed and predicted values was 0.99. When noise $\sigma^2(n)$=4 was added to the signal the correlation was 0.64. The effect

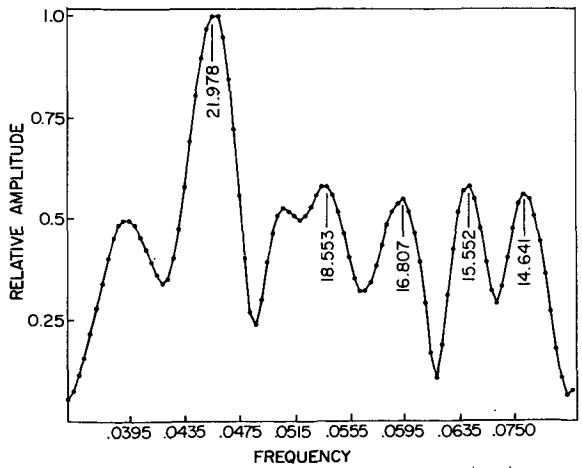


Fig. 4. Result of HRFA scan for 301 observations of $y(t)$, with noise introduced.

of noise on the predictions was reduced by smoothing with a 7 term binomial weighting formula. Thus, the right hand terms of (9) became $\tilde{y}(t-1)$ where 1 represents the lag and $\tilde{y}$ is a smoothed input value. Of course, in this smoothing procedure the high frequency content of the signal is attenuated as well. Figure 5 shows a graphical comparison of the observed and predicted values for the first 36 points. It should be noted that each prediction is based on information contained in values from 14 to 72 units previously. There is no fitting or estimation of regression coefficients. If all of the peaks shown by HRFA are spurious, there should be no correlation between the predicted and observed values.

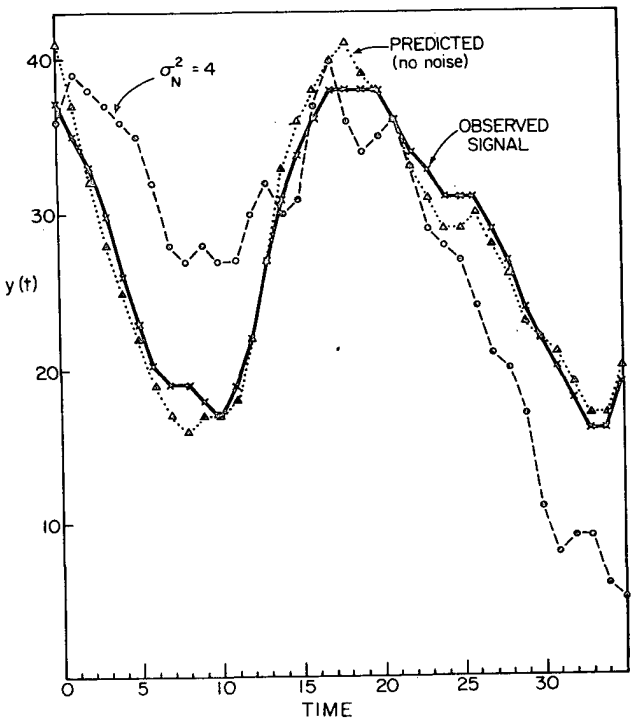


Fig. 5. Comparison of observed and predicted smoothed values of y(t), with and without the introduction of noise.

3.2 Astronomical Theory of Climate Change (ATCC)

In recent years a number of investigations have been reported that provide evidence supporting the ATCC. Quasi-periods found in deep sea cores appear to correspond quite closely with those determined by the mathematical-physical theory of orbital variations. Figure 6 presents a time series plot of the variations of eccentricity, tilt, and precession for the period from 4,000,000 BP to the present time. These curves are good examples of the AP functions discussed earlier -- no section of a curve will exactly reproduce another section of the curve although there may be some similarities as indicated by a simple inspection. To illustrate these points, we have performed a HRFA on the data for three sections of the orbital parameter time series involving time spans of 294,000 years specified as follows:

Span A - ending at present,
Span B - ending 1.6 millions years BP,
Span C - ending 2.95 million years ago.

(The data were taken at 2,000 year intervals for the 294,000 years in order to correspond with the isotopic ice volume record supplied to us by Professor John Imbrie). The orbital data were generated by the main terms of the trigonometric series given by Berger (1977).

Table 3 shows the frequencies and (relative) amplitudes of the three highest peaks found by the HRFA for time intervals A, B, and C. Only the results for eccentricity are shown, but similar variations among A, B, and C were found for the other orbital parameters. Note that changes take place in both frequency and amplitude as the data span changes. Of course, this is exactly what would be expected with AP functions. One should not interpret these variations as due to sampling or noise resulting from a stochastic process. The signal is deterministic and the commonly used "significance" tests have no meaning.

Table 3. Frequency information for major lines determined by HRFA in the eccentricity parameter for three different time intervals A, B, and C. Each interval spans 294,000 years. The periods (in years) are in parentheses.

Interval	Rank 1 Freq.	Rank 2 Freq.	Rank 3 Freq.	Rank 4 Freq.
A	0.0175 (114,286)	0.0313 (63,898)	0.0383 (52,219)	0.0451 (44.346)
B	0.0051 (392,157)	0.0185 (108,108)	0.0307 (65,147)	0.0379 (52,770)
C	0.0183 (109,290)	0.0283 (70,671)	0.0351 (56,980)	0.0417 (47,962)

A HRFA was next performed on the isotopic ice volume record represented by the time series of $\delta^{18}O$ for the last 294,000 years at 2000 year intervals. To compare with the orbital variations, it is necessary to perform the HRFA on the same 294,000 years. Although the usual spectral or harmonic analysis could have been used, it is interesting to note the additional detail given by the HRFA. The results are shown in Table 4. We think the agreement is remarkable, especially among the shorter periods. For the precession period near 23,000 years, the disagreement is less than 1/4 percent, which is much closer than one would expect because of dating errors in the sea core data. Low resolution spectrum analyses could not be expected to provide such detailed information since they are designed for a different purpose and usually assume that there are no lines or narrow bands in the spectrum.

3.3 Analyses of the Drought-Area Index (DAI).

Evidence has been found of a 22-year periodicity in large-scale drought in the western United States from a 360-year tree ring reconstruction of a drought area index (Mitchell et al, 1979). A link was indicated between the drought-area index (DAI) and the Hale or Double Sunspot cycle, but

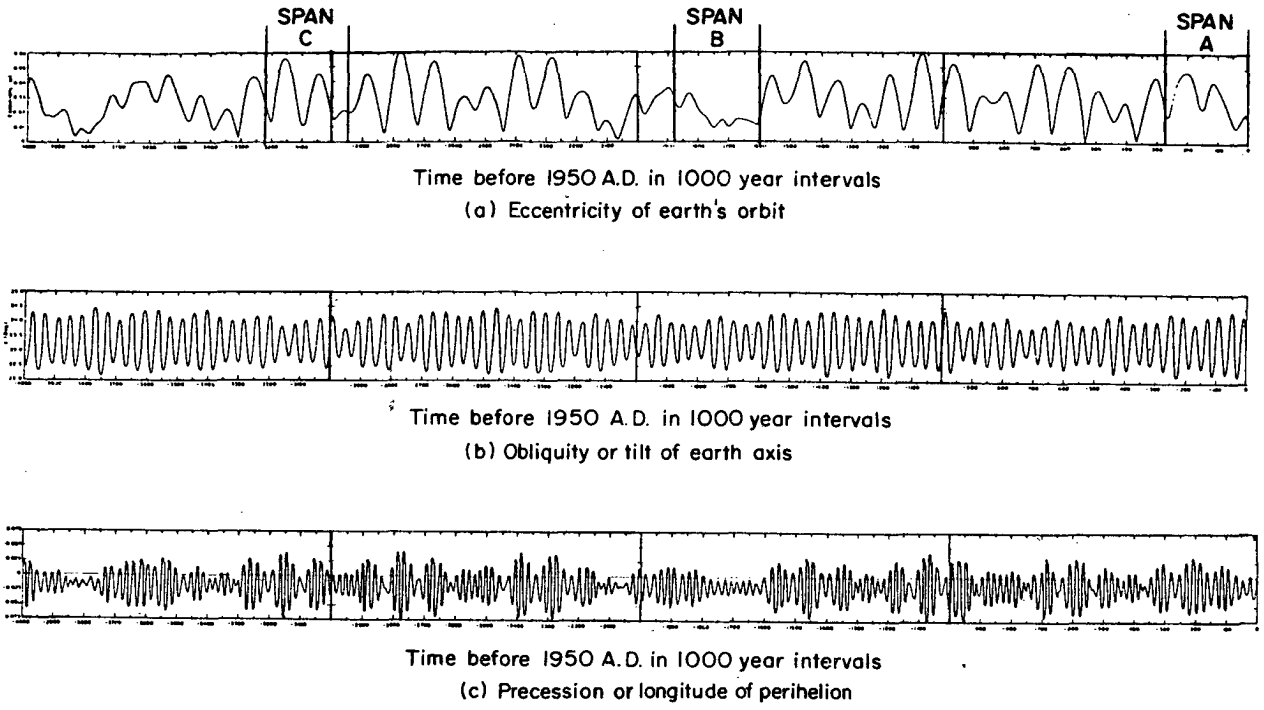


Fig. 6. Variations of eccentricity, tilt, and precession as a function of time for the period from 4,000,000 B.P. to the present time (from Vernekar, 1972).

there were appreciable variations in amplitude and phase of the DAI with respect to solar cycle. Currie (1981) suggested that the drought variations were a reflection of the 18.6-year lunar nodical forcing on the atmosphere while Bell (1981a, 1981 b) suggested a combination of solar and lunar influence. Since the reports discussed above were based on low resolution spectrum analyses, it seemed appropriate to examine the DAI series using HRFA. The data were supplied by Murray Mitchell.

The series investigated here covers the period 1600-1962 and was identified as "F-65(PDSI<-1)" in Mitchell, et al (1979). Figure 7 shows the result using HRFA for the frequency range from 1/25 cycle per year to 1/16.3 cycle per year. The peak at 22.026 years is the largest found in the entire frequency range from 0.0002 to the Nyquist frequency 0.5000. A most interesting feature of Fig. 7 is the indication that the spectral power around 18.6 years is split into two lines, which could result if the nodical period of 18.613 years was modulated by some low frequency component. Low resolution only spectra would show a broad peak around 18.6 years, perhaps merging into the peak at 22 years. Other peaks in the HRFA scan did not show sharpness comparable to the ones shown here.

Table 5 gives summary data on the ten largest peaks found by HRFA. There are a number of interesting observations that can be made about the periods shown. Many of them are harmonics or subharmonics of other periods, e.g.:

7.000 x 17.857 = 124.999 years
6.500 x 19.231 = 125.002 years
6.000 x 37.879 = 227.274 years
2(1/19.231 - 1/17.857) = .0080 (125 years)
3.000 x 18,613/2 = 27.933 years

It appears rather unlikely that all of these relationships are accidental results of sampling from a purely stochastic process. However, there appears to be no obvious connection between the period of 22.026 and the others.

4. SUMMARY AND CONCLUSIONS

We have presented here the results of some of the research leading to the development of statistical methods for analysis and prediction of geophysical and climatic time series. Although further analysis is needed, tests performed on real or artificial series with known signals show advantages over many commonly used procedures. The HRFA applied to the earth's orbital parameters and isotopic ice volume data provide detailed and convincing evidence in support of the ATCC in agreement with the findings of other investigators Application of HRFA to drought data for western United States shows evidence for a cycle around 22 years in combination with peaks at 19.231 and 17.857 years, which may be interpreted as indicating a modulated cycle around 18.6 years. Other significant periods indicated by the analysis of the DAI appear to be harmonics or subharmonics of 18.6 years. Prediction equations are presented for the extrapolation of AP functions and for

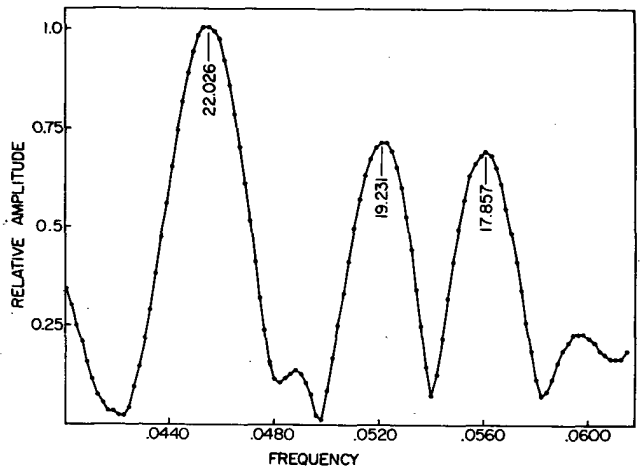


Fig. 7. Results of HRFA scan of frequencies corresponding to periods around 25 years to around 16 years.

Table 4. Periods determined by HRFA from the last 294,000 years of the isotopic ice volume record ($\delta^{18}O$) and the orbital parameter precession (prec.), obliquity (obl.), and eccentricity (ecc.).

Frequency (cycles/ 2000 yrs.)	Period (years)	Relative Amplitude
0.1062	18,832	0.0098 (prec.)
0.1057	18,921	0.0892 ($\delta^{18}O$)
0.0864	23,148	0.0299 (prec.)
0.0862	23,202	0.1473 ($\delta^{18}O$)
0.0692	28,902	0.0680 ($\delta^{18}O$)
0.0665	30,075	0.13109 (obl.)
0.0513	38,986	0.1691 ($\delta^{18}O$)
0.0492	40,650	0.9299 (obl.)
0.0391	51,151	0.15636 (obl.)
0.0352	56,818	0.1212 ($\delta^{18}O$)
0.0290	68,966	0.1430 ($\delta^{18}O$)
0.0186	107,527	0.3970 ($\delta^{18}O$)
0.0170	117,647	0.009446 (ecc.)
0.0065	307,692	0.009445 (ecc.)

cyclical phenomena with non-sinusoidal waveforms. Frequency lines detected by HRFA can be used for the OPMM predictions when physical theory does not specify the periods.

Standard tests are available for assessing the significance of frequency lines detected by HRFA, but the distinction should be made between significance testing and estimation. The value of HRFA appears to be its precision in estimating the location of a frequency line. Also, confusion often exists between fitting and prediction. OPMM is not a fitting procedure, but is directed at the prediction problem with minimum loss of degrees of freedom due to estimating a multiplicity of coefficients. In our opinion, the real test of a prediction procedure is how well it performs on independent data. We feel that both HRFA and OPMM hold up well in this respect.

Table 5. HRFA preliminary selection of significant frequencies for the DAI series. Frequencies are ordered by decreasing magnitude of amplitude. Notice that the combined power of the periods around 18.6 years is essentially the same as that for the 22 year period.

Rank	Frequency	Period	Amplitude	% Var.
1.	.0454	22.026	0.284	8.02
2.	.0264	37.879	0.241	5.80
3.	.0080	125.000	0.226	5.09
4.	.0520	19.231	0.204	7.95
5.	.0560	17.857	0.196	
6.	.0678	14.749	0.170	2.87
7.	.0126	79.365	0.163	2.65
8.	.1330	7.519	0.162	2.63
9.	.0044	227.273	0.157	2.46
10.	.0358	27.933	0.148	1.96

5. ACKNOWLEDGEMENTS

This research was supported in part by the National Science Foundation under Grant ATM-7911465 and by the National Oceanic and Atmospheric Administration under Grant No. NOAA-NA81AA-D-00039. We wish to thank Professor John Imbrie for supplying the isotopic ice volume data.

6. REFERENCES

Bell, P. R., 1981a: Predominant periods in the time series of drought area index for the western High Plains, AD 1700 to 1962. NASA Conference Publication 2191: Variations of the Solar Constant, S. Sofia (ed.), p. 257-264.

Bell, P. R., 1981b: The combined solar and tidal influence in climate. NASA Conference Publication 2191: Variations of the Solar Constant, S. Sofia (ed.) p. 241-256.

Berger, A. L., 1977: Support for the astronomical theory of climate change. Nature, 269, 44-45.

Brier, G. W., 1979: Use of difference equation methods for predicting sunspot numbers, Solar-Terrestrial Influence on Weather and Climate, B. M. McCormac and T. A. Seliga (eds.), D. Reidel Publishing Company, Dordrecht, Holland, 209-214.

Brier, G. W. and B. Kelbe, 1980: Statistical Prediction of Complex Phenomena Using a One Parameter Multiplicative Model (OPMM). Preprint -- WMO Symposium on Probabilistic and Statistical Methods in Weather Forecasting, Nice, France, 8-12 September, 1980. Pp. 85-92.

Currie, R. G., 1981: Evidence for 18.6 year M_n signal in temperature and drought conditions in North America since A.D. 1800. J. of Geophysical Res. 86 (C11)11055-11064.

Doob, J. L., 1953: Stochastic Processes. Wiley

Mitchell, J. M., Jr., C. W. Stockton and D. M. Meko, 1979: Evidence of a 22-yr. rhythm of drought in the western United States related to the Hale solar cycle since the 17th Century. 125-143 in B. M. McCormac and T. A. Seliga (eds). Solar Terrestrial Influences on Weather and Climate, D. Reidel Pub., Dordrecht.

Siddiqui, M. M. and A. J.Izenman, 1981: High-resolution frequency analysis of the monthly sunspot numbers. Submitted to Journal American Statistical Assoc.

Siddiqui, M. M. and C. C. Wang, 1983: Statistical Analysis of Almost-Periodic Random Functions. Preprints for First Saudi Symposium on Statistics and its Applications, Riyadh, Saudi Arabia, May 2-5, 1983.

Vernekar, A. D. 1972: Long-period global variations of incoming solar radiation. Meterological Monographs 12, No. 34.

SPECTRAL AUTO VARIANCE, CROSS VARIANCE AND COHERENCE ANALYSIS OF LONG AIR TEMPERATURE SERIES

Christian-Dietrich Schönwiese

Institute for Meteorology and Geophysics
J.W. Goethe University, D-6000 Frankfurt 1 (F R G)

1. INTRODUCTION

The air temperature near surface can be conceived to be a leading element used for definition and consideration of climate. Based on any thermal regime and particular ciculation patterns involved the characteristics of other climatic elements (e.g. precipitation) may be deterministically or statistically derived. From the statistical point of view and related to the recent modern period from where direct measurements are available (last few centuries) additional arguments for temperature emphasis are: Relatively accurate measurements (also in earlier times), longest time series available (since 1659 AD Central England record; Manley, 1974; since 1579 AD northern hemisphere assesments; Groveman and Landsberg, 1979), relatively small spatial variability and therefore relatively fair spatial representance.

Since a long time scientific and public people is concerned with the well-known fact of climate variability. In contrast to this fact, the time-space related details of this variability are only fragmentary known and there are a lot of speculations about the reasons. Easy to understand that the question of cyclical and non-cyclical variability components of climatic elements which can be answered by means of spectral statistics is of particular interest: There may arise some suggestions for climate forcing and the very problematic topic of climate predictability, in case of cyclical components, may be discussed.

In the following some remarks on long air temperature series spectral statistics are summarized, where a few examples from all over the world are put together in order to present preliminary results. In the first instance, the climate definition is a statistical one, where the statistics reflect the facts of climate behaviour. These statistics

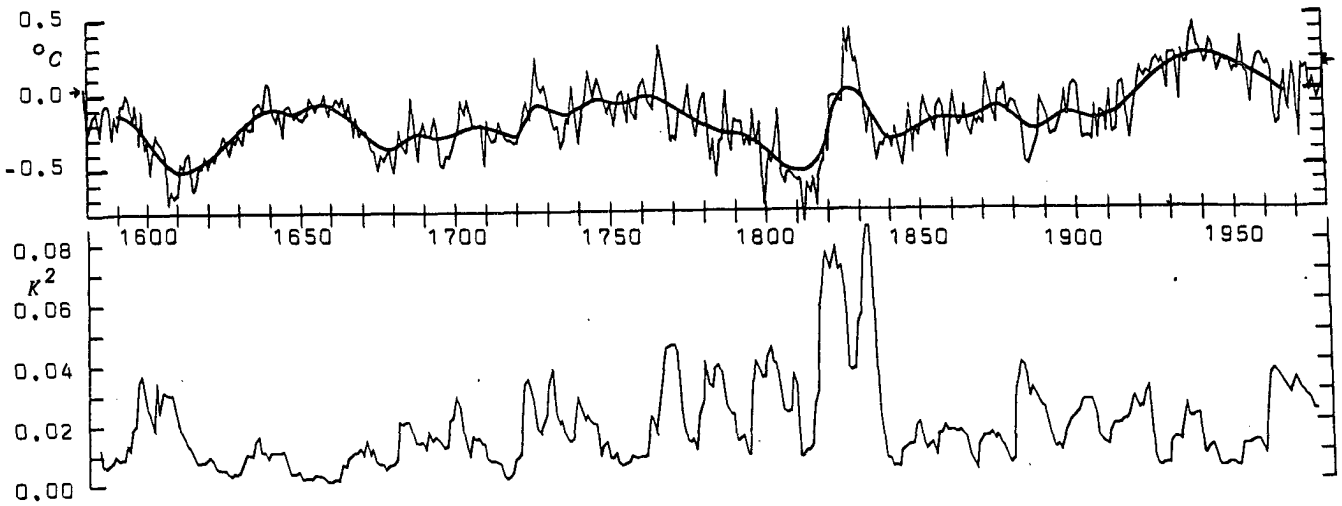


Fig. 1. Northern hemisphere yearly mean temperature assessments 1579-1980 after Groveman and Landsberg (1979), upper plot, including low-pass filtered data (Gaussian filter T > 30 a, T=period, a=years); in addition, 10 year moving variance, lower plot.

may serve to verify deterministic models, appraise physical theories and trigger further investigations in the statistical and non-statistical domain. (For climate forcing studies see Lamb, 1972,1977; Berger et. al.,1981; Schönwiese, 1983; and others.)

2. TIME SERIES STATISTICS

Fig. 1 (preceding page) presents an example of a climatic time series on the basis of yearly mean data: Northern hemipshere temperature (near surface) assessments after Groveman and Landsberg, 1979, where the 1881-1975 data are taken from Borzenkowa et al., 1976, 1976-1979 from Lauscher, 1981, and 1980 adopted after the Jones et al., 1980, data set. (See also Yamamoto, 1980.) The spatial scope (hemispheric) may both avoid regional (or local) randomness and provide a data basis for comparisons with regional (or local) data.

Note the three aspects of this figure: (1) interval (yearly) averages (first moment); (2) low-pass filtered interval averages (Gaussian low-pass filter suppressing periods $T \leq 30$ a; see e.g. Mitchell et al., 1966; Schönwiese, 1978); (3) variance moving interval (here 10 a) values (second central moment) used to derive another time series characterizing the original data set.

Both the low-pass filter and the moving variance aspect (in addition, of course, higher moments) not only describe the time series but also point to possible non-stationarity: There is a evident long-term variability component (see low-pass filtered data) causing time-dependent fluctuations of the average; there are also time-dependent fluctuations of variance (note the enormously increased level of variance c. 1815-1835).

Because overall (whole interval) trends of average and variance are unlike in this case, the time series considered here can be supposed to be quasi-stationary (although after c. 1920 the low-pass filtered component reached a level as never before in the interval considered). Other non-spectral statistics as well as stationarity problems shall not be discussed here.

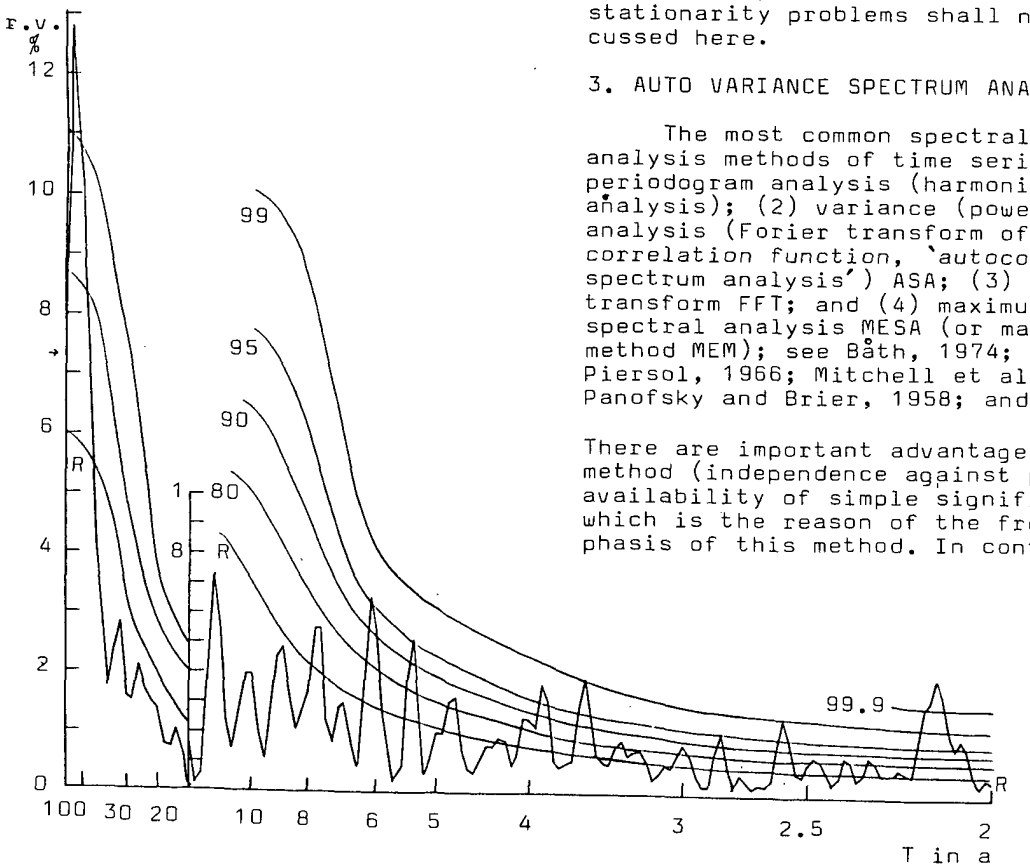


Fig. 2. Northern hemisphere yearly mean temperature (compare Fig. 1) variance spectrum; sample size n=402, maximum time lag M=150, one time lag autocorrelation coefficient r_1=0.77; abscissa period T in years (=a), ordinate relative variance r.v. in percent. R means theoretical Markov red noise spectrum. The curce numbers are values of significance in percent (significance $p = 1-\alpha$, where α=error probability).

3. AUTO VARIANCE SPECTRUM ANALYSIS

The most common spectral statistic analysis methods of time series are: (1) periodogram analysis (harmonic Fourier analysis); (2) variance (power) spectrum analysis (Forier transform of the autocorrelation function, 'autocorrelation spectrum analysis') ASA; (3) fast Forier transform FFT; and (4) maximum entropy spectral analysis MESA (or maximum entropy method MEM); see Båth, 1974; Bendat and Piersol, 1966; Mitchell et al., 1966; Panofsky and Brier, 1958; and others.

There are important advantages of the ASA method (independence against phase shifts, availability of simple significance tests) which is the reason of the frequent emphasis of this method. In contrast to

cross spectrum analysis, where two time series are analysed, ASA is a 'auto' variance spectrum method. In recent time MESA is also often performed because it allows a more detailed spectral disintegration.

In Fig. 2 the northern hemisphere temperature (1579-1980) variance spectrum estimate is plotted using the ASA method.

$$R = \bar{s}^2 \frac{1-r_1^2}{1+r_1^2-2r_1\cos(\pi k/M)} \qquad (1)$$

is the theoretical 'red spectrum' simulating the superposition of random(white) noise and persistency following the Markov model (background variance, population estimate). $\bar{s}^2$ is the averaged (white) sample spectrum, M the maximum time lag of autocorrelation, k/M (k=0,1,...,M) the relative frequency and r_1 the one lag autocorrelation coefficient. The sample peak confidence levels are tested against randomness by means of

$$\frac{\chi^2(p,\varphi)}{\varphi}, \quad \varphi = \frac{2N-M/2}{M}, \qquad (2)$$

where $p=1-\alpha$ is the significance (α = error probability), φ are the degrees of freedom and N is the sample size.

It can be seen in Fig. 2 that R increases enormously with period T (r_1 = 0.77) and that some spectrum peaks exceed the 90% level of significance (see also Tab. 1): 300, 6.1, 5.4, 3.9, 3.6, 2.8, 2.6, 2,13 a (a=years); similar results are obtained by MESA. Note that the width of the period intervals varies (H= 2M+1 spectrum estimates, H=0 → T=∞ (long-term residuum); H=1 → T=2M/H=300 a, if M= 150; H=2 → T=150 a; etc.) which means varying accuracy of the dominating periods (e.g. 300 a → ∞ ↔ 150 a; but 2.13 ± 0.01 a).

In addition to the northern hemisphere data set, Tab. 1 summarizes the results of ASA (variance spectrum analysis) in respect to some other temperature time series. There are the following problems: (1) Different sample size which complicates comparisons. (In Tab. 1, therefore, the common interval 1857-1980 is additionally used). (2) Questionable validity of the Markov red noise theory (that means, it is not certain that the relation $r_h = r_1^h$, $h > 1$, presupposition of formula (1) is valid for the background spectrum. (3) Question whether r_1 (sample) is a good estimate of the related ϱ_1 statistic. R should be some type of regression curve in respect to the sample spectrum. In case of significant deviations not only r_1 but also r_h, $h > 1$, should be included in the estimation of R. It is important, however, to include as few as possible sample statistics for population estimates because the former may be a random sample effect.

There are many results of variance spectrum analysis available in the literature, although the results are far from being complete in the climatological do-

Table 1. Some results of variance spectrum analysis based on yearly mean temperature series of the time intervals denoted (M = maximum time lag). The numbers in parantheses are significance levels (p=1-α) in percent exceeded by sample variance spectrum peaks (dominant variance components).

station/region	interval	M	periods of dominant variance components in years
northern hemisphere[1]	1579-1980	150	~300(90), 6.1(95), 5.4(95), 3.9(95), 3.6(99), 2.8(95), 2.6(99), 2.13(99.9)
	1857-1980	50	~100(95), 4.2(80), 3.2(80), 2.1(80)
Philadelphia[2]	1738-1980	100	~200(99),~50(90), 3.2(90), 2.7(90), 2.02(95)
	1857-1980	50	~100(99), 3.4(80), 3.2(80), 2.2(80)
Stykkisholmur[3]	1857-1980	50	~100(95), 3.7(99)
Central England[4]	1659-1980	100	~200(99),~25(90),~14(80), 5(80), 3.4(80), 3.1(90), 2.15(95)
	1857-1980	50	~100(90), 5(80), 3.0(80), 2.3(80), 2.1(95)
Hohenpeissenberg[5]	1781-1980	100	~200(95),~13(90), 8(80), 5(80), 3.4(95), 2.3(95), 2.2(95), 2.1(80)
	1857-1980	50	~100(90), 4.5(80), 3.2(80), 2.6(90), 2.1(90)
Tokyo	1876-1980	50	~100(90), 4.2(90), 2.6(95), 2.2(99)
Cape Towne	1857-1980	50	~100(95), 2.7(90), 2.5(80), 2.4(80)
New Zealand[6]	1857-1980	50	~100(80), 6(80), 3.4(95)

[1] see Fig.1 [2] after *Landsberg(1980)* as tabled in *Lamb(1977)*, updated [3] Iceland
[4] after *Manley(1974)*, updated [5] Germany (989 m, near Alps) [6] after *Salinger(1980)*

main. The main features of the spectra considered here (see also Schönwiese, 1974, 1978a, 1978b, 1980, 1980a, and others) are: (1) 'Red' background spectra but distinctly different r_1 values (which demands differenr R estimates; examples: Central England 1659-1980, M=100, r_1=0.26; Philadelphia 1738-1980, M=100, r_1=0.34; New Zealand 1857-1980, M=50, r_1=0.57; compare also Tab. 1). (2) Some similar significant sample spectrum peaks, particularly near the long-term residuum (roghly T > 30 a or >50 a) and in the interval 2.1-2.3 a (quasi-biennial oscillation, with the exception of New Zealand), in addition c. 2.6-2.8, c. 3.1-3.4, c.5-6, and less significant c. 8, c.15 and c. 25 a.

4. CROSS VARIANCE SPECTRUM ANALYSIS

When the cross correlation function is applied instead of the autocorrelation, the principles of cross spectrum analysis can be outlined as follows: Let $c_{ab}(\tau)$ be the cross correlation function, where a and b are two time series and b is lagged in relation to a ($\tau = k\Delta t$=time lag; k=0,1,...,M), and $c_{ab}(\tau)$ the corresponding function, where a is lagged in relation to b, then

$$E(k) = \frac{1}{2}(c_{ab}(k) + c_{ab}(k)) \qquad (3)$$

is the even and

$$O(k) = \frac{1}{2}(c_{ab}(k) - c_{ab}(k)) \qquad (4)$$

is the odd part of the cross variance function. Applying the Fourier transform on E(k) one gets the cospectrum Co(f), f= 1/T=frequency, and applying this transform on O(k) one gets the quadrature spectrum Qu(f). Similar to an auto variance spectrum, Co(f) and Qu(f) describe the spectrally disintegrated variance, but now the uniform variance of two time series is considered, where Co(f) registers variance in phase (phase angle 0^0, positive spectrum values) or antiphase (180^0, negative values), respectively. In a corresponding way Qu(f) is related to 90^0 (positive) and 270^0 phase angles. The term

$$Cr(f) = Co(f) - i\,Qu(f) \qquad (5)$$

is called the cross spectrum.

In this study only two applications of cross spectrum analysis shall be entered: The relation

$$\frac{Qu(f)}{Co(f)} = \tan\frac{\pi\tau f}{M} \qquad (6)$$

enables the evaluation of the phase spectrum of two time series and

$$Ch(f) = \frac{Qu^2(f)+Co^2(f)}{S_a(f)\cdot S_b(f)} \qquad (7)$$

is the squared coherence; that means, Ch(f) realizes the spectrally disintegrated squared correlation coefficient r^2 of the two time series analysed, where $S_a(f)$ and $S_b(f)$ are the auto variance spectra. Coherence spectrum peaks can be tested against non-correlation by

$$CL = \sqrt{1-\alpha^{1/(\varphi-1)}} \qquad (8)$$

(see e.g. Panofsky and Brier, 1958), where again α is the error probability and φ are the degrees of freedom (computed by means of formula (2)). Coherence spectrum peaks exceeding particular CL levels point to significant ($p=1-\alpha$) correlation within the specified period interval.

In Fig. 3 some examples of coherence analysis are shown; see also Tab. 2. In all these analyses the northern hemisphere temperature (see Fig. 1) is defined to be the basis time series and some local or regional series are compared with. In respect to the interval 1738-1980 it can be seen that the coherence between northern hemisphere and Central England temperature is very significant (p > 99%) in the long-term residuum (T > 100 a) and significant (p > 95%) in the period intervals c. 3.6 and 2.2 a. In case of Philadelphia the peak at c. 7 a is dominating and the coherence T > 100, c. 14, 3.1 and 2.3 a is less significant (p > 90%). In respect to Tokyo and New Zealand there is no significant coherence in the long-term part of the spectrum but well-established, however, at c. 3 a in relation to Tokyo and at c. 20 a in relation to New Zealand. It can be read from Tab. 2 that similar periods dominate as in case of auto spectrum analysis.

Two examples of phase spectra are plotted in Fig. 4, where the outcome is that in the period range T ≤ 8 a the phases are ≤1 a but up to ~2 a (Central England) or ~5 a (Philadelphia), respectively, in the T > 8 a or >30 a range. Roughly the phases increase with period.

5. OUTLOOK

The examples of spectral statistics of climatic time series presented here are not more than preliminary and fragmentary. The studies are continued by support of the 'Deutsche Forschungsgemeinschft' (German Science Foundation) and it is hoped that also the European Community will support these efforts. Long temperature and precipitation series are emphasized hitherto and more detailed results will be published elsewhere. It is planned to extend these investigations also on pressure and circulation parameters. Moreover, statistical forcing studies are on the way (see Schönwiese, 1983) also supported by the 'Deutsche Forschungsgemeinschaft'.

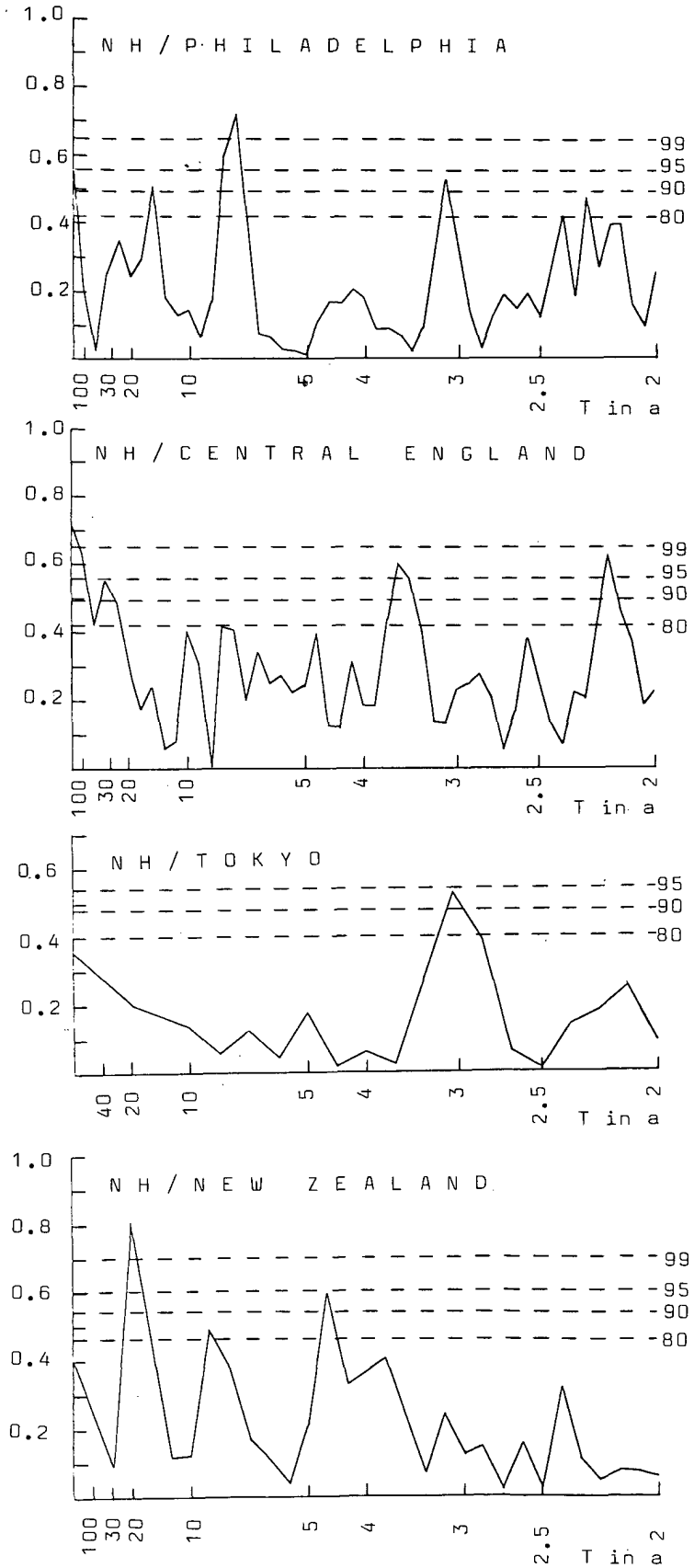


Fig. 3.

Coherence of yearly mean temperature series, where the northern hemisphere data set (see Fig. 1) is correlated with the time series (stations or regions) denoted. NH = northern hemipshere, other symbols see Fig. 2. The right-hand numbers indicate the significance ($p=1-\alpha$) of the coherence spectrum peaks in percent. The time intervals and maximum time lags M are: NH/Central England and NH/Philadelphia 1738-1980, M = 50; NH/New Zealand 1857-1980, M = 30; NH/Tokyo 1876-1980, M = 20.

Table 2. Some results of coherence analysis based on yearly mean temperature series. All series denoted (station/region) are compared with the northern hemisphere data set, see Fig. 1. Symbols and notes same as in Tab. 1; see also Fig. 3.

station/region	interval	M	periods of significant coherence in years (p)
Philadelphia	1738-1980	50	100(95), 14(90), 7(99), 3.1(90), 2.3(80)
	1857-1980	50	100(99), 12(95), 10(90), 6(80), 2.7(95), 2.1(99)
Stykkisholmur	1857-1980	50	100(99), 5(90), 4.3(95), 2.4(80)
Central England	1659-1980	100	100(95), 33(95), 8(99), 6(90), 3.5(99), 2.6/2.2/2.1(95)*
	1857-1980	50	100(99), 25(99), 3.4(80), 2.6(99)
Hohenpeissenberg	1857-1980	50	100(99), 17(95), 7(95), 4(95), 3.6(80)
Tokyo	1876-1980	20	3(90)
Cape Towne	1857-1980	50	100(99), 9(80), 7(80), 3.3(99), 3.0(80), 2.7(95), 2.4(80)
New Zealand	1857-1980	30	20(99), 9(80), 5(95)

*) without p=80% significance level

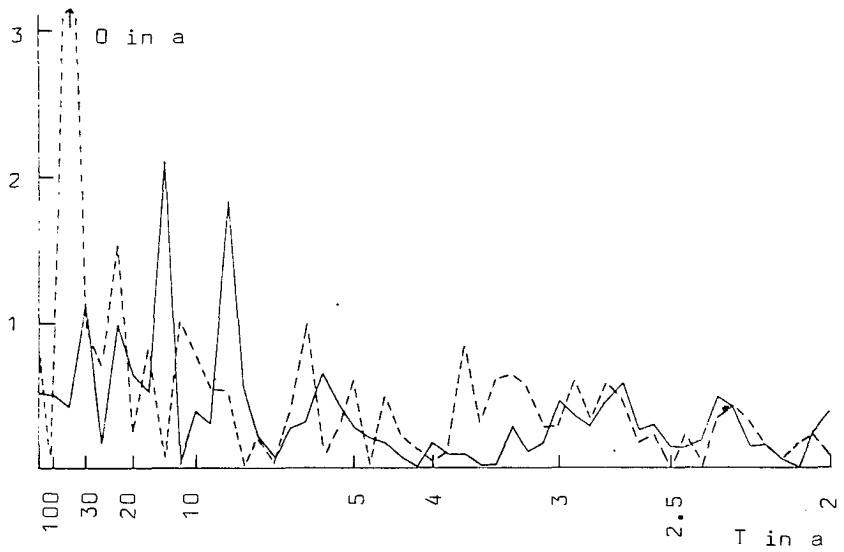


Fig. 4.

Phase spectra between northern hemisphere (see Fig. 1) and Central England, solid line, or Philadelphia, dashed line, respectively. Data basis: yearly mean temperature series 1738-1980. The arrow at T ≈50 a is related to the pase 5.4 a. Most phases are below 1 a but three peaks (Central England T=8 and 12 a; Philadelphia T ≈25 a) point to phases in the order of 2 a and one peak (Philadelphia T ≈ 50 a) exceeds even 5 a.

6. REFERENCES

Båth,M.,1974: Spectral analysis in geophysics. Elsevier, Amsterdam etc.

Bendat,J.S.,and A.G.Piersol,1966: Measurement and analysis of random data. Wiley, New York.

Berger,A.(ed.),1981: Climatic variations and variability: Facts and theories. Reidel, Dordrecht etc.

Borzenkova,I.I.,et al.,1976: Met. Gidrol., 7, 27-35.(In Russian.)

Groveman,B.S.,and H.E.Landsberg,1979: Univ. of Maryland, Coll. Park, Publ. No. 79-181/2.

Lamb,H.H.,1972/1977: Climate: Present, past and future. Methuen, London.

Lauscher,F.,1981: Zentralanstalt Meteorol. Geodynamik, Wien(Vienna), Publ. No. 252.

Manley,G.,1974: Quart. J. R.Met.Soc.,100, 389-405.

Mitchell,J.M.,et al.,1966: WMO Publ. No. 195 (Tech. Note No. 79), Geneva.

Panofsky,H.A.,and G.W.Brier,1958: Some applications of statistics to meteorology. Pennsylvania State Univ., University Park.

Schönwiese,C.D.,1974: Mitt. Meteorol. Inst. München No. 24. - *1978:* Meteorol. Rdsch., 31, 73-84. - *1978a:* Arch. Met. Geoph. Biokl.,Ser. B,26, 1-16. - *1978b:* Contr. Atmosph. Phys. (Beitr. Phys. Atm.), 51, 139-152. - *1980:* Meteorol. Magazine, 109, 101-113. - *1980a:* Meteorol. Rdsch., 33, 167-174. - *1983:* Arch. Met. Geoph. Biokl., in print (Northern hemisphere temperature statistics and forcing. Part A: 1881-1980 AD; Part B in prep.).

Yamamoto,R.,1980: In *Ikeda,S.,et al. (eds.):* Statistical Climatology. Elsevier, Amsterdam etc., pp. 307-324.

In addition, data from *Jones,P.D.,and T.M.L.Wigley,1980:* Clim. Monitor,9, 43-47; and later updatings in Clim. Monitor. - *Salinger,M.J.,1980:* Clim. Monitor,9, 112-118 and priv. comm. - *NCAR* (Boulder, USA) world monthly surface clim. data collections. - *National Weather Services* data publications and priv. comm.

DYNAMIC MESA-SPECTRA OF CLIMATOLOGICAL TIME SERIES

H. P. Junk

Meteorological Institute
University of Bonn,
Auf dem Hügel 20,
D-5300 Bonn 1

1. INTRODUCTION

The maximum-entropy-spectral-analysis (MESA), also known as maximum-entropy-method (MEM), has found numerous applications in meteorology in recent years, especially in the analysis of climatological time series (e.g. Currie, 1974; Bain, 1976; Dyer, 1976; Olberg, 1979). Because of the considerably better frequency resolution properties, the MESA gives the possibility to investigate the low-frequency part of the power spectrum in much greater detail as it can be done by conventional spectral estimation methods (Blackman & Tukey method, discrete Fourier-transformation). It was therefore obvious to use this attribute of the MESA for the analysis of the stability of power spectra. The first "dynamic" MESA was applied on geomagnetic micropulsations by Radoski et al. (1976). Olberg & Rosenow (1980) used this method for the investigation of four climatological time series. However, they solved the Yule-Walker equations direct by estimating the autocovariances with a conventional estimator to make statements about confidence intervals. But this contravenes the principle of maximum entropy, as several authors have already stated, for example Ulrych & Clayton (1976). In this paper two time series, the rainfall series Nauru and the area-averaged rainfall series Black Forest, will be investigated by means of the dynamic MESA.

2. CONCEPT OF DYNAMIC MESA

The following figure shows the principle of the dynamic MESA.

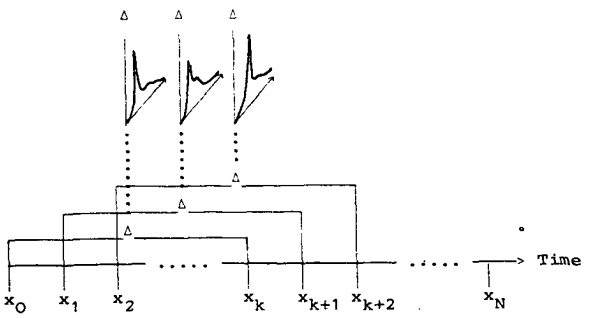


Fig. 1. Scheme of dynamic MESA

A time series consisting of N data samples is given as an example. With the MESA the spectrum or a part of the spectrum for the first k data points is computed. Then this "data-window" is moved forward, say 1 data points (in Fig. 1 1 = 1) and the next spectrum is computed for data points 1 to k + 1 and so on. This dynamic spectral analysis gives an overview about shifts and persistence of spectral peaks with time.

3. EXAMPLES

3.1. NAURU-SERIES

The first series is the rainfall series of the little Pacific island Nauru, which is situated about 2000 km east of New Guinea (0° 34'S, 166° 55'E). Fleer (1975) shows within the scope of power spectrum analysis of long monthly rainfall series in the area of Indonesia, Australia and the western Pacific that the power spectrum of the Nauru series has a remarkable significant peak at 60 months, whose variance contribution to the total variance is twice as big as the variance of the relative weak annual peak. In the following figure a Blackman & Tukey spectrum is given for a lag of 120 months, which corresponds to 13 percent of the series length. The series consists of 948 monthly data points from 1892 to 1970. To smooth the raw spectrum values a Hamming-filter was used.

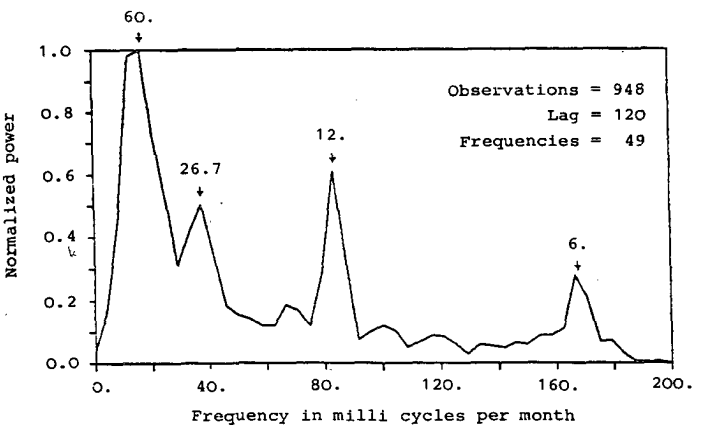


Fig. 2. Blackman & Tukey spectrum for Nauru-series with lag = 120 months.

One can recognize four marked spectral peaks at 6, 12, 26.7 and 60 months. The last two are the frequently cited quasi-two year and five year oscillations. They are found in many time series of meteorological elements (see Fleer, 1981). According to Fleer (1975) merely the half-year wave and the 60 months oscillation are significant with 3 σ (by the assumption of a red "null" continuum) in the Nauru spectrum.

The MESA spectrum for 50 filter coefficients (Fig. 3) has about the same spectral resolution as the Blackman & Tukey spectrum mentioned above. It has been computed with the well-known Burg-formula. In addition to this a search-algorithm for the spectral maxima has been used, described by Radoski et al. (1975). A spectral grid of 500 equally spaced discrete frequencies and their corresponding spectral values were computed. The search-algorithm yields 26 additional frequency points, so that the neighbouring spectral values of the maxima reach up to 99 percent of the maxima height.

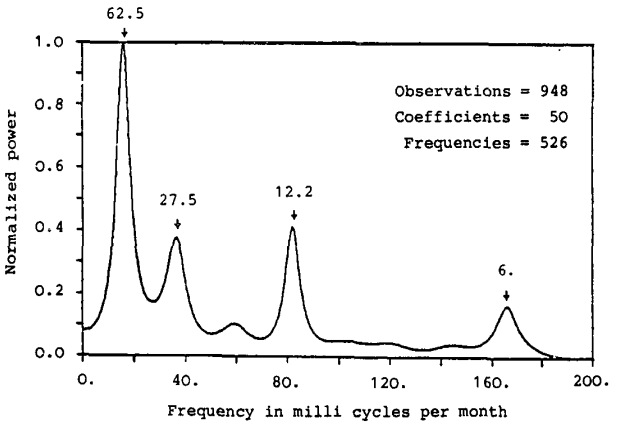


Fig. 3. MESA-spectrum for Nauru-series with 50 filter coefficients.

The peaks are located at 6, 12.2, 27.5 and 62.5 months. Using 120 filter coefficients (Fig. 4) these peaks lie at 6, 12.1, 26.7 and 69.4 months. Additionally a peak at 44.6 months can be seen.

Increasing the filter length, no appreciable spectral shifting of these peaks can be determined. The peak at 60 months, cited by Fleer (1975), appears to be traced back to an overlapping of the peaks at about 70 and 45 months, because of the poor resolution for lag = 120 months. This phenomenon occurs too at the analysis of synthetic signals (see for instance Junk, 1982). Another possibility, the "spectral shifting" described by several authors (e.g. Fougere, 1977), which often appears when using the Burg-algorithm to compute MESA-spectra, can

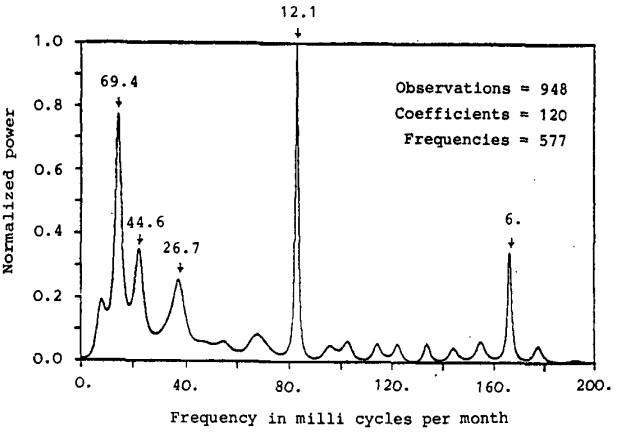


Fig. 4. MESA-spectrum for Nauru-series with 120 filter coefficients.

not be excluded. Therefore it was decided to use an independent method, the numerical band-pass-filter analysis. The filter weights were computed corresponding to the formulas specified by Schönwiese (1974, 1975). The following figure shows the variances of the filtered Nauru-series for periods from 55 to 85 months.

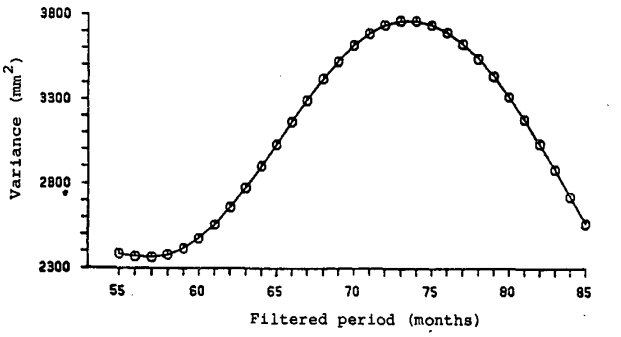


Fig. 5. Variances for the band-pass-filtered Nauru-series for 55 to 85 months.

The maximum is located at 74 months with 3750 mm^2. In Figs. 6 and 7, the band-pass-filtered series for 60 and 74 months are given. 88 filter weights were used for the first series (60 months) and 109 filter weights for the second one (74 months). While at 60 months several phase discontinuities can be seen, there are non at 74 months. Thus the band-pass-filter analysis confirms the result of the MESA. A "spectral shifting" to higher periods (from 60 to 70 months) due to the Burg-formula can be excluded.

As one can see from the band-pass-filtered series the amplitudes are subdued more or less on temporal fluctuations. The question arises, whether the single oscillations respectively the spectrum are persistent. Fig. 8 shows a dynamic MESA-spectrum of the Nauru-series for 45 filter coefficients. The subseries have

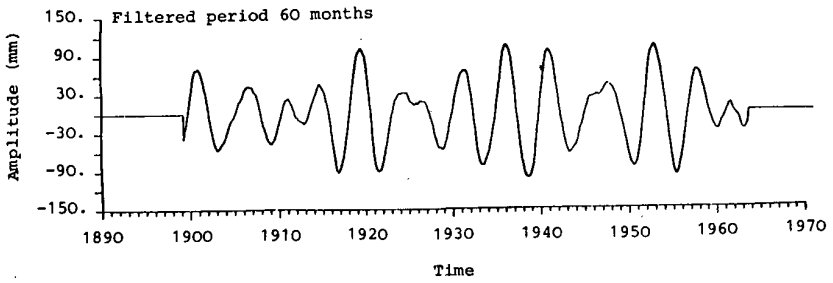


Fig. 6. Band-pass-filtered Nauru-series (60 months).

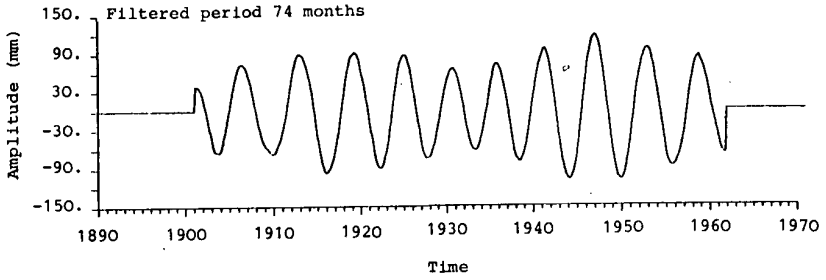


Fig. 7. Band-pass-filtered Nauru-series (74 months).

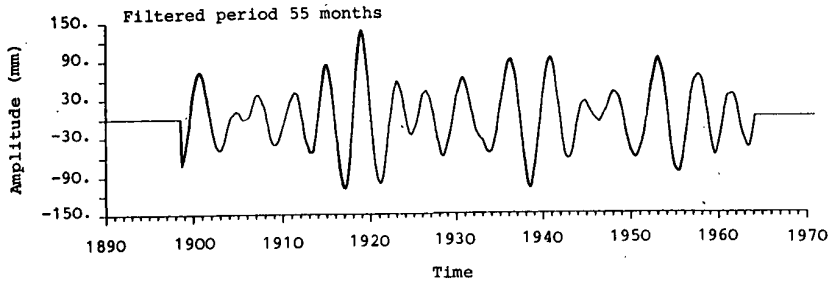


Fig. 9. Band-pass-filtered Nauru-series (55 months).

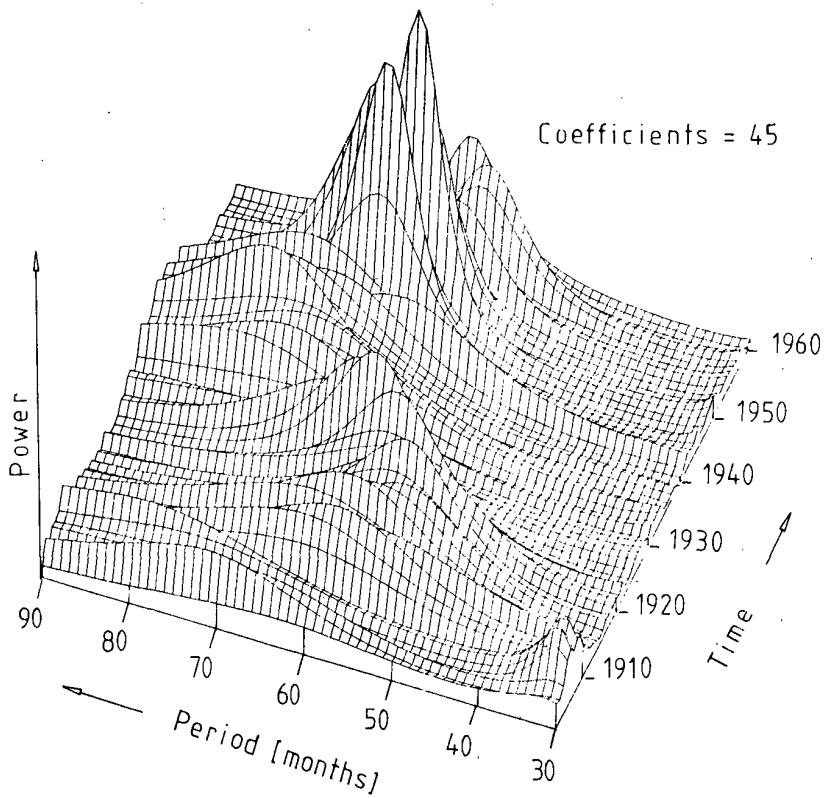


Fig. 8. Dynamic MESA-spectrum of the Nauru-series with 45 filter coefficients.

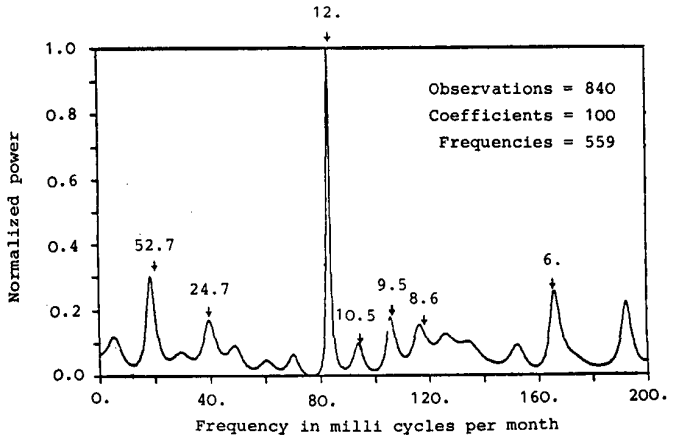


Fig. 10. MESA-spectrum for Black forest-series with 100 filter coefficients.

a length of 240 months and the time shift of the data-window amounts to 12 months. The three-dimensional plot gives the spectral density in dependence on period and time. The centre of each data interval is taken as the temporal offset of the respective spectrum (X-axis). Beside the large varying height of the peaks particularly one has to point to the large shifting of the spectral maximum with time (55 to 85 months). Very striking is the minimum at 1920 with about 55 months. Also for different filter lengths (45-60 coefficients) dynamic MESA-spectra were computed. All of them showed this minimum. The band-pass-filtered series for 55 months confirmes the result. Fig. 9 shows the filtered series. Clearly one can see a maximum of the amplitude at 1919. The amplitude reaches their largest value for the whole filtered series (140 mm).

3.2. Black Forest-series

This time series was compiled by Ambs 1978. It is an area-averaged rainfall series which represents the rainfall conditions in the Black Forest (SW-Germany). Ambs found a significant 8-months oscillation in the Blackman & Tukey spectrum of this series. Several authors have cited this oscillation in other series. Wacker (1979) computed a significant 8-months variance component in spectra for global radiation series in the southwest German area. Also Hillebrand (1978) found evidence for this oscillation in rainfall series from Bavaria. This seems to give rise to a more precise investigation with the dynamic MESA.

The series consists of 840 monthly data points from 1900 to 1970. Fig. 10 gives a MESA-spectrum for 100 filter coefficients. Besides the annual wave a broad spectral maximum with several peaks can be seen. The peaks are located at 10.5, 9.5 and 8.6 months. The integral over the interval 7.5 to 8.5 months (4.8 percent of the total variance) is larger than the variance contribution of the annual wave (11.4 to 12.6 months: 4.2 percent).

The band-pass-filtered series for 8 months (Fig. 11) shows that the amplitude is subjected to considerable fluctuations. Also a lot of phase discontinuities can be seen. Therefore, there is no evidence for a rythmic nor a periodic behaviour. The following dynamic MESA confirms this (Fig. 12). On the X-axis the frequency interval from 0.1 to 0.15 cycles per month (period 6.7 to 10 months) is given. The subseries have a length of 240 months, the time shift is 12 months and the filter length amounts to 40 coefficients. Up to 1930 no peak at 8 months can be seen. The spectra show almost continuously two peaks at about 9.1 and 7.1 months. From 1930 to 1943 a single peak at 8 months is recognizable. In this period the amplitude of the band-pass-filtered series (Fig. 11) is relatively well marked. After 1943 again a splitting in two peaks has taken place.Now they are located at 7 and 8.4 months. A periodic or quasi-periodic oscillation can therefore be excluded. A physical interpretation of this variance component is still unknown.

4. SUMMARY AND FUTURE WORK

Two rainfall-series were investigated with the dynamic MESA. The focal point on the Nauru-series was the peak in the low-frequency part of the power spectrum (according to Fleer, 1975: 60 months). The dynamic MESA showed that we are dealing with a quasi-periodic oscillation, whose period is located at about 70 months in the spectrum of the whole series. The maximum deviations from this "middle position" lie at $\pm$ 15 months (1905: 85 months, 1920: 55 months). In contrast to this the investigation of the 8 months variance component in the spectrum of the Black Forest-series gave no evidence for periodic nor a quasi-periodic oscillation.

Future analyses with the dynamic MESA will be engaged with several other time series of climatological elements (for instance tree-ring series). Furthermore, the Burg-algorithm will be replaced by the Marple-algorithm (Marple 1980). This algorithm is a recursive solution of

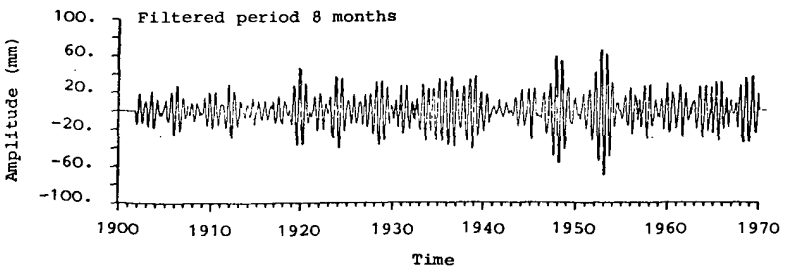


Fig. 11. Band-pass-filtered Black Forest-series (8 months).

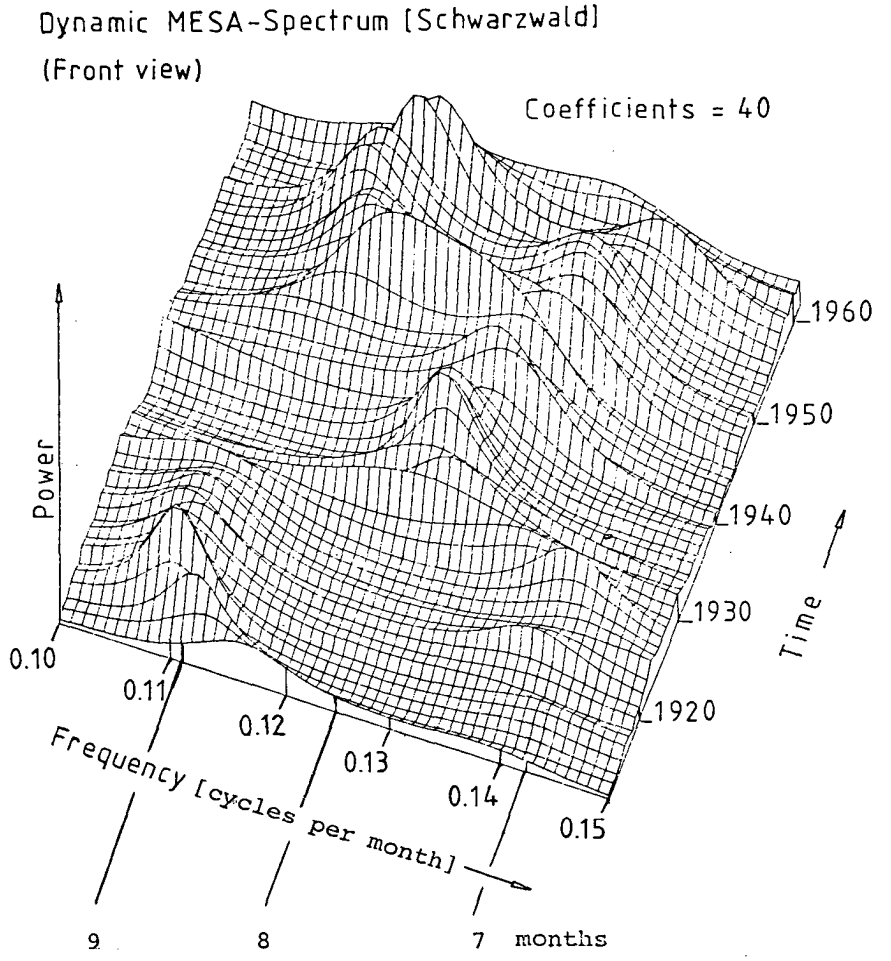


Fig. 12. Dynamic MESA of the Black Forest-series with 40 filter coefficients.

the Least-square algorithm cited by Ulrych & Clayton (1976) and Swingler (1979). The advantage of this algorithm is that it avoids spectral shifts and the spectral splitting, which can occur by using the Burg-algorithm.

5. REFERENCES

Ambs, A., 1978: Statistische Analyse von Zeitreihen monatlicher Niederschlagssummen im Oberrheingebiet und Bildung von Gebietsmitteln. Diploma thesis, Met. Inst. Univ. Bonn.

Bain, W.C., 1976: The power spectrum of temperatures in Central England. Quart. J. Roy. Met. Soc., 102, 464-466.

Currie, R.G., 1974: Solar cycle signal in surface air temperature. J. Geophys. Res., 79, 5657-5660.

Currie, R.G., 1981: Evidence for 18.6 year Mn signal in temperature and draught conditions in North America since AD 1800. J. Geophys. Res., 86, 11055-11064.

Dyer, T.G.J., 1976: An analysis of Manley's central England temperature data: I. Quart. J. Roy. Met. Soc., 102, 871-888.

Fleer, H., 1975: Spektrum- und Kreuzspektrumanalyse von Niederschlagsreihen aus Indonesien, Australien und Westpazifik. Diploma thesis, Met. Inst. Univ. Bonn.

Fleer, H., 1981: Large-scale tropical rainfall anomalies. Bonner Met. Abhandl. No. 26.

Fougere, P.F., Zawalick, E.J. and Radoski, H.R., 1976: Spontanous line splitting in maximum entropy power spectrum analysis. Phys. Earth. Planet. Int., 12, 201-207.

Hillebrand, Ch., 1978: Statistische Untersuchungen von Niederschlagsreihen in Mitteleuropa und von Gebietsmitteln für den Niederrhein. Diploma thesis, Met. Inst. Univ. Bonn.

Junk, H.P., 1982: Die Maximum-Entropie-Spektral-Analyse (MESA) und ihre Anwendung auf meteorologische Zeitreihen. Diploma thesis, Met. Inst. Univ. Bonn.

Marple, L., 1980: A new autoregressive spectrum analysis algorithm. IEEE Trans. Acous. Speech.Signal.Proc., 28, 441-454.

Olberg, M., 1979: Langperiodische Schwankungen der Strahlungsbilanz des Niederschlags und des Trockenheitsindexes in Potsdam. Z. Met., 29, 50-55.

Olberg, M. und Rosenow, W., 1980: Zeitliche Änderungen im spektralen Verhalten meteorologisch-klimatologischer Datenreihen. Z. Met., 30, 296-307

Radoski, H.R., Zawalick, E.J. und Fougere, P.F., 1976: The superiority of maximum entropy power spectrum techniques applied to geomagnetic micropulsations. Phys. Earth. Planet. Int., 12, 208-216.

Schoenwiese, C.D., 1974: Schwankungsklimatologie im Frequenz- und Zeitbereich. Wiss. Mitt. Met. Inst. Univ. München, No. 24.

Schoenwiese, C.D., 1975: Bemerkungen zur numerischen Bandpassfilterung von Zeitreihen unter besonderer Berücksichtigung meteorologischer Daten. Arch. Met. Geophys. Biokl. Ser. A; 24, 311-320.

Swingler, D.N., 1979: A comparison between Burg 's maximum entropy method and a nonrecursive technique for the spectral analysis of deterministic signals. J. Geophys. Res., 84, 679-685.

Ulrych, T.J. and Bishop, T.N., 1975: Maximum entropy spectral analysis and autoregressive decomposition. Revs. Geophys. Space Phys. 13, 183-200.

Ulrych, T.J. and Clayton, R.W., 1976: Time series modelling and maximum entropy. Phys. Earth. Planet. Int. 12,188-200.

Wacker, U., 1979: Untersuchung langfristiger Schwankungen der Sonnenscheindauer und Abschätzung der Globalstrahlung im südwestdeutschen Raum. Diploma thesis, Met. Inst. Univ. Bonn.

RED AND WHITE NOISE AND TURBULENCE OF LOW-LEVEL WINDS FOR TIME INTERVALS OF SECONDS

Oskar M. Essenwanger

U. S. Army Missile Command
Research Directorate
Redstone Arsenal, Alabama 35898

1. INTRODUCTION

Wind fluctuations in the boundary layer related to turbulence have found extensive treatment in the literature. It is well documented (e.g., Kolmogorov, Obukhov and Corrsin, (Hinze, 1959), etc.) that turbulence provides a slope of -5/3 in a diagram power spectrum lines versus (inverse) wavelength in double logarithmic coordinates for the inertial subrange of turbulence. Thus it appeared logical to assume that wind observations made on towers by instrumentation with short response time (e.g., seconds) would show turbulent features on a day with some convective activity and surface friction. Indeed, some power spectra from that day produced this -5/3 slope, but others did not. In these cases, another mechanism appeared dominant, and interpretation from turbulence theory alone was not sufficient.

It was postulated that the statistical background for the wind fluctuations was a random process with or without persistence (i.e., white or red noise, respectively). This assumption provided a suitable interpretation for the slope angle of the power spectrum (in double logarithmic coordinates). In fact, an almost linear relationship exists between the slope angle of the power spectrum and the first autocorrelation coefficient.

In order to gain a deeper understanding of the interactions between white and red noise processes and slope angles, and to support the interpretation of the observed tower data, simulated data sets were produced by modelling white and red noise. It will be illustrated that red noise and turbulence may differ in the contributions of the longer waves to the variance. Turbulence in the inertial subrange may also require a specific scale related to the time sequence of fluctuations. The details of the modelling effort and the comparison with empirical data will be presented in the subsequent sections.

2. SLOPE ANGLE AND FIRST LAG CORRELATION

In the last two decades the development of wind sensors with high sensitivity and short response time such as the Gill u-v-w anemometers (Gill, 1975) made it possible to measure "instantaneous" fluctuations of the wind in short time intervals (e.g., seconds). Various authors (e.g., Panofsky, 1973; Busch, 1973; Lumley and Panofsky, 1964; Pasquill, 1962; Priestley, 1959, etc.) have assumed that the criteria for the existence of the

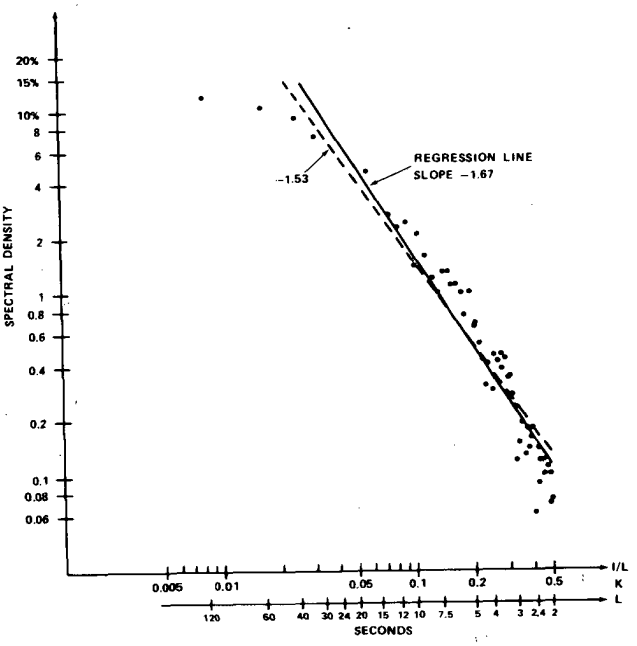


FIGURE 1. POWER SPECTRUM OF WINDSHEAR Δ U, LEVEL 3-2 TOWER DATA, 19 AUGUST 1974 1 SECOND RECORDINGS, MAX. LAG 60 SECONDS

inertial subrange conditions are generally fulfilled in the atmosphere (provided measurements are not too close to the ground). According to Kolmogorov, Obukhov and Corrsin (e.g., Hinze, 1959; Tennekes and Lumley, 1972, etc.), under these conditions the power spectral density would produce a slope of -5/3 in a diagram of the logarithm of the spectral density versus the logarithm of the (inverse) wavelength valid for the three-dimensional energy spectra or the one-dimensional components. The energy spectrum is:

$$E(k) = \alpha\varepsilon^{-\frac{2}{3}} k^{-\frac{5}{3}} \qquad (1)$$

where α is the Kolmogorov constant, ε the dissipation (here constant) and k the wave number. The power spectrum density L_j is related to (1) by:

$$\ln E(k) = \ln L_j \qquad (2)$$

Examination of "instantaneous" wind observations taken on towers by Gill u-v-w anemometers at Huntsville, Alabama, and at Otis Air Force Base

revealed that some power spectra for the wind components or shears followed the -5/3 law (e.g., Fig. 1, solid line), but others did not. Thus the question arose whether these measurements at short time intervals show fluctuations which may need a different interpretation. Figure 1 represents the windshear of the U-component measured at 30 feet and 18 feet height.

It was tested whether the wind fluctuations and their power spectra could have been produced by random processes such as white noise or by random processes with persistence (red noise). For red noise, an exponential pattern was assumed which is equivalent to a first order Markov chain (Essenwanger, 1980). Persistence is found in many meteorological elements. Since the first autocorrelation (lag) correlation ρ_1 defines the exponential red noise pattern, a series of power spectra for exponential red noise was constructed for a uniform maximum lag m = 60 and plotted in double logarithmic coordinates (Fig. 2). As delineated by Fig. 2, the slope angle changes as a function of the first autocorrelation. White noise, with $\rho_1 = 0$, would produce a horizontal line, i.e., the slope angle is zero.

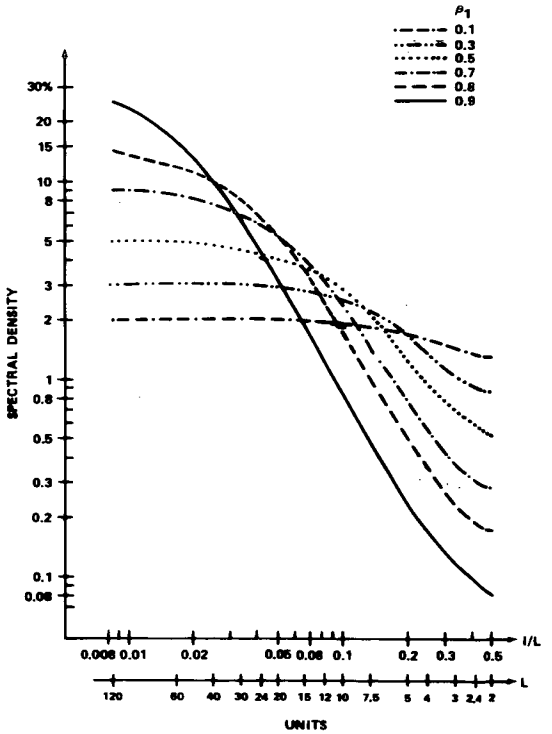


FIGURE 2. POWER SPECTRA OF EXPONENTIAL RED NOISE

Figure 2 is based on data whose maximum lag in the autocorrelogram is kept greater than a defined threshold. This implies that the truncation takes place after the lag correlation has decreased to a small value. As shown by Essenwanger (1983), the maximum lag and the first autocorrelation for exponential red noise are governed by:

$$m \geq (\ln R_t / \ln \rho_1) \qquad (3)$$

This leads to $m \geq -6.2146/\ln \rho_1$ for a truncation correlation coefficient of $R_t = 0.002$, or for a maximum lag of m = 60 the corresponding $\rho_1 \leq 0.90$. If $\rho_1 > 0.90$, then $m > 60$, and m must be adjusted. Red noise curves for $\rho_1 > 0.90$ were omitted from Fig. 2.

If the background of exponential noise is suitable, then the slopes of the power spectra for our data taken at short time intervals should display some association with the first lag correlation. Indeed, examination of the slope angles of power spectra for various data samples (single wind components, total wind, and shears) from Huntsville and Otis Air Force Base led to the data in Fig. 3. The slope angles θ were calculated from fitting a linear regression line to power spectra (double logarithmic coordinates, see Fig. 1). In order to obtain the slope angle of the shorter waves, the spectra were truncated successively at the long wave part (e.g., first, second, third, etc.) until stability of the slope angle was achieved. The dashed line in Fig. 1 exhibits the slope angle without the elimination of the first two waves. Although in this case the difference between the angle of the regression line of the full spectrum and the truncated spectrum is only about 3°, the slope angles can be considerably different if turbulence spectra assume the shape as illustrated by Busch (1973, p. 39). The graphs in Busch's text show that long waves decline considerably exhibiting a parabolic shape of the power spectra with a peak somewhat in the center of the spectra. From these types of spectra the slope angle of turbulence (in the inertial subrange) cannot be calculated by regression analysis without proper truncation of the longer waves.

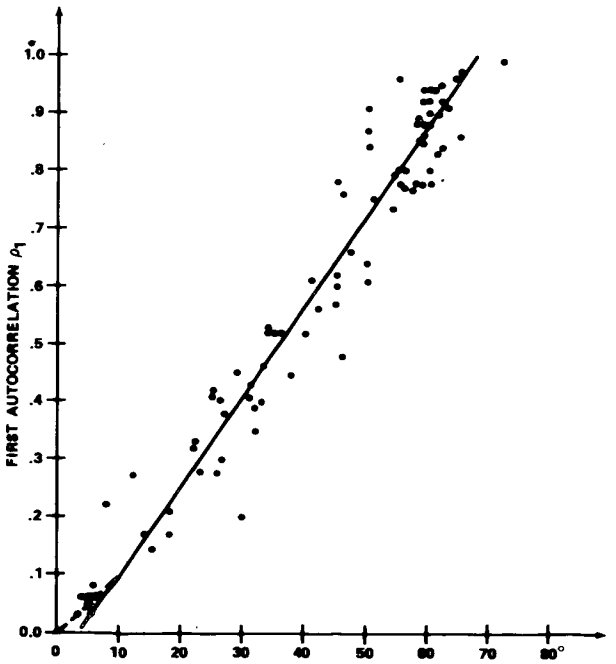


FIGURE 3. RELATIONSHIP BETWEEN FIRST AUTOCORRELATION COEFFICIENT AND SLOPE ANGLE OF POWER SPECTRUM

Figure 3 displays an almost linear relationship between the slope angle and the first autocorrelation. The solid line is the linear regression line of the data points in Fig. 3. It should be noted that this linear regression line from 90 power spectra intersects the abscissa for $\rho_1 = 0$

at a point for θ of about 4°. For $\rho_1 = 0$, however, we have white noise which necessitates a slope angle of $\theta = 0$. Therefore, an adjustment and a deviation from the linear line as indicated by the dashed curve was made bending towards the origin for the correlations $\rho_1 < 0.1$.

As exhibited by Fig. 3, the slope angle of the power spectra is a function of the first lag correlation. This result implies that our wind data may be produced by exponential noise. Therefore, the interpretation that the fluctuations of the wind in our data are governed by persistence is confirmed.

The idea to relate turbulence and exponential red noise is not new, e.g., Fichtl and McVehil (1970), Pasquill (1962), Hanna (1979), or Stewart (1981). A further discussion will follow later.

3. INTERMITTENT AND AVERAGED OBSERVATIONS

The results exhibited in the diagram of Fig. 3 have significant applications. Suppose we record the wind components every sixth second. This corresponds to an intermittent recording at six-second intervals and produces an autocorrelogram of every sixth coefficient from the autocorrelogram of one-second readings. This series shows a completely different slope in the spectrum. The sixth autocorrelation of the data series in Fig. 1 was $\rho_6 = 0.267$, which is now converted to ρ_1. We would expect a slope angle of 22° according to Fig. 3. The slope angle from the data sample (Fig. 4) turned out to be 22.9° (slope -0.422). This is an excellent agreement.

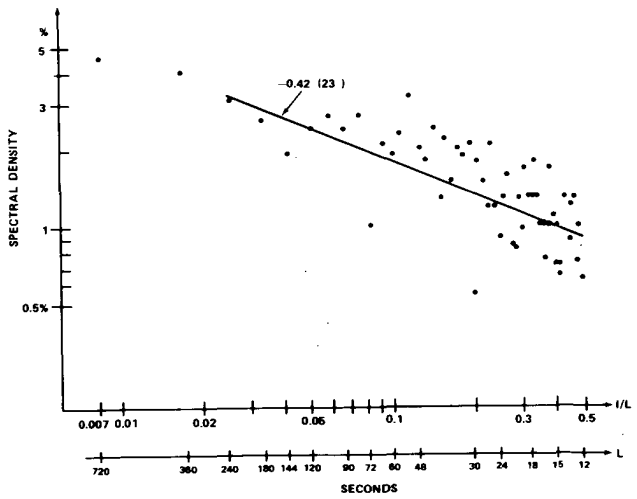


FIGURE 4. POWER SPECTRUM OF THE WINDSHEAR Δ U, LEVEL 3-2
TOWER DATA, 19 AUGUST 1974
6 SECOND RECORDINGS, MAX. LAG 360 SECONDS

One conjecture is that the difference rests in the longer range of waves for m = 60 because now the first wave has a length of 720 seconds compared with 120 seconds for the one-second data. However, the slope is essentially the same (Fig. 5) if the power spectrum is constructed for a maximum lag m = 10. The range of waves comprises now a length of 120 to 12 seconds, with a truncation of waves from 12 to two seconds compared with m = 60 in Fig. 1. Although the slope angle of 27° is 4° different, it is not a significant change.

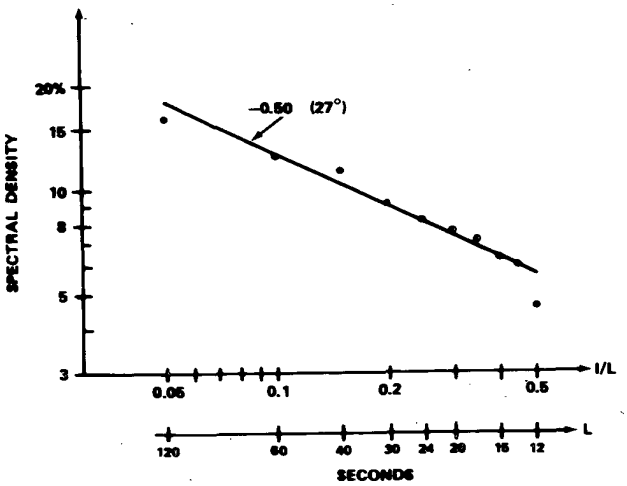


FIGURE 5. POWER SPECTRUM OF THE WINDSHEAR Δ U, LEVEL 3-2
TOWER DATA, 19 AUGUST 1974
6 SECOND RECORDINGS, MAX. LAG 60 SECONDS

The independence of the slope angle from the maximum lag (unless m violates (3)) has been discussed in detail by Essenwanger (1982).

In consequence of Fig. 3, we would also expect that averaging of data would affect the slope. Figure 6 discloses the slope for six-second (non-overlapping) average values which may correspond to a slower responding instrument resembling the measurement of a time average. In our case, the first lag correlation $\rho_1 = 0.445$ which leads to an expected slope angle of 34° (Fig. 3). The regression line in Fig. 6 discloses an angular slope of 38.5°, which is not a significant departure.

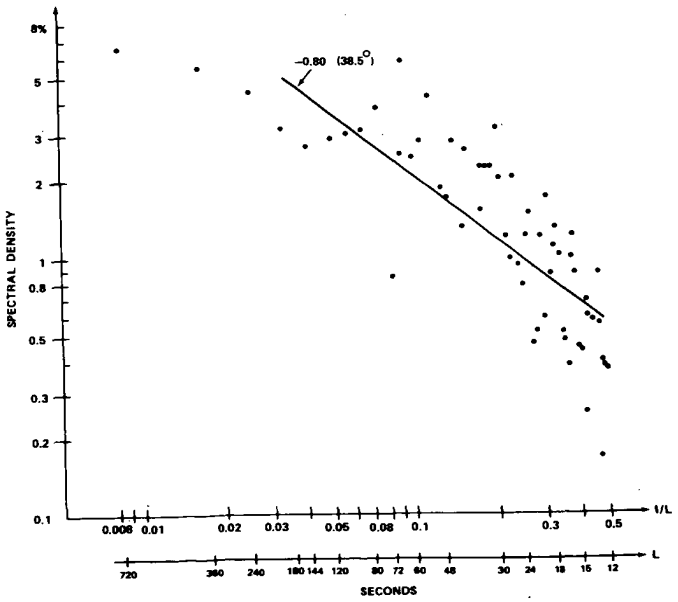


FIGURE 6. POWER SPECTRUM OF THE WINDSHEAR Δ U, LEVEL 3-2
TOWER DATA, 19 AUGUST 1974
6 SECOND AVERAGES, MAX. LAG 360 SECONDS
NON-OVERLAPPING

The particular example for this article was so chosen that the one-second recordings produced a slope angle of -5/3, and only intermittent measurements (six seconds) or averaging produced a different slope angle. Thus, any other interpretation than different conditions of red noise could be excluded. Otherwise, one may have suspected that the disagreement with the -5/3 law could be in the length of the record. The measurements stretched over one hour at Huntsville, Alabama, on 19 August 1974, and the observations at Otis Air Force Base at six-second intervals extended over four hours. These time periods have been considered by other authors as reasonably homogeneous (e.g., Wyngaard, 1973). The mean $\overline{U}$ value (turbulence component) did not vary significantly within the time frame of the measurements for the Huntsville data. Fluctuations in the wind direction during the measurements at Otis Air Force Base over the longer time period on that particular day were taken into consideration by breaking the data into homogeneous sections. However, in order to produce data samples for controlled homogeneous conditions, a simulation study of red noise was performed.

4. WHITE AND RED NOISE SIMULATION MODEL

As stated before, a study of the wind fluctuations for data which were independent from observational deficiencies such as hidden inhomogenities was desirable. Thus, data samples for white noise and exponential red noise were generated from controlled conditions with known properties.

A prior investigation by Essenwanger (1982) served as the basis for the modelling. In this article the author documented that the construction of the power spectrum via the Fast Fourier Transform (FFT) (see Cooley and Tukey, 1965) may very often be considered as a suitable replacement for the power spectrum obtained via the Fourier transform of the autocorrelogram (Tukey, 1949; Blackman and Tukey, 1958). Some differences may exist if "quasiperiodicity" is present such as in some time series of meteorological elements (see also Essenwanger, 1976, p. 234) because the phase angles are standardized in the autocorrelogram but they are not for the FFT. In this case, amplitudes of the Fourier series from the FFT and autocorrelogram will differ.

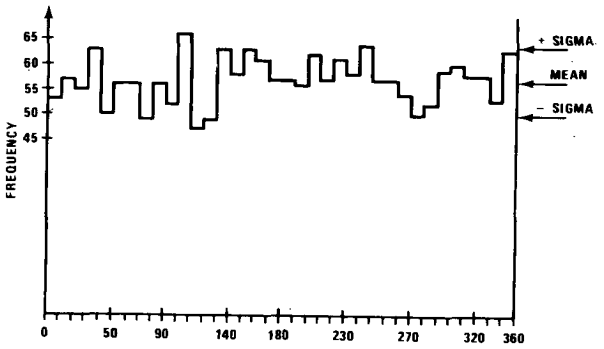
Data samples y_i, i = 1,...N for modelling were generated from:

$$y_i = A_o + \sum_{j=1}^{n} A_j \sin (\alpha_{ij} + \beta_j) \qquad (4)$$

where n = N/2, $\alpha_{ij} = 2\pi ji/N$, and β_j is a set of randomly produced phase angles. Since power spectra are based on the Fourier series, (4) is the most appropriate function.

A set of 2048 random phase angles was obtained by a random number generator converted to the scale 0 to 360° or 0 to 2π. The frequency distribution of these random angles is exhibited in Fig. 7. As evident from Fig. 7, the frequency density of the set provides a rectangular distribution with no significant departures at any groups of angles (at the 2 sigma level of significance). This data set was utilized for the construction of several data sets y_i by (4) as follows.

FIG. 7 HISTOGRAM OF PHASE ANGLES



The amplitudes of the Fourier term are governed by two equations. First, we observe that:

$$\sum_{j=1}^{n} A_j^2 = 2\sigma_y^2 \qquad (5)$$

where σ_y^2 is the variance of y_i (to be chosen arbitrarily without effects on the power spectrum). The second equation is a mathematical representation of the amplitudes for the power spectrum of white or red noise. For white noise, the amplitudes are constant:

$$A_j^2 = 2\,\sigma_y^2 / n \quad \text{for all } j. \qquad (6)$$

For red noise, the curves in Fig. 2 were broken into quasi-linear sections. It may be recalled that L_j of the power spectrum corresponds to $A_j^2 / 2\sigma^2$. In scaling for N, the correct value to fulfill (5) is not known. Therefore, an auxiliary amplitude B is obtained first. The following two basic equations were utilized:

$$B_j^2 = \exp (g_1) \text{ for } j = 1,\ldots,n_1 \qquad (7a)$$

and:

$$B_j^2 = [\exp (g_2)] / c_1 \text{ for } j = n_1 + 1,\ldots,n \qquad (7b)$$

(7a,b) can be further broken into subsections if necessary. Finally:

$$C^2 = 2\,\sigma_y^2 / \sum_{j=1}^{n} B_j^2 \qquad (8a)$$

and:

$$A_j = C\,B_j \qquad (8b)$$

(8a,b) serve to standardize the power spectrum so that $\Sigma L_j = 1.0$.

In our case, red noise for $\rho_1 = 0.9$ and N = 4096 required one section of $n_1 = 30$ (determined from the scaling of the spectrum from 120 to 4096). For convenience, $c_1 = 10$ to tie (7a) and (7b) together. From Fig. 2 it was derived:

$$g_1 = \text{constant} = 1.0986 \qquad (9a)$$

or (slanted):

$$g_1 = (\ln j) / 3.096 \text{ for } j = 1,\ldots,30. \quad (9b)$$

For (7b) two subsections were designed:

$$g_2 = -0.26 - 0.74 \ln (j/4096)$$
$$\text{for } j = 31,\ldots,84 \quad (10a)$$

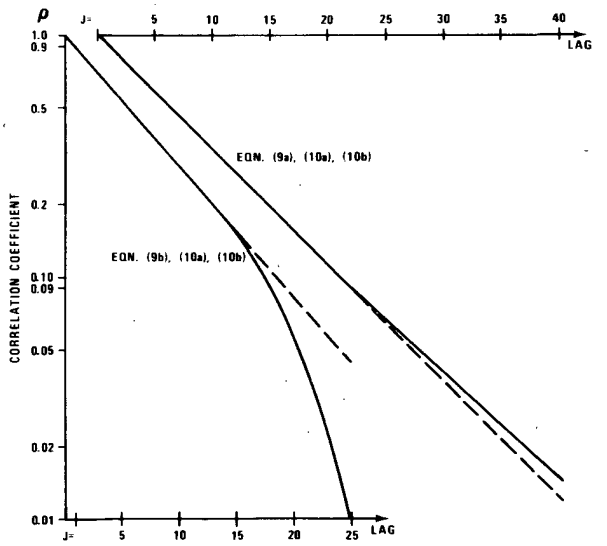
and:

$$g_2 = -3.80 - 1.67 \ln (j/4096)$$
$$\text{for } j = 85,\ldots,2048 \quad (10b)$$

The substitution of (9b) for (9a), i.e., L_j increases with increasing j, did not lead to a significant change in the first autocorrelation coefficient, although the autocorrelograms diverge for $\rho_j < 0.10$ (Fig. 8). The spectra of both series, however, are virtually identical (Fig. 9) with slope of -1.66.

The power spectra for the data sets with (9a) have a first autocorrelation coefficient $\rho_1 = 0.89$, while the substitution (9b) lowered it insignificantly to $\rho_1 = 0.88$. Both series show the features of exponential noise which corresponds to a straight line decrease in Fig. 8.

FIG. 8 AUTOCORRELOGRAM OF MODELS (SEMILOGARITHMIC COORDINATES)



$$\rho_j = \rho_0^j \quad (11)$$

where ρ_0 corresponds to the first lag correlation of ρ_1. The divergence below $\rho_j = 0.1$ is caused by the difference between (9a) and (9b), and the values of the β_j for $j = 1,\ldots,n_1$.

It may be noticed in Fig. 2 that toward the short wavelength end the red noise curve shows a slight upwards bend. This bend was not incorporated into (10b). The data series (9a), (10a,b) was utilized for modelling in the following section.

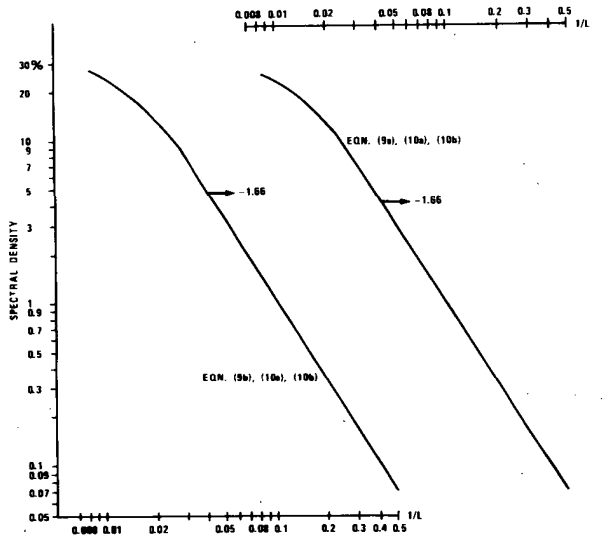
The white noise set (6) produced $\rho_1 = 0.005$ and an autocorrelogram with fluctuations $-0.005 < \rho_j < 0.005$. Since the power spectrum is defined by (6), a graph of the spectrum of the simulated data is omitted.

The primary purpose of the modelling effort was the study of red noise conditions. The set of equations (5) through (8) can be applied, however, to model any turbulence spectrum, e.g., the spectra illustrated by Busch (1973).

5. RESULT OF MODELLING

The following discussions will primarily refer to red and white noise processes. Because a close relationship exists between slope angle and first lag correlation (Fig. 3), the analysis can be restricted to the study of the first lag correlation.

FIG. 9 POWER SPECTRA OF MODELS



It is apparent from Fig. 9 that exponential red noise (or white noise) data samples can be properly constructed based on (4) for any N and specifying conditions for the slope. If the single component of turbulence were to follow a red noise pattern with a slope angle of -5/3, the first lag correlation should be in the range $0.85 \leq \rho_1 \leq 0.90$ for strictly red noise conditions, but could probably be $0.8 \leq \rho_1 \leq 0.9$ in empirical cases.

Under the assumption of red noise with $\rho_1 = 0.89$ for the U and V component of turbulence (series in Fig. 9), the total horizontal wind vector $V_H = U^2 + V^2$ would not produce a slope of -5/3. In this case the first lag correlation for V_H is $\rho_1 \sim 0.79$ with slope around -1.38 assuming the same magnitude of variance σ_y^2 for U and V. Since σ_V is usually smaller than σ_U we find $0.79 < \rho_1 < 0.89$, depending on the ratio σ_U to σ_V.

The energy flux is a function of the slope angle in the power spectrum (e.g., Tennekes and Lumley, 1973, p. 268, etc.). The slope angle for -5/3 is 59°. The flux increases when $\theta < 59°$ and

decreases for $\theta > 59°$. This fact may be interpreted (assuming a red noise background) that a decrease of persistence would be observed with increasing energy flux and vice versa. However, more modelling effort may be necessary, especially simulating deviations from the exponential red noise pattern such as a rapid decrease toward the long wavelength end and a steeper slope towards the short wavelength end (dissipation spectral region).

The experimentation with white and red noise model data samples revealed one interesting fact concerning spectra of the wind shear. Suppose we calculate the wind shear $\Delta U_{ji} = U_{ji} - U_{1i}$ for the U-component of the wind, where U_{1i} is represented by the data series y_i, and $j = 2$, U_{2i} by y_{i+t}, $j = 3$, U_{3i} by $y_i + t_o + t$. This concept is equivalent to the observed fact that persistence decreases either with increase in time or distance or a combination of both. In our data samples, where $\rho_1 = 0.89$ for the U_{1i} series, the first lag correlation for ΔU_i increases from an original value around zero to 0.89 after $t \sim 30$ (Table 1). In the second series with U_{3i}, where $t_o > 30$, the first lag correlation remains constant (Table 1).

Table 1

FIRST LAG CORRELATION FOR SIMULATED WINDSHEAR ΔU AND ΔS

	Red Noise, $\rho_1 = 0.89$				White Noise
	ΔU_2	ΔU_3	ΔS_2	ΔS_3	ΔS_2
t		$t_o > 30$		$t_o > 30$	
1	-0.02	0.89	0.05	0.79	0.28
2	0.43	0.89	0.24	0.79	0.03
4	0.68	0.89	0.40	0.79	0.05
4	0.81	0.89	0.55	0.79	0.03
16	0.86	0.89	0.70	0.79	0.06
32	0.89	0.89	0.77	0.79	0.03
64	0.89	0.89	0.79	0.79	0.05
128	0.89	0.89	0.79	0.79	0.05
256	0.89	0.89	0.79	0.79	0.05

ΔU_{ji} and ΔS_{ji}, see text.

This increase from a low first lag correlation to a constant value after a certain minimum threshold t_o can also be observed for the vector shear of two components U, V,

$$\Delta S_{ji} = (U_{ji} - U_{1i})^2 + (V_{ji} - V_{1i})^2,$$

$j = 2,3$ (Table 1). Again, if $\sigma_V \sim \sigma_U$ the assymptotic value for ρ_1 stays around 0.78, while $\sigma_V < \sigma_U$ would increase the first lag correlation, $0.79 < \rho_1 < 0.89$, depending on the magnitude of the ratio σ_V/σ_U (provided $\sigma_V < \sigma_U$).

The results of Table 1 may at first appear as a contradiction to empirical results where the correlation decreases with increase of the shear interval (in time or space). E.g., the U-components at levels 2 and 3, tower 3, measured on 19 August 1974 at Huntsville, displayed a high persistence of $\rho_1 = 0.97$ each. The components were highly intercorrelated with each other, $\rho_{23} = 0.89$. These conditions lead to $\rho_1 = 0.82$ and the power spectrum of Fig. 1.

$\rho_{23} = 0.89$ corresponds to a shift of 5 seconds in the autocorrelogram of the individual component. In our case of modelling, it is five units. A glance at Table 1 reveals that ΔU_2 would be of the magnitude around 0.7 for $t = 5$. This result is equivalent to the empirical conditions described above. It indicates that the turbulence structure does not only depend on the persistence of the components, but at different altitude levels the phase angles of the different waves are shifted or changed. A shift of y_i by t, i.e., y_{i+t}, is equivalent to a shift in phase angles of the Fourier terms in (4). After a sufficient shift (e.g., in our case, 30 units), the correlation of ρ_1 stabilizes.

The difference in the phase angle between two levels of tower observations can be confirmed by a FFT analysis of wind components which are utilized in the shear analysis. As expected, the fluctuations in the amplitude were not significant and the slope of the power spectrum produced from the FFT rendered an almost identical slope angle as the spectrum via the autocorrelogram.

Table 1 discloses further that for white noise ρ_1 remains around zero except for $t = 1$. Again, the value for $t = 1$ is due to an insufficient shift of the phase angles in modelling. All empirical data analyzed by this author showed a sufficient phase shift. The fluctuation around $\rho \sim 0.04$ rather than an expected $\rho = 0$ is an effect of the "imperfections" of the set of phase angles utilized for modelling. A perfect rectangular distribution would very likely lead to the proper value $\rho \sim 0$.

A slope angle of 59° produces a slope of -5/3. According to Fig. 3, ρ_1 would be 0.86. Under consideration of statistical fluctuations, we would expect exponential red noise to produce a slope of around -5/3 for $0.8 < \rho_1 < 0.9$. This leads to the following interpretation for the scale of turbulence.

In turbulence analysis

$$U = \overline{U} + U' \tag{12}$$

and

$$\overline{R}_s = \int_0^\infty \rho(t)dt \tag{13}$$

For time series records at a given point, $\overline{R}_s$ is the Eulerian integral scale (e.g., Tennekes and Lumley, 1972, p. 275, etc.). It can be related to the Eulerian length scale of turbulence:

$$L_E = \overline{U}\,\overline{R}_s \tag{14}$$

(e.g., Ivanov and Klinov, 1962; Tennekes and Lumley, 1972; Hanna, 1979; Weber et al., 1982, etc.).

In the exponential model $\overline{R}_s$ is a unique function of ρ_1. If turbulence were to follow the

exponential red noise model, it would imply that turbulence in the inertial subrange requires a special value of $\overline{R}_s$ and the length scale of turbulence L_E would change only with a change of $\overline{U}$. If $\overline{U}$ is constant, however, the slope of the power spectrum is a function of $\overline{R}_s$ alone, which implies that it is a function of ρ_1, and may differ from -5/3. Thus the power spectrum may show other turbulence than the inertial subrange, a variation of the slope according to the type of stability conditions (e.g., Tennekes and Lumley, 1972; Weber et al., 1982, etc.), or no turbulence at all. In the latter case, the fluctuations of the wind velocity follow a random pattern with persistence (red noise).

One final point may be worthwhile for consideration. We learn from Fig. 3 that the slope angle changes with a change of ρ_1. Consequently, any empirical autocorrelogram following an approximate (empirical) red noise pattern could be converted by scaling into a series where the slope angle is -5/3. This scaling can be based either on a modification of (3) or a suitable interpolation of the autocorrelogram. E.g., linear scaling based on (3) would use the maximum lag m, and we calculate a ratio m_1/m_2. We find from (3) that $m_{0.9} = 65.5$ and $m_{0.5} = 10.0$ (the subscript representing ρ_1). Thus, a suggested scaling would be the addition of six to seven values between every correlation coefficient of the autocorrelogram for an original $\rho_1 = 0.5$.

Since the interpolation procedure would have no effect on $\overline{U}$, only $\overline{R}_s$ in (13) would change, provided dt means a unit time interval. This would change the Eulerian length scale of turbulence L_E. It is evident, however, that conversion of the six-second intermittent measurement to one-second intervals, although increasing $\overline{R}_s$, would decrease dt, now to be one second. Thus the effect on L_E may be a point of interpretation rather than substance.

The investigation is continuing whether interpolation of the autocorrelogram or utilization of the FFT is the best suitable technique. At the present time, results are inconclusive.

One word of caution must be added. If the data series follows basically a red noise pattern, scaling by the ratio of the maximum lag is appropriate. If the autocorrelogram decreases slower than the red noise series (i.e., the data follow a higher order Markov chain), Fig. 3 may not be suitable and (3) would lead to a maximum lag which is too small.

6. CONCLUSIONS

It is customary to consider the fluctuation of the wind in the boundary layer during short time intervals (e.g., seconds) to be in agreement with turbulence in the inertial subrange. While this association was confirmed for some tower data at Huntsville, other data, especially that taken at Otis AFB, delineated a different slope in the power spectrum (in double logarithmic coordinates). The detailed investigation revealed that these small-scale fluctuations are more likely to be produced by a random process with persistence (exponential red noise). This concept would explain that the angle in the power spectrum is not -5/3, but is a function of the first lag correlation (Fig. 3).

A second interpretation is possible. The turbulence structure in the inertial subrange requires a specific persistence to be present corresponding to red noise with a first lag correlation around 0.8 to 0.9.

The slope angle for averaged data cannot very simply be defined, but measurements with less sensitive instrumentation would resemble an averaging process where the slope angle is not necessarily -5/3. The deviation may be caused alone by the averaging process. The study of the effects has not been completed.

The investigation indicates that a slope of -5/3 in a diagram of power spectrum lines versus (inverse) wavelength (double logarithmic coordinates) can be obtained from a random process with persistence. Thus, it is not sufficient evidence to conclude that turbulence is present, although some authors assume modelling of the autocorrelation function by exponential red noise, even for turbulence. Other turbulence characteristics such as surface roughness, stability, or the intensity of turbulence should supplement any diagnostic process.

7. REFERENCES

Blackman, R. B., and J. W. Tukey, 1958: The Measurement of Power Spectra. Dover Pub., Inc., New York, 190 pp.

Busch, Niels E., 1973: On the Mechanics of Atmospheric Turbulence. Workshop on Micrometeorology (D. A. Haugen, Ed.), American Meteorological Society, Boston, 1-65.

Cooley, J. W., and J. W. Tukey, 1965: An Algorithm for the Machine Calculation of Complex Fourier Series. Math. Comput., 19, 297-301.

Essenwanger, Oskar M., 1976: Applied Statistics in Atmospheric Science, Part A, Frequencies and Curve Fitting. Elsevier, Amsterdam, 412 pp.

Essenwanger, O. M., 1980: On Red Noise and Quasi-Periodicity in the Time Series of Atmospheric Temperature. Statistical Climatology (S. Ikeda, E. Suzuki, E. Uchida and M. M. Yoshino, Ed.). Elsevier, Amsterdam, 165-181.

Essenwanger, Oskar M., 1982: Turbulence Analysis by Use of the Fast Fourier Transform. Proceedings Army Science Conference, 15-18 June 1982, Vol. I. Publ. Deputy Chief of Staff for Research, Dev. & Acqn., DA, 367-381.

Essenwanger, Oskar M., 1983: Elements of Statistical Analysis. World Survey of Climatology, I, (O. M. Essenwanger, Ed.). In press, Elsevier, Amsterdam.

Fichtl, G. H., and G. E. McVehil, 1970: Longitudinal and Lateral Spectra of Turbulence in the Atmospheric Boundary Layer at Kennedy Space Center. J. Appl. Meteor., 9, 51-63.

Gill, G. C., 1975: Development and Use of the Gill UVW Anemometer. Bound.-Layer Meteor., 8, 475-495.

Hanna, S. R., 1979: Some Statistics of Lagrangian and Eulerian Wind Fluctuations. J. Appl. Meteor., 18, 518-525.

Hinze, J. O., 1959: Turbulence, An Introduction to its Mechanism and Theory. McGraw Hill, New York, 586 pp.

Ivanov, V. N., and F. Ya. Klinov, 1962: Some Characteristics of a Turbulent Velocity Field in the Lower 300 Meter Layer of the Atmosphere. Bull. Acad. Science, USSR, 1026-1030.

Panofsky, Hans A., 1973: Tower Micrometeorology. Workshop on Micrometeorology (D. A. Haugen, Ed.), American Meteorological Society, Boston, 151-176.

Pasquill, F., 1962: Atmospheric Diffusion. Van Nostrand, New York, 297 pp.

Priestley, C. H. B., 1959: Turbulent Transfer in the Lower Atmosphere. University of Chicago Press, Chicago, 130 pp.

Stewart, D. A., 1981: A Survey of Atmospheric Turbulence Characteristics. US Army Technical Report RR-81-6, Redstone Arsenal, 23 pp.

Tennekes, H., and J. L. Lumley, 1973: A First Course in Turbulence. MIT Press, Cambridge, 300 pp.

Tukey, J. W., 1949: The Sampling Theory of Power Spectrum Estimates. NAVEXOS-P-735, Woods Hole. Office of Naval Research, 47-67.

Weber, H. H., J. S. Irwin, W. B. Petersen, J. J. Mathis, Jr., and J. P. Kohler, 1982: Spectral Scales in the Atmospheric Boundary Layer. J. Appl. Meteor., 21, 1622-1632.

Wyngaard, John C., 1973: On Surface-Layer Turbulence. Workshop on Micrometeorology (D. A. Haugen, Ed.), American Meteorological Society, Boston, 101-149.

8. ACKNOWLEDGEMENTS

The author appreciates with sincere thanks the contributions by the following individuals: Dr. Dorathy A. Stewart for her critical review and discussion of the manuscript; Mrs. Alexa M. Mims for the preparation of the computer programs to calculate the power spectra; Mr. Roger D. Betts for the computer program to derive the simulated data series; Mrs. Mona L. White for her patience in the preparation of the manuscript draft, her expeditious typing of the manuscript, and her editorial comments; Major James C. Weyman (Air Force Geophysics Laboratory) for furnishing the data sets from Otis Air Force Base.

NONINTEGER SPECTRAL ANALYSIS OF CLIMATOLOGICAL DATA

G. Padmanabhan

Department of Civil Engineering
North Dakota State University
Fargo, ND 58105, U.S.A.

and

A. Ramachandra Rao

School of Civil Engineering
Purdue University
West Lafayette, IN 47907, U.S.A.

1. INTRODUCTION AND OBJECTIVES

Blackman and Tukey (BT) and Fast Fourier Transform (FFT) methods of spectral analysis are commonly used in the analysis of climatalogical and hydrometeorological time series data. These methods have the disadvantage that by using them the spectral estimates may be computed only at specific frequencies. Consequently spectra computed by using these methods lack good definition, especially at low frequencies, which are of considerable interest to climatalogists and hydrologists. The frequencies at which spectra can be estimated depend on the maximum number of lags up to which autocorrelation coefficients are computed in BT method (Blackman and Tukey, 1959) and on the number of data points available in FFT method (Jenkins and Watts, 1968). In addition, these methods use subjectively selected window functions for smoothing. In order to use FFT method, data are often augmented with zeros. Some versions assume that the data repeats itself periodically (Jones, 1975). In the discrete Fourier transform used in FFT method, spectral estimates are obtained only at harmonic frequencies which are integer multiples of 1/n, n being the total number of data points. However, in both of these methods considerable spectral details may be missed in the frequencies which are between the frequencies at which spectral estimates are computed. Consequently, methods which do not have these limitations are being recently developed.

Noninteger (NI) method is one such method of spectral analysis which has not yet been tested with climatalogical data. The basic objectives of this study are to investigate the power of resolution and computational requirements of NI method by using observed climatalogical data and to compare the results with those obtained by using BT method. Some well-known time series which have been extensively analyzed in the past by using conventional methods are used in the study.

2. NONINTEGER METHOD

Noninteger method is a variation of the periodogram method of spectral analysis. However, unlike FFT, NI method is designed such that the spectrum can be estimated at nonharmonic frequencies also.

In the traditional spectral analysis methods based on discrete Fourier Transform only integer values are used for i in estimating the harmonic coefficients by using 1.

$$\left.\begin{aligned} a_i &= \frac{2}{n}\sum_{j=1}^{n} x_j \sin(2\pi ij/n),\ i = 1,2,\ldots[n/2] \\ b_i &= \frac{2}{n}\sum_{j=1}^{n} x_j \cos(2\pi ij/n),\ i = 1,2,\ldots[n/2]-1 \\ b_{[n/2]} &= \frac{1}{n}\sum_{j=1}^{n} x_j \cos(2\pi ij/n) \end{aligned}\right\} \quad (1)$$

In Eqs. 1, n is the number of observations, i is the harmonic number, [n/2] is the integer part of n/2, and a_i and b_i are the ith harmonic coefficients. The "raw" spectral estimate at the ith harmonic, or correspondingly at frequency i/n, is given by (2).

$$\left.\begin{aligned} S_i &= (a_i^2 + b_i^2)/2,\ i = 1,2,\ldots[n/2]-1 \\ S_{[n/2]} &= b_{[n/2]}^2 \end{aligned}\right] \quad (2)$$

The result in (2) is valid because of the orthogonality property of sine and cosine functions, which in turn is valid when integer values are used for harmonic number i, in (1). Equivalently, the fundamental period T must be an integral multiple of the harmonic periods. Therefore the harmonics will have periods T, T/2, T/3,..., T/[n/2]. Correspondingly, on the frequency scale spectral estimates are computed at frequencies 1/T, 2/T, 3/T,..., [n/2]/T. It is quite possible that considerable interest may lie in frequencies

which are between the harmonic frequencies. Spectral estimates at these frequencies can be obtained if noninteger values are used for i in (1). However, when noninteger values are used for harmonic number, equations in (1) are not valid.

NI method uses a multiple regression approach to obtain spectral estimates at all frequencies. The methodology used in NI spectral analysis is as follows. The coefficient of determination R_i^2 of the regression of sine and cosine terms XS_{ij} and XC_{ij} associated with the ith wave form of the data is the explained variance of the data due to that wave form. XS_{ij} and XC_{ij} are the jth ordinates of sine and cosine components of the ith wave form. The explained variance R_i^2, is the spectral contribution at frequency $f_i = i/n$ corresponding to the ith wave form. In the NI method, XS_{ij} and XC_{ij} defined below are computed first. In defining XS_{ij} and XC_{ij}, i is harmonic number which can take integer or noninteger values.

$$\left.\begin{aligned} XS_{ij} &= \sin(2\pi ij/n); \\ XC_{ij} &= \cos(2\pi ij/n); \end{aligned}\right\} \quad j = 1,2,\ldots n \qquad (3)$$

For any value of i, the period of the wave form is n/i and the corresponding frequency f_i is given by i/n. XS_{ij} and XC_{ij} computed above are used in the regression relationship 4 to obtain $\hat{X}_j$ which are estimates of X_j. R_i^2 is computed by using (5).

$$\hat{X}_j = a + b_{si}\,(XS_{ij}) + b_{ci}\,(XC_{ij}), \quad j = 1,2,\ldots,n \qquad (4)$$

$$X_{ij} = \hat{X}_j + n_i$$

$$R_i^2 = \sum_{j=1}^{n} (\bar{X}_j - \hat{X}_j)^2 \Big/ \sum_{j=1}^{n} X_j^2 \qquad (5)$$

Estimating the regression coefficients are of secondary concern. Of greater utility are the estimates of coefficient of determination R_i^2 due to regression. The coefficients b_{si} and b_{ci} of the wave forms considered above enable one to recombine the wave forms to predict X_j. In a similar manner, R_i^2 can be estimated for as many wave forms as needed at desired frequencies. Distribution of R_i^2 with frequency f_i associated with ith wave form gives the variance spectrum. The computations involved are shown in the flow chart (Fig. 1). However, the spectrum thus obtained will be approximate because the estimate of R_i^2 is approximate when XS_{ij} and XC_{ij} are non-orthogonal or correlated. This occurs whenever noninteger values are used for harmonic numbers, although the correlation between XS_{ij} and XC_{ij} is usually small.

Since the NI approach is based on multiple regression, well-known tests of significance can be used to test the significance of correlation between XS_{ij} and XC_{ij} and to test R_i^2 itself. Since R_i^2 can vary between 0 and 1 it corresponds to the normalized spectral density. The non-normalized spectral density is given by $2\,R_i^2\,S_x^2\,(n-1)/n$ where S_x^2 is the variance of X_j, j=1,2,...,n and n is the number of observations. The computational details can be found elsewhere (Schiekedanz and Bowen (1977), Padmanabhan and Rao (1980 and 1982)).

3. DATA USED IN THE STUDY

Data used for spectral analysis in the present study include the well-known time series of Zurich Sunspot numbers and Central England temperature series which have been extensively analysed by several earlier investigators. Hydrologic time series of annual rainfall measured at several stations, monthly river flows and monthly and annual drought index are also used. These data are listed in Table 1. The hydrologic time series used in the study come from different geographic locations.

Central England rainfall and Kew rainfall constitute a very long time series and are analysed to investigate long term oscillations in them. The spectra of Kelburn rainfall computed by using traditional methods indicate several ususual periodicities. In order to test these periodicities in Kelburn rainfall, it was analysed further. The Waltair and Mangalore rainfall data are analysed to detect the presence of biennial oscillations present in Indian rainfall. Palmer's Drought Index series for Kansas and Iowa are analysed to delineate their long term behavior. Krishna and Godavari river flows are included for their highly periodic nature. Wold series which are generated by using autoregressive models are analysed to test the computational accuracy of the method. The models from which Wold series are generated are given at the bottom of Table 1. The plots of some of the data series are shown in Fig. 2.

4. RESULTS

In this study, conventional spectra are computed by using Blackman and Tukey mtheod. Hamming window is used for smoothing. Conventional Spectra computed by using autocorrelations up to lag m are referred to as CS(m) in this discussion. Examples of NI and conventional spectra are shown in Figs. 3 and 4 for comparison.

Since Wold series 1,2 and 3 are synthetically generated by using autoregressive models of known coefficients (table 1), their true spectra could be calculated (Jenkins and Watts (1968)). The true, conventional and NI spectra of Wold series 1,2 and 3 are shown in Fig. 3. The true spectra have peaks at 0.167 cycles/Δt, with the peaks becoming sharper from series 1 to 3. Conventional spectra of orders as high as CS(50) are required to reasonably reproduce the true spectra. The NI spectra, however, clearly bring out the prominent peak. NI spectrum of series 1 appears to be highly variable.

The well-known 11-year cycle in the Zurich sunspot data is clearly brought out in the NI spectrum (Fig. 4). The corresponding peak in the NI spectrum is sharper them even CS(50) shown alongside. The splitting of spectral density estimate around 11-year period may be due to changes in periodicity of sunspots over the duration spanned by the data. Also an 80-year cycle is clearly brought out by the NI spectrum of sunspot data whereas it is not so obvious in the CS estimate.

TABLE 1. Data Used for Spectral Analysis and Modeling.

No.	Data	Abbreviation Used	Number and Type of Observations	Period	Source
1	Sunspots	SSPOT	316(Y)	1660-1975	Schonweise (1978)
2	Beveridge's Wheat Index	BWI	370(Y)	1500-1869	Anderson (1971)
3	Central England Temperature	CET	315(Y)	1660-1974	Schonweise (1978)
4	Central England Rainfall	CER	244(Y)	1727-1970	Meteorological Magazine (1971)
5	Kew (England) Rainfall	KEW	274(Y)	1697-1970	Wales-Smith (1973)
6	Kelburn (New Zealand) Rainfall	KELB	104(Y)	1872-1975	New Zealand Meteorological Service Annual Reports
7	Waltair (India) Rainfall	WALT	75(Y)	1866-1940	Smithsonian Miscellaneous Collections
8	Mangalore (India) Rainfall	MANG	88(Y)	1853-1940	Smithsonian Miscellaneous Collections
9	Kansas Drought-Index	KAN	76(Y) & 912(M)	1887-1962	Palmer (1965)
10	Iowa Drought-Index	IOW	33(Y) & 396(M)	1930-1962	Palmer (1965)
11	Godavari (India) Flow	GOD	720(M)	1901-1960	UNESCO (1971)
12	Krishna (India) Flow	KRI	720(M)	1901-1960	UNESCO (1971)
13	Wold Series - 1	WOLD1	200(S)	-	Anderson (1971)
14	Wold Series - 2	WOLD2	200(S)	-	Anderson (1971)
15	Wold Series - 3	WOLD3	200(S)	-	Anderson (1971)

Y: Annual data M: Monthly data S: Synthetic data

Wold Series have been generated by using the following AR models:

WOLD1: $y(t) = 0.25\ y(t-1) - 0.0625\ y(t-2) + n(t)$
WOLD2: $y(t) = 0.7\ y(t-1) - 0.49\ y(t-2) + n(t)$
WOLD3: $y(t) = 0.9\ y(t-1) - 0.81\ y(t-2) + n(t)$; $n(t) = N(0,1)$

CS(50) and NI spectra of Central England annual temperature, annual rainfall and Kew annual rainfall are also shown in Fig. 4. CS(50) of central England temperature data indicates long term oscillations but the components are not clearly resolved. However, NI spectrum clearly shows a cycle of approximately 100-year period in the central England data. Central England rainfall data show concentration of variance in two frequency ranges - one around a low frequency of 0.02 cycles/year and the other around 0.42 cycles/year. These two components are clearly resolved in the NI spectrum. For Kew rainfall no significant peaks are apparent from CS(50). However, NI spectrum, although appearing to be highly variable shows a discernible peak at 0.42 cycles/year frequency.

Conventional spectra of Kelburn and Waltair annual rainfall series indicate multiple peaks. The details are clearly evident for these data at all frequency ranges in the corresponding NI spectra. The results are similar for the spectra of annual rainfall from Mangalore, although multiple peaks are not quite evident.

Both CS(20) and NI spectra of Iowa and Kansas drought index data show a prominent peak near 20-year period. However, the frequency resolution is higher in the NI spectra.

Krishna and Godavari monthly river flows have strong yearly periodicity and this is clearly seen in the conventional spectra. All the peaks seen in CS(100) also appear in NI spectra.

5. DISCUSSION

The results of this study indicate that the NI method is very sensitive and can indicate the details in the low frequency region of the spectrum much better than the conventional methods. Subjective choice of window functions is avoided in this method and the computions are quite simple.

In BT method, the number of spectral estimates is controlled by the specified maximum lag. Resolution can be improved by increasing the maximum number of lags up to which autocorrelations are estimated, resulting in spectral estimates at close frequency intervals. However, at larger lags, the autocorrelation estimates are weak and hence will introduce instability in the spectrum. In the NI method the spectrum is computed directly from the data at any desired frequency. There is no need

of windowing in NI method, since the regression method used in estimating the spectra yields smooth estimates. It should be noted that in the NI method, stronger the correlation between sine and cosine functions for a particular wave form, the less reliable are the spectral estimates at the corresponding frequencies. Since the method principally uses multiple regression, standard signficance tests can be used to test the correlation between sine and cosine terms of a given wave form and based on the results, spectral estimates at certain frequencies can be discarded. However, in the results reported in this study, no estimate was discarded. In addition, as only a few wave forms are available for sampling at the low frequency range, the spectral estimates at low frequencies are not very reliable. However, this will be a serious problem only in very short samples. For moderately long series, results from NI method are quite dependable.

6. ACKNOWLEDGMENT

The authors wish to acknowledge that the Non-integer Spectral Estimation Computer Program developed by Dr. P.T. Schiekedanz and Ms. E.G. Bowen has been used in this study. The Atmospheric Sciences Section of Illinois State Water Survey, USA provided the computer program.

7. REFERENCES

Anderson, T.W. "The Statistical Analysis of Time Series", Wiley, New York, 1971.

Blackman, R.B. and J.W. Tukey, "The Measurement of Power Spectra", Dover, New York, 1959.

Jenkins, G.M. and D.G. Watts, "Spectral Analysis and its Applications", Holden-Day, San Francisco, 1968.

Jones, R.H., "Spectrum Estimation and Time Series Analysis - A Review", 4th Conference on Probability and Statistics in Atmospheric Sciences, Tallahassee, Florida, American Meteorological Society, pp. 35-39, 1975.

Meteorological Magazine, 100, p. 362, 1971.

Padmanabhan, G. and A.R. Rao, "Stochastic Analysis and Modeling of Hydrologic and Climatologic Time Series". I."Maximum Entropy Spectral Analysis of Hydrologic and Climatalogic Data", Tech. Rept. CE-HSE-80-1, pp. 185, 1980. III. "Computer Programs for Analysis and Modeling of Hydrologic Time Series", Tech. Rept. CE-HSE-82-9, pp. 126, 1982. School of Civil Engineering, Purdue University, W. Lafayette, IN 47907, USA.

Palmer, W.C., "Meteorological Drought", Research Paper No. 45, Weather Bureau, U.S. Department of Commerce, Washington, D.C., pp. 1-58, 1965.

Schiekedanz, P.T. and E.G. Bowen, "The Computation of Climatological Power Spectra", Journal of Applied Meteorology, Vol. 16, No. 4. pp. 359-367, April 1977.

Schonwiese, C.B. "Central England Temperature and Sunspot Variability 1660-1975", Arch. Met. Geoph. Biokl. Series - B, 26, pp. 1-16, 1978.

"Smithsonian Miscellaneous Collections", Smithsonian Institution, Vol. 79, 1927, Vol. 90, 1934 and Vol. 105, 1947.

UNESCO, "Discharge of Selected Rivers of the World", Vol. II, Paris, 1971.

Wales-Smith, B.G., "An Analysis of Monthly Rainfall Totals Representative of Kew, Survey from 1967-1970", Meteorological Magazine, Vol. 102, pp. 157-171, 1973.

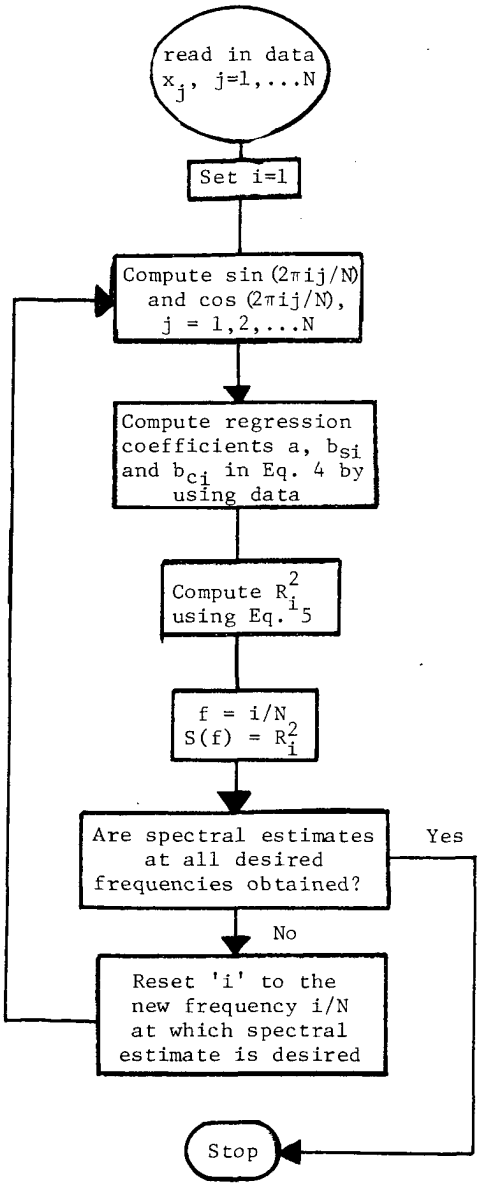


FIGURE 1. Flow Chart for Computations in NI Method.

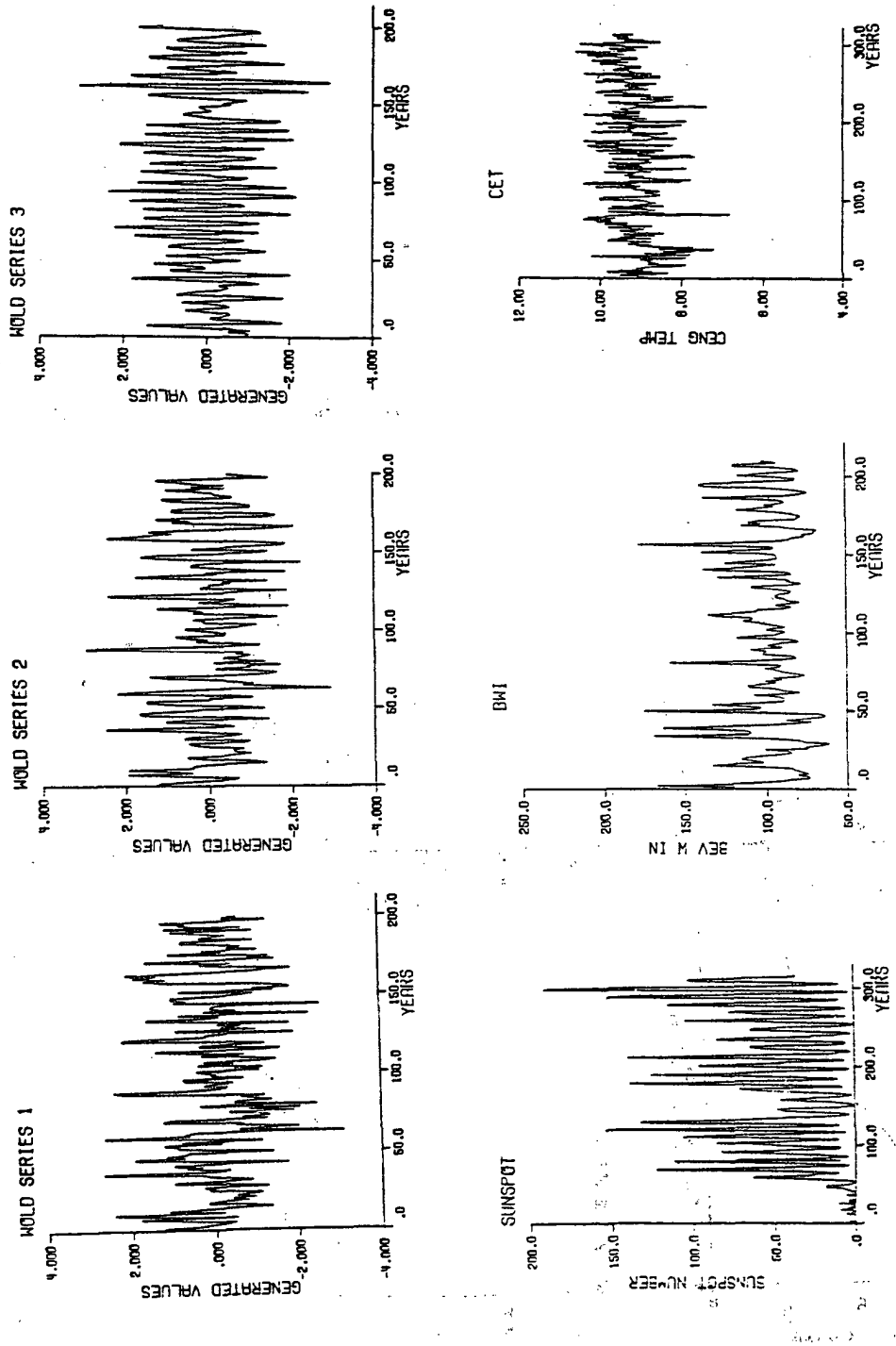


FIGURE 2. Plots of some of the data used in the study.

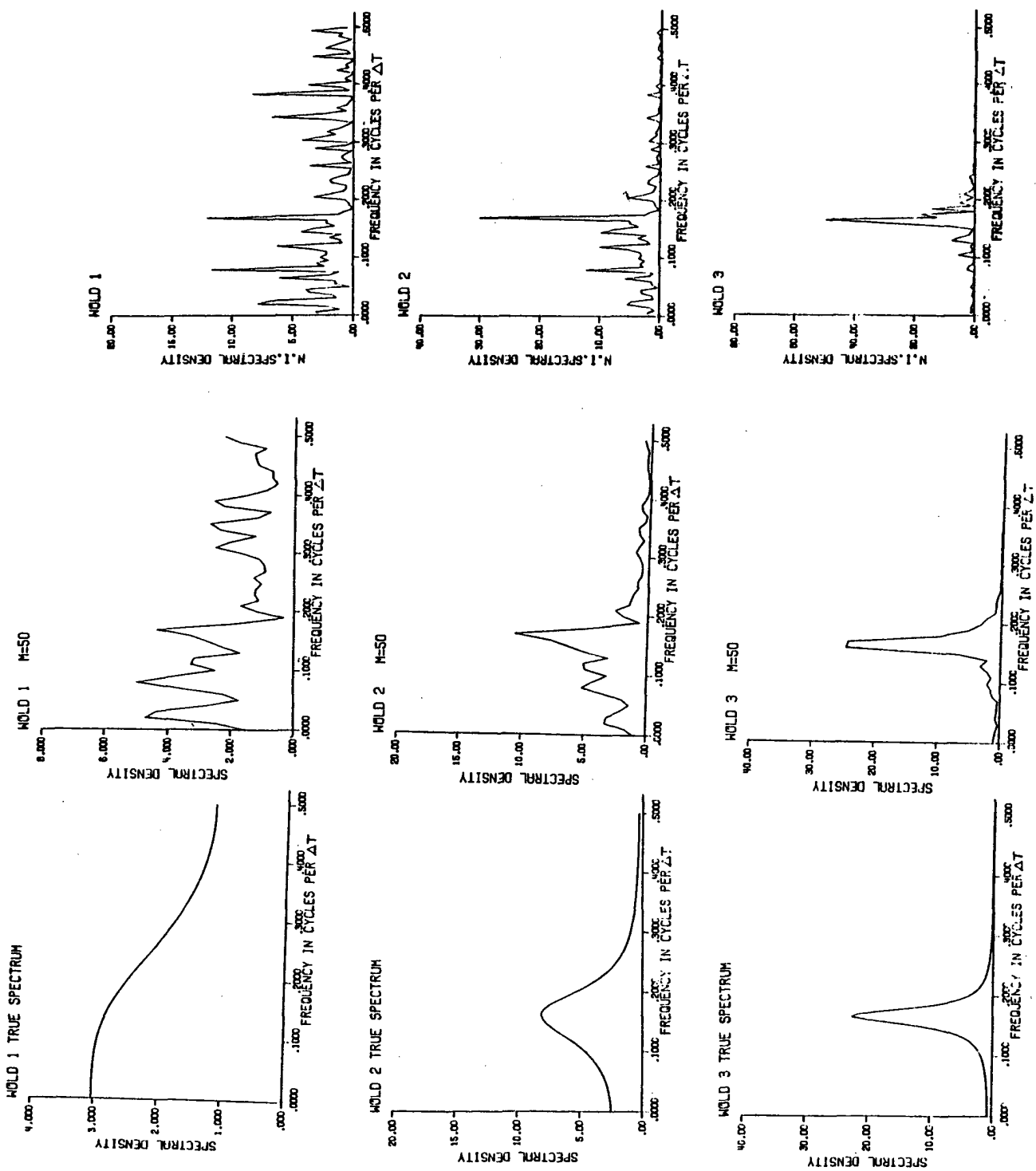


FIGURE 3. True (I column), Conventional (II column) and noninteger spectra of Wold series.

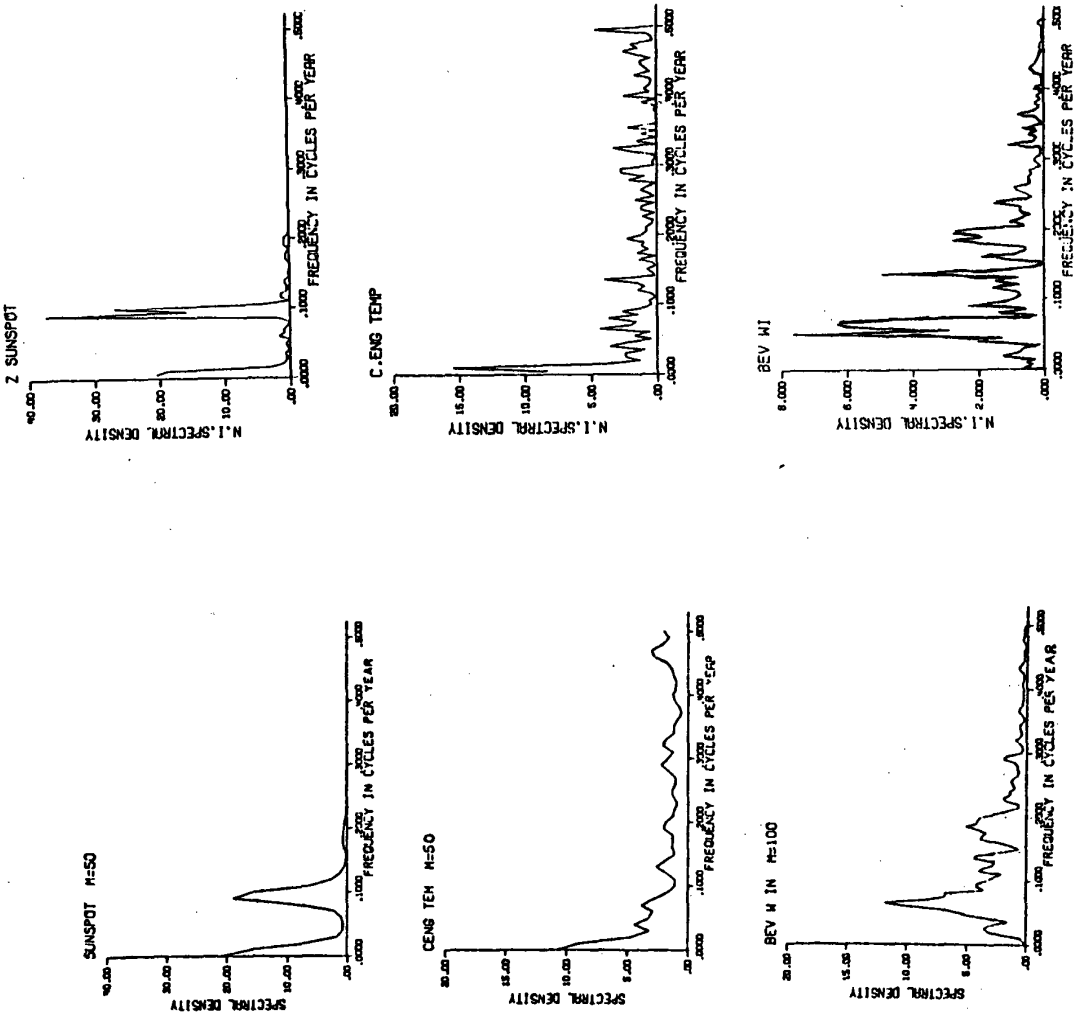


FIGURE 4. Conventional (I column) and Noninteger (II column) spectra of sunspot, central England temperature and Beveridge Wheat Index Series.

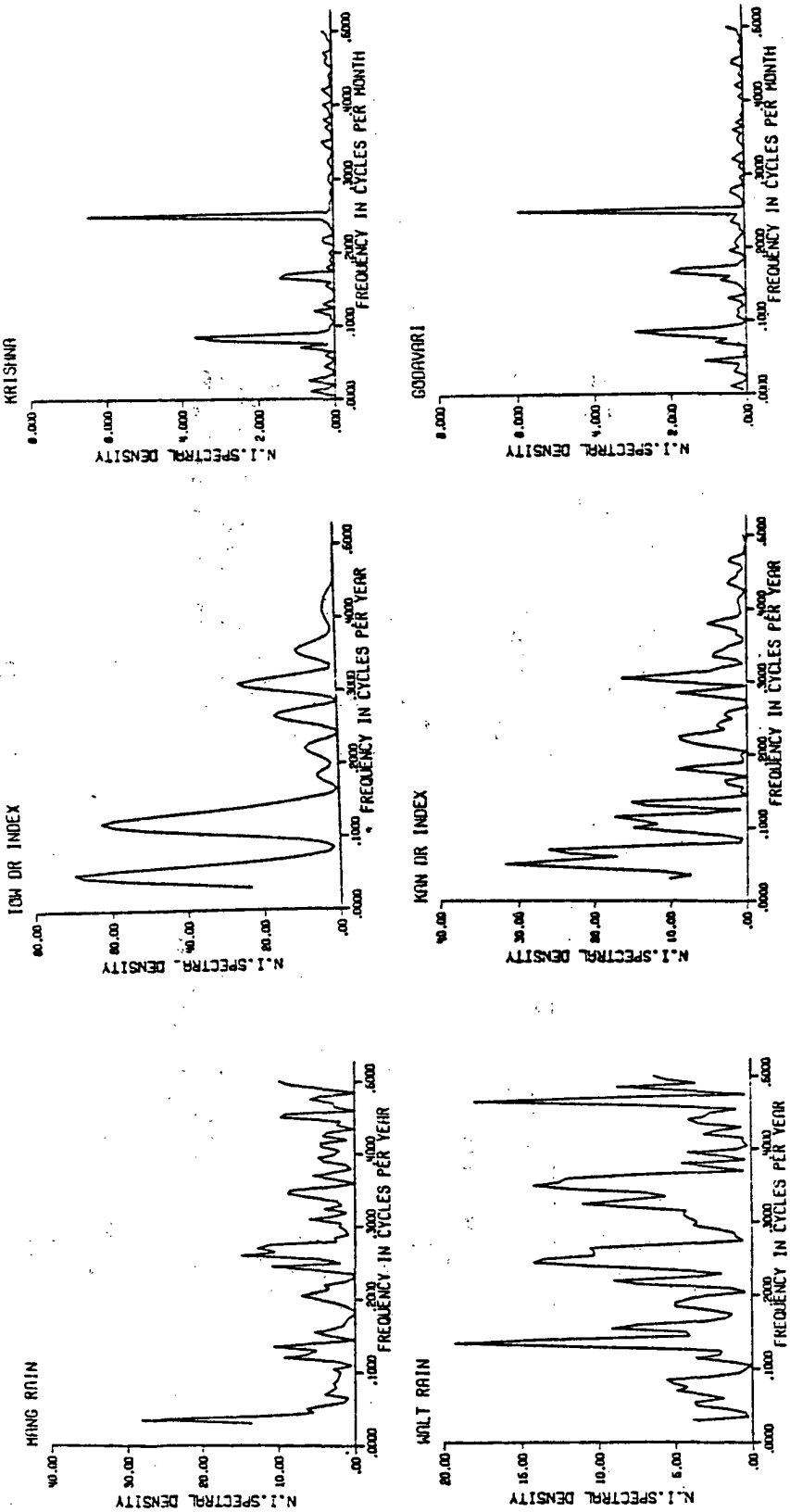


FIGURE 4. (contd.) Noninteger spectra of annual rainfall at Mangalore and Waltair, Palmer's drought Index for Iowa and Kansas and monthly Krishma and Godavari river flows.

SPECTRAL ANALYSIS OF SUMMER PRECIPITATION SERIES

Rosa H.Compagnucci

Instituto Argentino de Nivología y Glaciología (IANIGLA)
C.C.330, 5500 Mendoza, Argentina

Walter Mario Vargas
Facultad de Ciencias Exactas y Naturales, Dpto.de Meteorología
Universidad de Buenos Aires
1428 Nuñez, Buenos Aires, Argentina

1. INTRODUCTION

Summer rainfall series in the middle-west zone in Argentina, known also as Cuyo, have special characteristics in regard to its variability.

Compagnucci and Boninsegna (1979) studied the precipitation records of the Mendoza Observatory. This study showed the various characteristics in the series of precipitations: winter rainfall -total amount of rain fallen from April to September-, and the summer series -total precipitation from October to March-. While the first presents random characteristics as a result of the different aleatory tests used, the second, the summer series, is not fully represented by this process. This fact can not be attributed to the existence of lineal tendency as none of the tests proved a positive result in this sense. Sneyers' test and the sign test for the median, showed instead the existence of possible significant fluctuations.

In the analysis of total summer rainfall, obtained in eight different stations in the Cuyo region, specially chosen because of their reliable and extensive records, Compagnucci et al(1982) observed significant and quasi-periodical variations. The results of smoothing the total summer precipitation series, with a 10-term Gauss low-pass filter, showed that the long-wave fluctuations have a quasi-periodicity of 18 years and the series are in phase (Fig.1). The percentual deviations are similar in the series for the maximums and minimums.

A discussion on summer precipitation fluctuations is affected by: instrumental records of only 100 years long (approximately), great year to year precipitation variability; little density of the area network; lack of continuity of some series. Nevertheless, we study raw summer precipitation series of Cuyo applying spectral analysis to them, in order to prove the existence of quasi-periodical fluctuations in the series.

These results are also compared with those obtained by Tyson and Dyer (1978) and Tyson (1980) in South Africa.

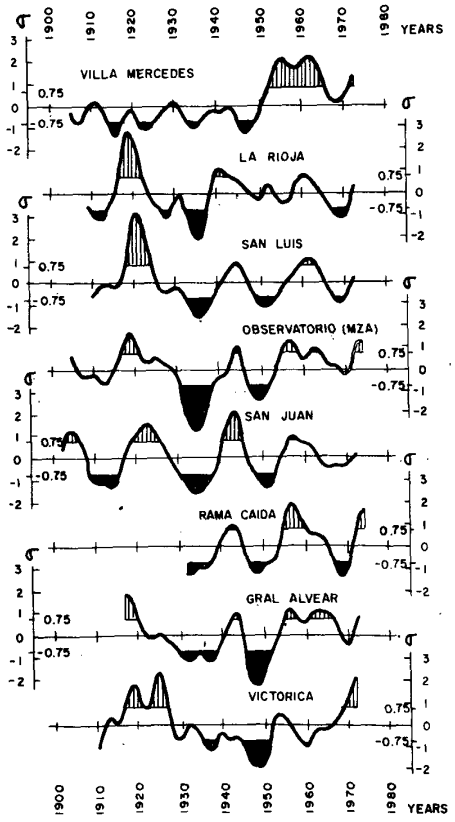


Fig. 1: SERIES OF SUMMER PRECIPITATION IN CUYO, SMOOTHING THE DATA USING 10-TERM GAUSS LOW-PASS FILTER - BLACK AREAS ARE MINIMUNS LESS THAN -0.75 SHADED AREAS ARE MAXIMUMS MORE THAN +0.75

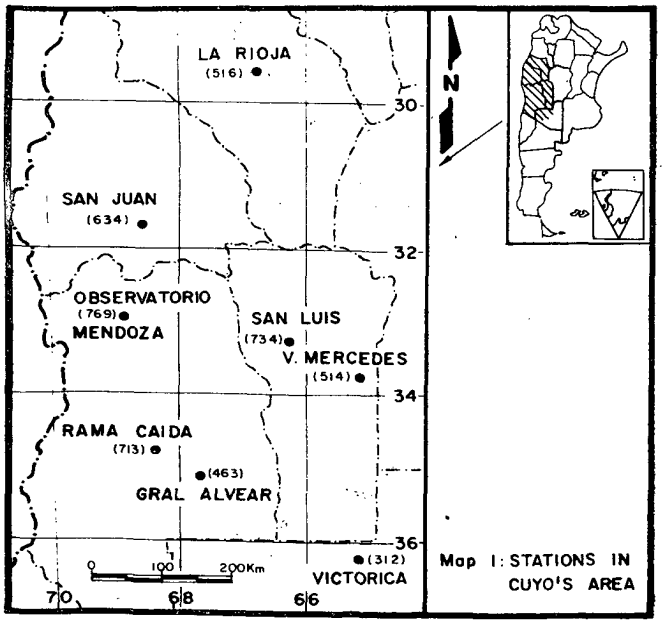
2. DATA

The data used in this paper were provided by the "Servicio Meteorológico Nacional, Fuerza Aérea Argentina" (National Meteorological Service, Argentine Air Force)and by the "Instituto Nacional de Tecnología Agropecuaria, I.N.T.A."(National Institute of Agricultural Technology).

The series analysed were the following:

STATIONS	RECORD
Colonia Alvear	1912-1977
La Rioja	1904-1977
Mendoza Observatorio	1900-1978
Rama Caída	1927-1979
San Juan	1897-1977
San Luis	1905-1977
Victorica	1905-1977
Villa Mercedes	1900-1977

The map (1) shows the area and stations studied in this paper.

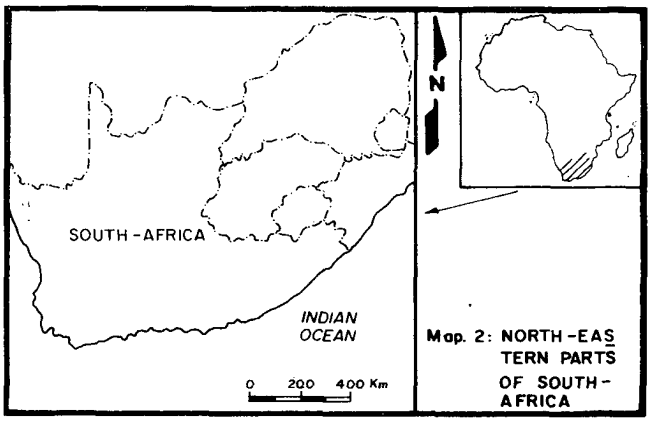


Map 1: STATIONS IN CUYO'S AREA

Summer precipitations spatial average series of the north-eastern parts of South Africa, from Tyson's and Dyer's (1978) works,(Fig.11),was also included as data. It was obtained by calculating the 1910-1977 mean for each of the 62 stations,and later, the percentage of the 1910-1977 mean of a yearly summer rainfall average across the 62 stations. Map (2) shows the South Africa region over which the mean time series has been determined (Dyer and Tyson, 1977).

Tyson and Dyer infer in their paper that this series is representative of the South African region.

With the same methodology the precipitation spatial average series of the Cuyo region was calculated (Fig.11).



Map. 2: NORTH-EASTERN PARTS OF SOUTH-AFRICA

3. METHODS OF ANALYSIS

The raw series of summer precipitations have been analysed using the following procedures:

3.1 Trend

Linear trend has been determined by conventional least squares regression analysis, Panofsky (1968), and the significance of the slope tested using a test t-Student, Johnston (1977).

3.2 Serial Correlation

Serial correlations r(k) have been determined from:

$$r(k) = \frac{\sum_{t=1}^{N-k} \left[X(t) - \bar{X}\right] \cdot \left[X(t+k) - \bar{X}\right]}{N \cdot S_x^2}$$

where k=1,.....,M where M is the maximum lag and S_x^2 is the variance of the series.

3.3 Spectral Analysis

Estimates of smoothed spectral density R(f) have been determined from raw data, according to Jenkins and Watts (1968), using:

$$\bar{R}(f) = 2\left\{1 + 2\sum_{k=1}^{L-1} r(k)\, w(k) \cdot \cos\left(2\pi f k\right)\right\}$$

$$0 \leq f \leq \frac{1}{2}$$

where r(k), w(k) and L are serial correlations, the lag window, and the truncation point respectively; and choosing 2f=1/M and M=L-1, we have

the following equation:

$$\bar{R}(f) = 2\left\{1 + 2\sum_{k=1}^{M} r(k).w(k).\cos\frac{k\pi}{M}\right\}$$

The Parzen's window was used, Jenkins and Watts (1963-1968):

$$w(k)\begin{cases} 1 - 6\left(\frac{k}{M}\right)^2 + 6\left(\frac{k}{M}\right)^3 & 0 \leq k \leq \frac{M}{2} \\ \text{for} & \\ 2\left(1 - \frac{k}{M}\right)^3 & \frac{M}{2} \leq k \leq M \end{cases}$$

The choice of M=1/3 (N= total length of the series), was determined by following the suggestions of Mitchell et al (1977), in an attempt to maximize and minimize spectral detail and minimize spectral inestability.

Significance of spectral estimates is determined by using the null hypothesis of a first order Markov generating process, as outlined by Mitchell et al (1977).

The spectral densities and the significant level was determined for the raw series as well as for the series without tendency. The before said was applied to each of the eight series and to the average spatial precipitations series of the Cuyo region and South Africa.

3.4 Relation or dependency test

In order to analyze the relation between the precipitation series of Cuyo and South Africa, the following methods were applied:

3.4.1 Contingency table

The contingency table analysis is made for square coefficient contingency (X^2). Applying the hypothesis of independence, X^2 is asymptotically distributed in the χ^2 form, Kendall and Stuart (1977). The degrees of freedom are given by:

$(rc-1)-(r-1)-(c-1) \circ (r-1)(c-1)$

n_{11}	..... n_{12}	..... n_{1j}	..... n_{1c}	$n_{1.}$
n_{i1}	..... n_{i2}	..... n_{ij}	..... n_{ic}	$n_{i.}$
n_{r1}	..... n_{r2}	..... n_{rj}	..... n_{rc}	$n_{r.}$
$n_{.1}$	..... $n_{.2}$	..... $n_{.j}$	..... $n_{.c}$	n

Contingency table

$$X = n\sum_{ij}\left(\frac{n_{ij}}{n_i\ n_j} - 1\right)$$ Square coefficient contingency

Besides, Pearson's coefficient of contingency (P), maximum coefficient or complete asociation (P_{max}), and the corrected coefficient of contingency (C_c) were calculated as follows:

$$P = \frac{X^2}{n + X}^{1/2}$$ Pearson's coefficient

$$P_{max.} = \frac{r-1}{r}^{1/2}$$ Maximum coefficient or complete association

$$Cc = P / Pmax$$ Corrected coefficient

3.4.2 Correlation coefficient and the significance level

If $r > r_c$ for the significance level given by t_g, the null hypothesis that r=0 is disregarded (Mitchell et al,1977).

$$r = \sum_{i=1}^{n} \frac{(x_i - \bar{x}).(y_i - \bar{y})}{n.S_x.S_y}$$

$$r_c = \frac{-1 \pm tg\sqrt{n-2}}{n-1}$$

4. RESULTS

4.1 The result of calculating the tendency and of the tests applied to the slope to verify if they are significantly different from cero, proves

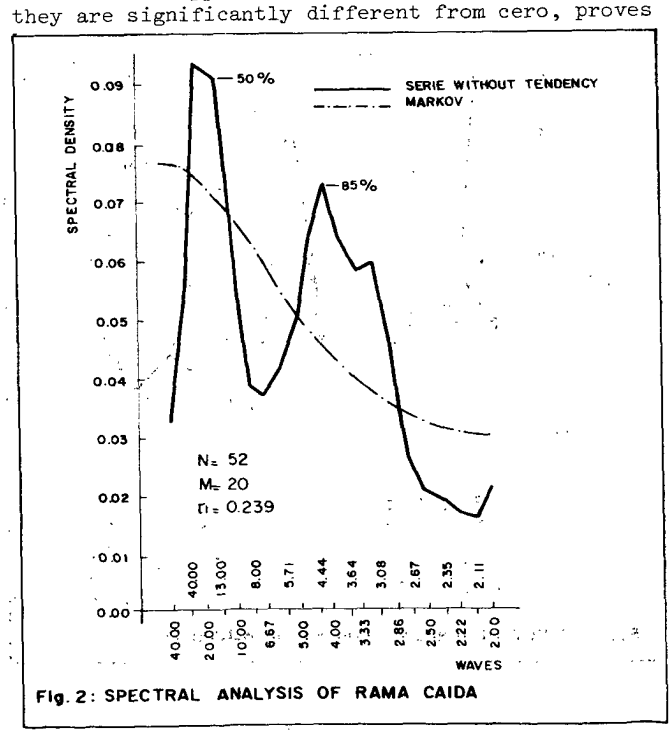


Fig. 2: SPECTRAL ANALYSIS OF RAMA CAIDA

that only Rama Caída at 95% level and Villa Mercedes at 99% present a tendency, which in both case is positive.

The explanation in the case of Rama Caída (Fig.2) could be that this series is a shorter record than the others and has been started during one of the dryer periods (Fig.1) ending in one of the damper periods. This might explain the tendency in this series which does not come up in the other one. This hypothesis is strengthen by the fact that the spectral analysis (Fig.2)with or without tendency is the same, the maximum in long waves for the spectrum of raw data do not appear as it would in the case of series dominated by tendency.

The case of Villa Mercedes is different.The spectrum (Fig.3) of the raw series shows a maximum in longer waves which does not come up in the spectrum of the same series with no tendency.This is a sign of a real tendency in the series possibly due to the change of the station location occurred in 1950. In the paper Compagnucci et al (1982) an interruption is pointed out in the double mass curve for this particular station and year. Moreover a difference can be observed in the data for the filtered low-pass series (Fig.1) between 1900-1950 and 1951-1977 periods.

4.2 The variance in the series now being studied could be explained mostly by random process. However there is a part, sometimes a considerable one that could be explained by quasi-periodical effects and of persistence.

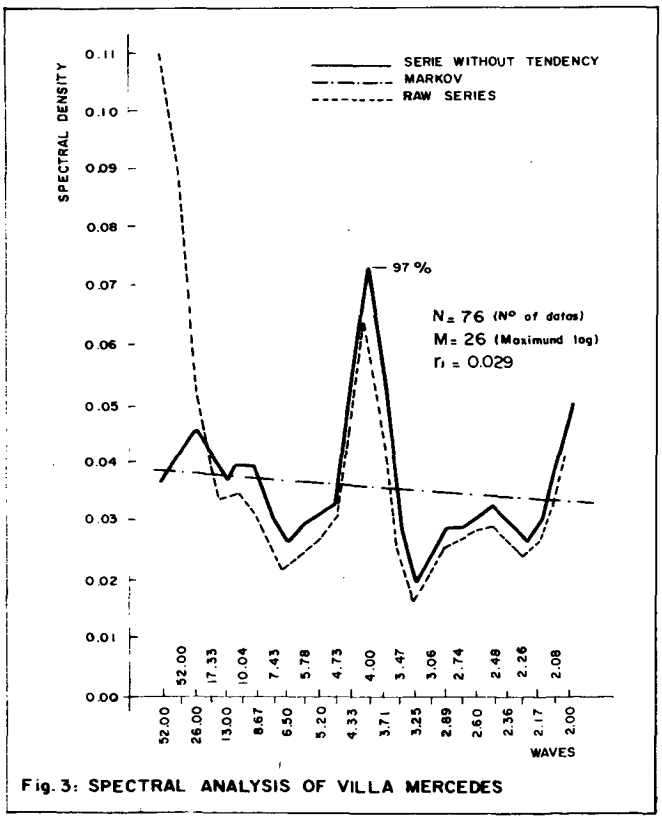


Fig. 3: SPECTRAL ANALYSIS OF VILLA MERCEDES

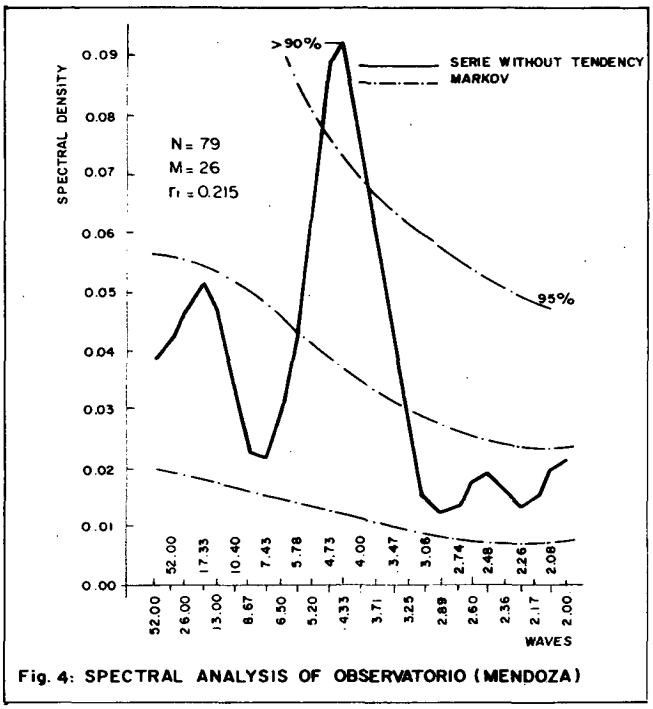


Fig. 4: SPECTRAL ANALYSIS OF OBSERVATORIO (MENDOZA)

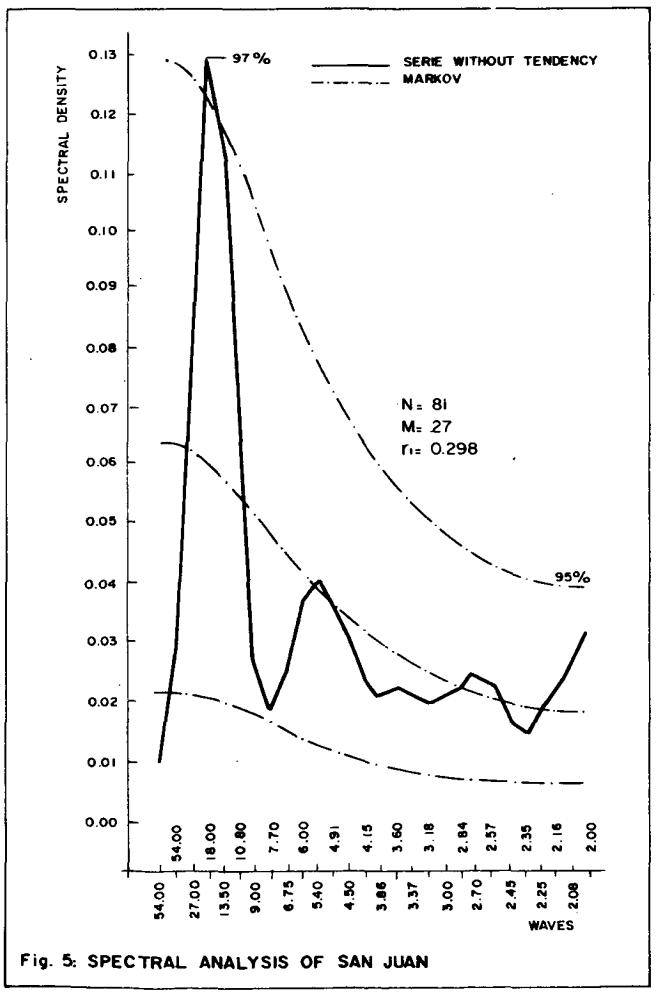


Fig. 5: SPECTRAL ANALYSIS OF SAN JUAN

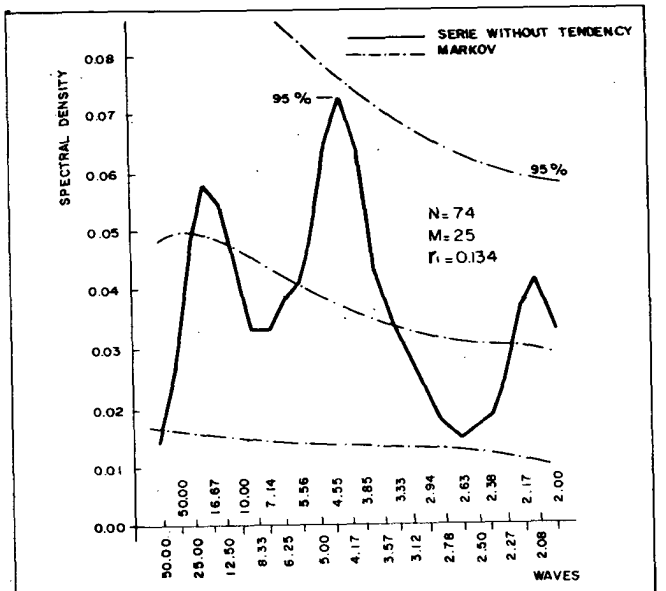


Fig. 6: SPECTRAL ANALYSIS OF LA RIOJA

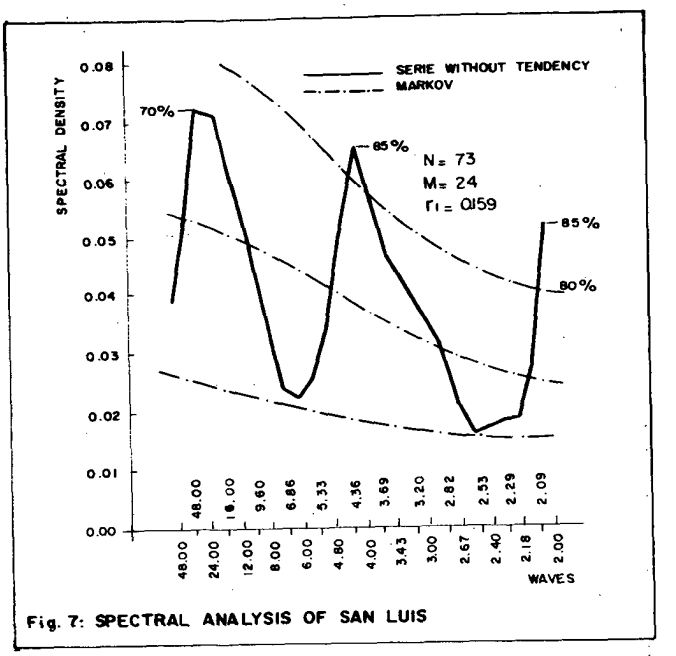


Fig. 7: SPECTRAL ANALYSIS OF SAN LUIS

The first autocorrelation coefficient is 90% significant for San Luis and more than 95% for San Juan, Rama Caída and Observatorio. Having accepted the significant level, the series show not only random processes but also a persistence process of the first order, at least.

Victorica, San Luis and San Juan series also present markovian characteristics as they prove that $r_1^2 = r_2$. In the particular case of San Juan, r_1 and r_2 are more than 95% significant, and moreover $r_1^2 = r_2$ and $r_1^3 = r_3$; in this case the persistence in the series is of second order.

Alvear presents negative r_1 and r_2, which means that it is a highly fluctuating series. In the case of the Observatorio, Rama Caída, Alvear and La Rioja, r_2 is negative and in San Luis r_2 is not very significantly different from cero. On the other hand, these series present a significant maximum spectral for the frequencies centered in 4 years. The negative or null value of r_2, can be explained because it means the possible correlation of half a wave in 4 years.

4.3 Generally the spectrum of the series that are being studied present a maximum quasi-periodical fluctuation of 18 years, another of 4 years and a very dim one of approximately 2 years.

It is possible that the maximum of 4 years as well as that of 2 years could be an indication of the biannual oscillation; in San Luis (Fig.7) it has 85% significance and is faintly outlined in other stations as La Rioja (Fig.6), Victorica (Fig.8), Villa Mercedes (Fig.3), San Juan (Fig.5) and Alvear (Fig.9).

The stations can be divided in three sets due to the difference between the spectrums.

Victorica (Fig.8), located very much in the southeast of the Cuyo area, does not present a maximum centered in 18 years, but has maximums at 4 and 2 years which are not very significant compared with Markov.

A second group made out of the series that prove a significant maximum for almost 4 years, taking into account the markovian hypothesis is formed by Observatorio (Fig.4), Alvear (Fig.9), Rama Caída (Fig.2) and La Rioja (Fig.6). Moreover, they present a small maximum of roughly 18 years.

The stations of San Luis (Fig.7) and San Juan (Fig.5) form a third group which has a very significant maximum of quasi-periodical 18 years. In the case of San Juan there is a 98% significance.

4.4 Tyson and Dyer (1978) postulated that the spectral analysis of the average-space series of summer precipitations of South Africa, present very significant quasi-periodical fluctuations proving three important maximums in the spectrum of the series.

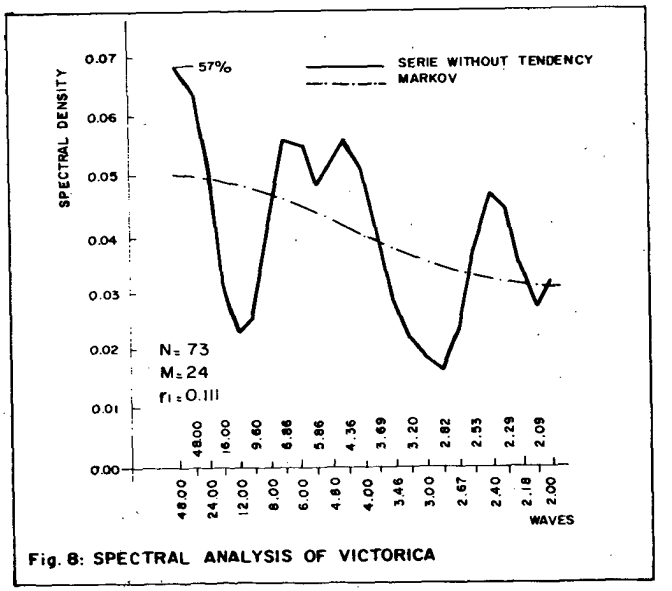


Fig. 8: SPECTRAL ANALYSIS OF VICTORICA

The most significant, between 16 and 22 years another between 3.5 and 4 years and a third one between 2.5 and 2 years.

The average spatial series of middle-west in Argentina, presents maximums in the same quasi-periods, as can be observed in Fig.10.

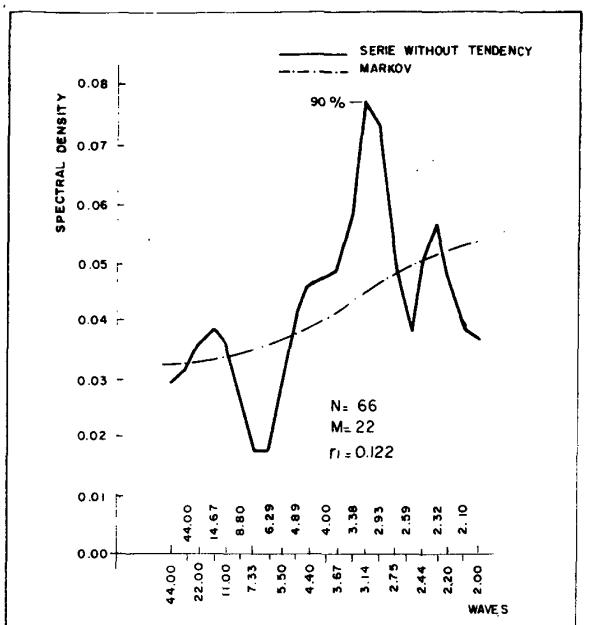


Fig. 9: SPECTRAL ANALYSIS OF GENERAL ALVEAR

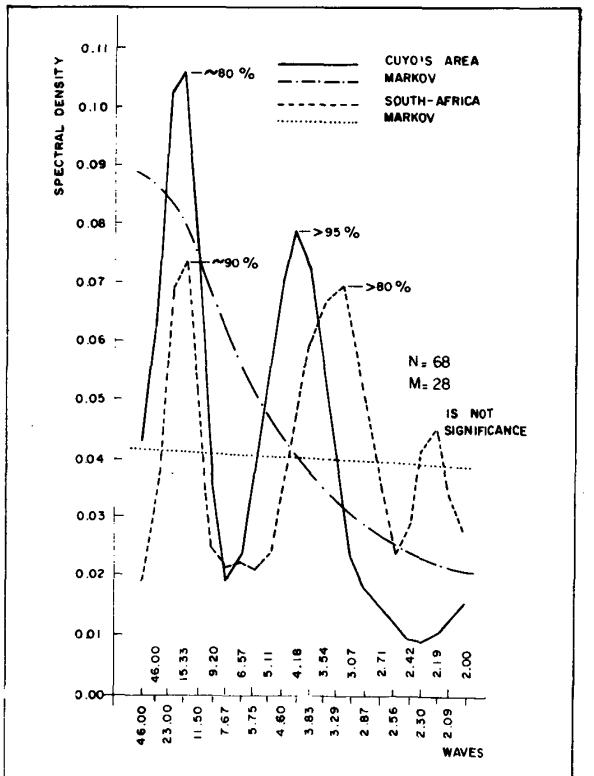


Fig. 10: SPECTRAL ANALYSIS OF SUMMER SPATIAL-AVERAGE SERIES OF CUYO'S AREA AND SOUTH-AFRICA.

In the graphic of the Fig.10 the full line draws spectral density of average spatial series of Cuyo, and the dotted lines, the results of spectral analysis in the South African series, both for 1910-1977;they were calculated by the spectral analysis of Tukey with window of Parsen taking M=23 (maximum lag).

In the diagram, the two curves present a maximum between 2.3 and 1.5 years, belonging to the quasi-period of 18 years as well as the maximum centered in 4 years.

The maximum in South Africa is between 3.5 and 3.2 wave lengths, while in the Cuyo series it is between 4.6 and 3.8.

The difference presented by both spectrums comparing them with Markov is that in the Cuyo case the maximum of 4 years has a higher significance level (95% probability) than the South Africa case (80% probability).

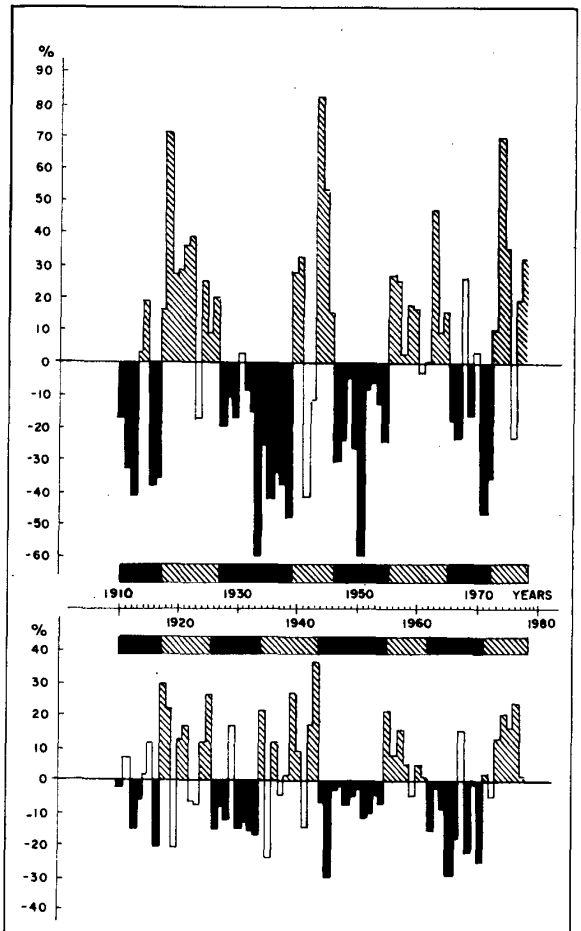


Fig. 11: SUMMER PRECIPITATION SPATIAL-AVERAGE SERIES OF CUYO'S AREA IS SHOWN IN THE UPPER SECTION OF THE GRAPHIC. MEAN ANNUAL RAINFALL OF NORTH-EASTERN PARTS OF SOUTH-AFRICA FOR THE SUMMER RAINFALL REGION, AS A WHOLE, LOWER SECTION. BOTH EXPRESSED AS A PERCENTAGE DEVIATION FROM THE 1910-1977 MEAN, FOR RAW DATA. INDIVIDUAL WET (BLACK) AND DRY (SHADED) YEARS ARE SHOWN TOO.

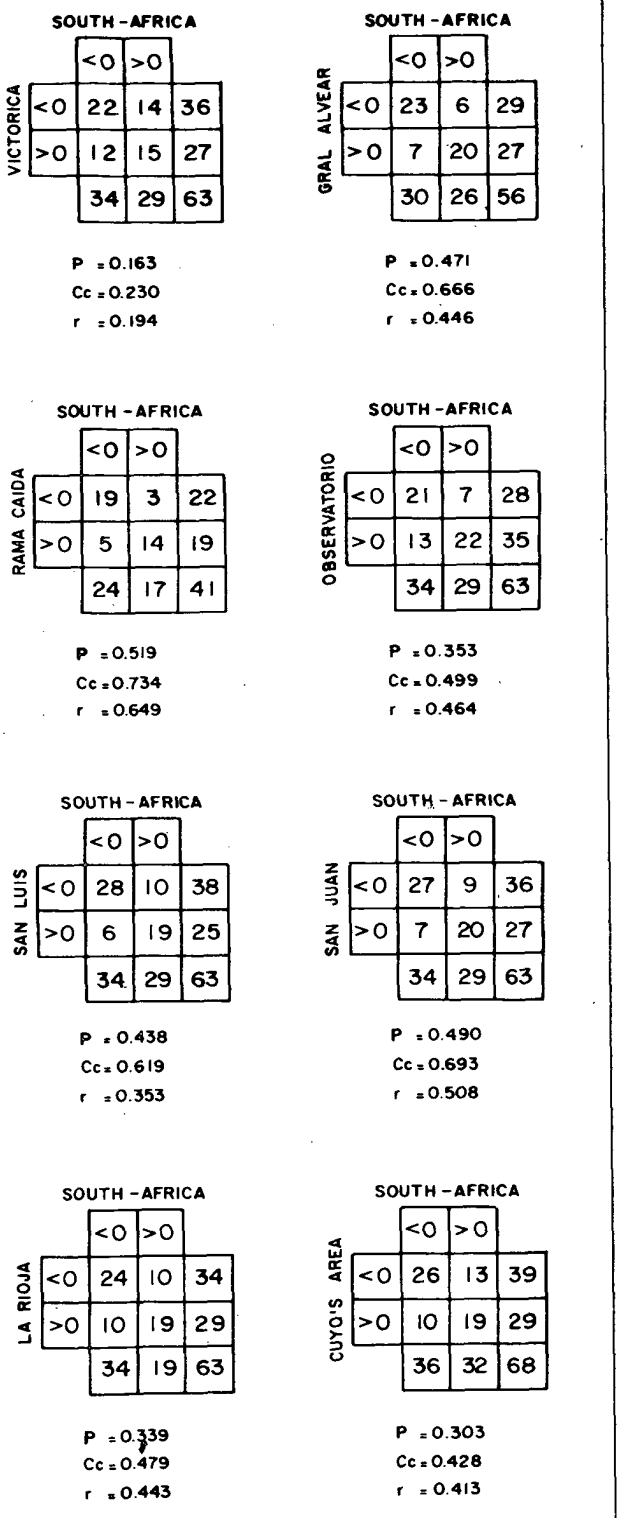


Fig. 12: CONTINGENCY TABLES, BETWEEN CUYO'S AREA AND SOUTH-AFRICA.
P : PEARSON'S COEFFICIENT
Cc: CORRECTED COEFFICIENT OF CONTINGENCY
r : CORRELATION COEFFICIENT

For the South African series the maximum of 18 years (approximately 90%) is more significant than for the Cuyo series (80% probability or less).

The differences are partly due to the fact that few series were available for this paper, while in calculating the spatial average series for South Africa, 62 series were used.

4.5 In Fig.11, the average spatial series, where the dry and damp periods for both areas, Cuyo and South Africa, are shown, it can be seen that the dry and damp periods are almost in phase.

Furthermore, the values of the deviation percentage with respect to the mean (abscissa) are greater in Cuyo than in South Africa. This would have the same explanation as for the small difference in both series' spectra. That is, the Cuyo spatial average is done with 8 stations while for South Africa it is done with 62 stations. In this way this last curve is smoother than the first.

The year 1944 achieves the maximum value for both areas.

The contingency tables (Fig.12) show the high correspondence between South Africa summer precipitation spatial average series with each one of the series and spatial average of Cuyo. It can be observed from the corrected coefficient results that Rama Caida (C_c = 0.734) and San Juan (C_c = 0.69) are the ones that show a mayor relation with South Africa.

These results agree with those obtained from the spectral analysis, where San Juan presented a marked maximum centered in 18 years in the same way as the South African one.

Brief, the results of the coefficient of Pearson, the corrected coefficient and the correlation coefficient show a highly significant relation between both regions.

5. CONCLUSIONS

The summer precipitation series in the Cuyo region present quasi-periodical characteristics.

It can be observed in the spectral analysis maximums centered in the 18 and 4 years and another slightly sketched centered in the 2 years.

These periods are practically in phase with the summer rain periods in South Africa, if the analysed series are considered representative of the respective region.

Some of the series, specially San Juan, present markovian characteristics with a strong persistence.

San Luis and San Juan are parctically at the same latitude with the studied area in South Africa, which is also placed leewards to the mountains.

The reason why these both series are (Cuyo and South Africa) in phase is probably due to the comparable orography of both countries in regard to the general circulation of the atmosphere.

This particular relation will be analysed in the future from the study of different general circulation index and their correspondence with dry and damp periods.

ACKNOWLEGMENTS

The authors wish to thank Dr.Dyer who kindly offered the program used in this work.

REFERENCES

Compagnucci y Boninsegna, 1979: Estudio de la Serie de Precipitación del Observatorio Meteorológico de la Ciudad de Mendoza. Informe de las Primeras Experiencias de Ajuste de Defensa Temporada Granicera 1978-1979, Comisión Nacional de Investigaciones Espaciales, FAA, Tomo 2, 19-36.

Compagnucci, Boninsegna, Heinrich, 1982: Series de Precipitación de la Región Centro-Oeste y sus Variaciones, Geoacta, 11, 2, 219-231.

Jenkins y Watts, 1968: Spectral Analysis and its Applications, San Francisco, Holden-Day, 552 p.

Johnston, J., 1977: Métodos de Econometría, Barcelona, España, Vincens Vives, 464 p.

Kendall y Stuart, 1977: The Advanced Theory of Statistics, London, Charles Griffin and Company Lim, Vol 2, 723 p.

Panofsky y Brier, 1968: Some Applications of Statistics to Meteorology, University Park Pennsylvania, 224 p.

Tyson y Dyer, 1978: The predicted above normal rainfall of the seventies and the likelihood of droughts in the eighties in South Africa. S.Afr.J.Sc. 74, 372-377.

Tyson, 1980: Temporal and Spatial variation of rainfall anomalies in Africa, South of Latitude 22°during the period of meteorological record. Climatic Change, 2, 363-371.

THE 16-DAY PEAK IN THE SECOND EOF OF THE 500 MB GEOPOTENTIAL HEIGHT

Juhani Rinne and Annakaisa Sarkanen

Department of Meteorology
University of Helsinki
Hallituskatu 11-13
SF-00100, Helsinki 10, Finland

1. INTRODUCTION

The existence of a 16-day wave has been verified both theoretically (eg. Kasahara,1980, his Fig.3) and observationally (Madden and references therein, 1978). In this work we present a 16-day wave, which was initially found in the spectrum of an empirical orthogonal function (EOF). The horizontal structure of the wave and the interference between the wave and the quasi-stationary time-mean wave are studied with the aid of composite maps. These are computed from data of seven years, extending from 1963 to 1970 in the Northern Hemisphere.

2. THE DATA AND THE ANALYSIS METHOD

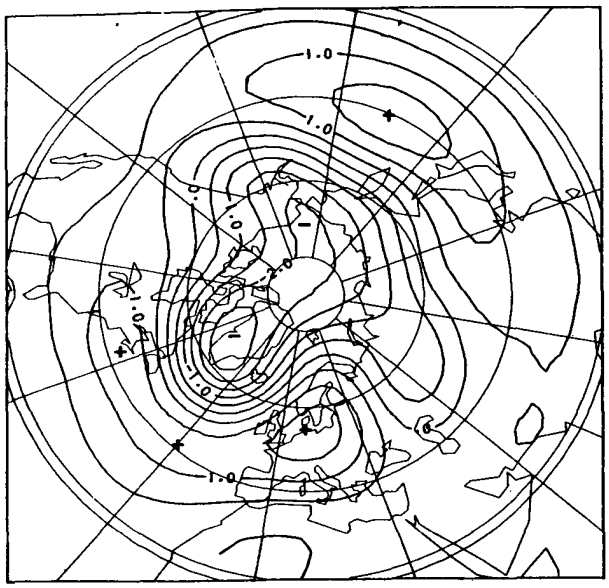
The basic data consists of 14672 FNWC (Fleet Numerical Weather Center) analyses of the 500 mb geopotential height for the 25-year period 1946-70. In addition to that, NMC (National Meteorological Center) data for the years 1966-79 has been applied. The analyses are given in the FNWC grid, a total of 1404 points north of 20°N. Data values are given twice a day, at 00 and 12 GMT, except in 1946-55 only once, at 12 GMT.

The empirical orthogonal functions (EOF) have been computed using 81% of the available FNWC analyses of the 500 mb height (Rinne et al.,1981). When the functions (g_ν) are known, the time-dependent coefficients can be determined from

$$C_\nu(t) = \sum_i g_\nu(i)\left\{z(i,t)-\bar{z}(i)\right\}dA(i), \quad (3)$$

where z is the height of the 500 mb surface, $\bar{z}$ is the long-term mean field and dA is the area element. Indices i and t refer to gridpoint and time, respectively.

The missing coefficients have been interpolated linearly (less than 2.6% of cases). The spectra of the time dependent coefficients C_ν have been computed applying the Maximum Entropy Method with the Final Prediction Error Criterion (Ulrych and Bishop,1975). The spectra show a number of relatively weak peaks superimposed on a red noise spectrum in the frequency interval of 5-200 cycles per year.


1.0
1.0
1.0

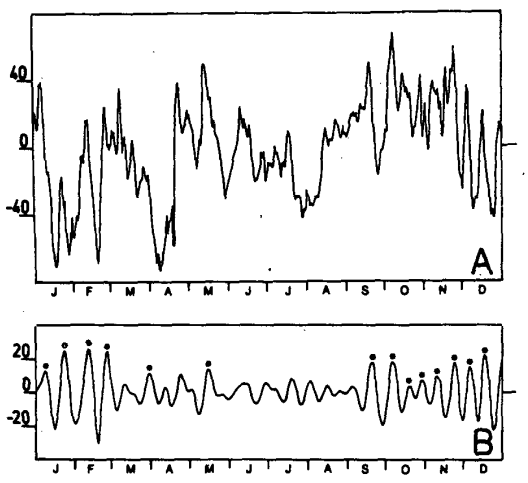


FIG.3: The original (a), and band-pass filtered (b) time series of the 2nd EOF during 1966. The selected maximum days (see text) are indicated. Unit: gpm.

zero points between them were picked out. In the following, the corresponding analyses will be symbolically referred to as 0, $\pi/2$ (maximum), π , and $3\pi/2$-analyses (minimum). Thus these terms correspond to a hypothetical sine-cycle with a period of 16 days.

4. RESULTS

The twice-daily time series of the original and filtered data during 1966 are given in Figs.3a and 3b, respectively. The year is quite representative. There are times with a more regular periodicity, for instance during period 3.12.1963-5.5.1964 11 regular cycles were repeated,similar in shape and size. During summer the amplitudes are always small, though the oscillation may be regular (e.g. 1968).

Composite maps were determined by computing the means of 0, $\pi/2$, π and $3\pi/2$-analyses. Only cycles with well defined maxima and minima were considered. No summer cases (June-Aug) were taken into account. The cycles accepted in 1966 are indicated by $\pi/2$-days in Fig.3b. The total numbers of cycles and analyses picked out during 1963-70 were 117 and 4*117, respectively. The mean of all these analyses was removed from the composite maps. The resulting anomaly maps are given in Figs. 4 and 5.

Typical maximum values of the 16-day period are 20m or less (Fig.3b). Multiplying the pattern in Fig.1 with +20m and -20m gives the approximations of Figs.4 and 5. However, these differ from Fig.1. In Figs. 4 and 5 there is practically no flow around the North Pole, or in other words, the zonal wave number 0 is absent. Instead there is a wave with zonal number 1. In the meridional direction a node is situated between 40 and 50°N. The main activity is in the north.These features of the zonal and meridional structure correspond to the theoretical 16-day wave (Madden, 1978). Because no other peaks are present in the spectrum of the 2nd EOF, the wave corresponds obviously only to one single mode.

We propose that the present wave is actually the 16-day wave as derived directly from the routine 500 mb analyses. The 2nd EOF has only been playing the rôle of an indicator, which has been used in tracing the key days.

The present wave does not exactly correspond to a theoretical mode with zonal wave number one. The Atlantic side cells are more intense than in the theoretical wave. It is interesting to note that those cells locate over a secondary ridge (Iceland) and trough (Scandinavia) of the mean winterly flow (Fig.8). Further, a cell is found over the weak ridge in the Bering Straits. Thus the cells interfere with the secondary ridges and troughs of the stationary field in its winter position. Correspondingly, the composite maps of 37 winter cases including their mean field

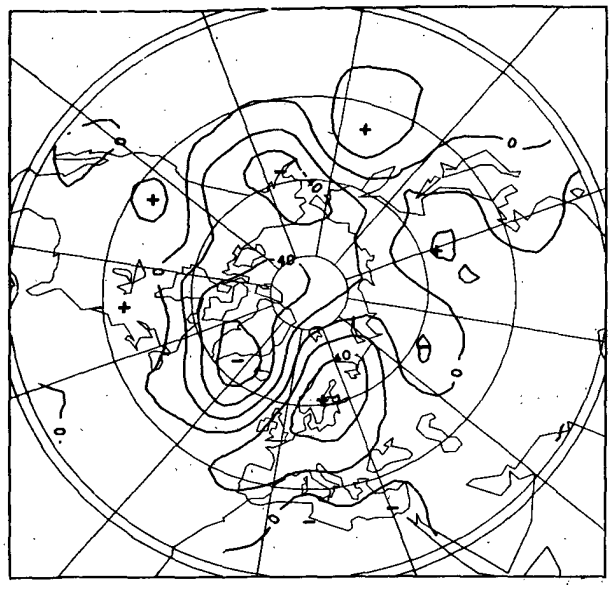


FIG.4: The composite map of 117 maximum days. The anomaly map is computed from the corresponding 500 mb analyses. Analyzed with 20 gpm intervals.

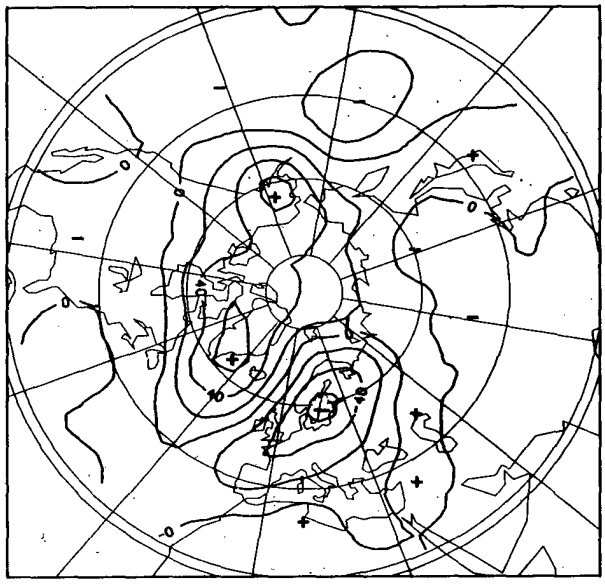


FIG.5: As in Fig.4 but for minimum days.

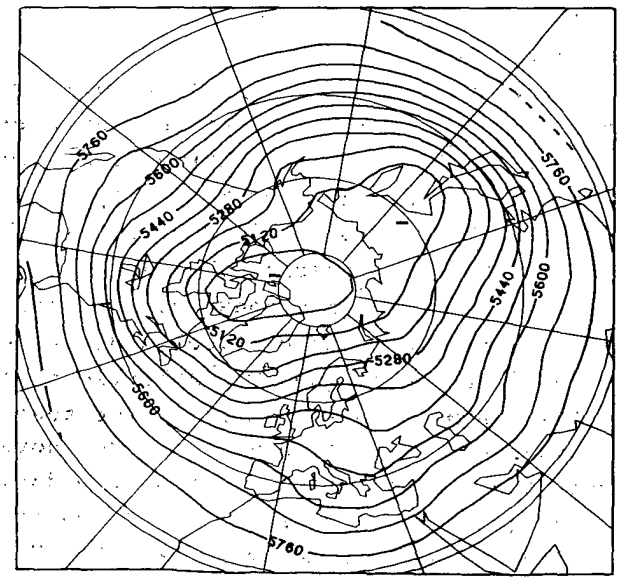


FIG.6: The composite 500 mb map of 37 maximum days during winter (Dec.-Feb.). Analyzed with 80 gpm intervals.

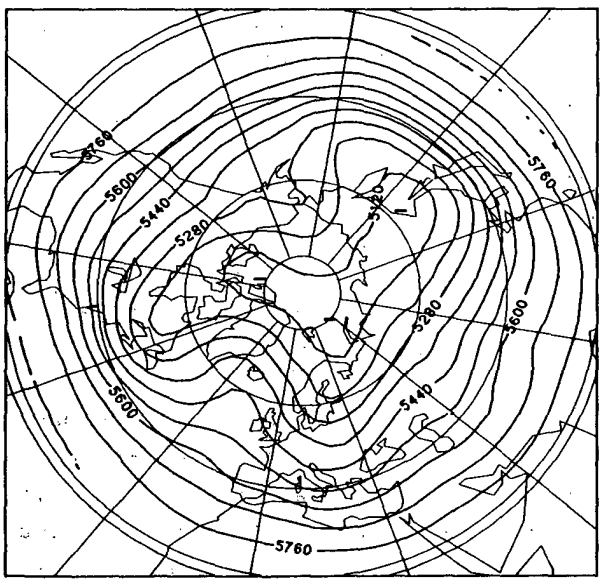


FIG.7: As in Fig.6 but for minimum days.

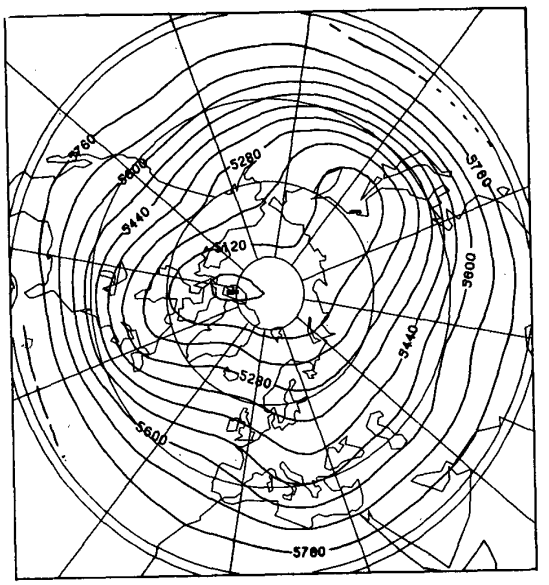


FIG.8: The winterly mean 500 mb field as computed from the 4*37 analyses. Analyzed with 80 gpm intervals.

show a zonal (Fig.6) and meridional (Fig.7) flow pattern. The interference makes the effect of the wave locally very pronounced. The result resembles the so-called index cycle with its zonal and meridional phases. Such changes in zonality are found in connection with the vacillation. Here the vacillation is produced by the interference of the standing waves and the travelling 16-day wave (cf. Lindzen et al.,1982). Similar arguments have been presented by Madden and Labitzke (1981).

The composite maps of 0- and π-phases indicate, that the wave moves westward (Figs.9 and 10). Note, that the identification of the key days is less unique than in the case of maximum and minimum days. The maps corresponding to Figs. 6 and 7 show intermediate stages of the development of the index cycle.

The significance of the result is not studied in this paper. The 16-day peak is found separately in four samples. The number of cases in the composite maps is large. The patterns in Figs.4 and 5 are constructed with the aid of an indicator, with which (Fig.1) they are only partly correlated. Nevertheless the two patterns are similar (excluding their sign). The Atlantic side activity can also be seen in the case study (Jan. 1979) presented by Madden and Labitzke (1981). All of these points hint that the results are significant.

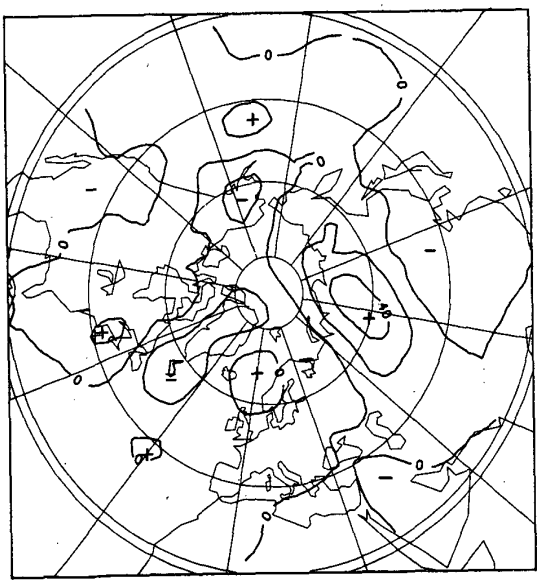


FIG.9: As in Fig.4 but for 0-days.

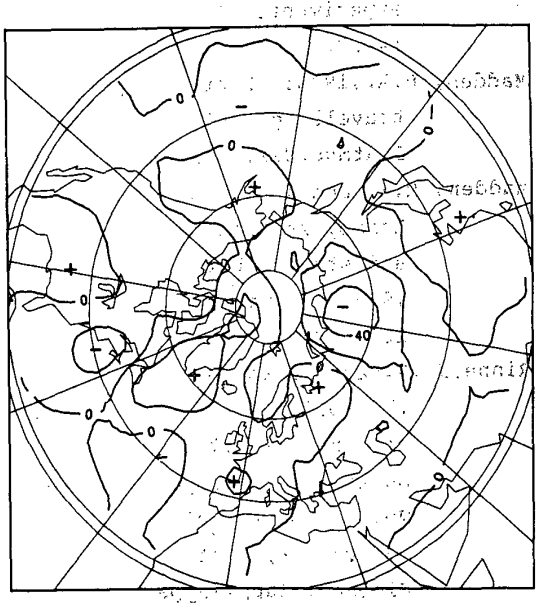


FIG.10: As in Fig.4 but for π-days.

5. SUMMARY

The spectrum of the 2nd EOF of the 500 mb geopotential height shows a permanent peak at 16 days. The composite maps show that the peak is due to a westward travelling wave with a zonal wave number one. The interference of the wave with secondary ridges and troughs of the winterly mean field results in a vacillation resembling the index cycle. During summer the oscillation is weak.

ACKNOWLEDGEMENTS: This work has been financially supported by the Academy of Finland.

6. REFERENCES

Kasahara, A.,1980: Effect of zonal flows on the free oscillations of a barotropic atmosphere. J.Atmos. Sci., 37, 917-929.

Lindzen, R.S., B.Farrell and D.Jacqmin, 1982: Vacillations due to wave interference: Applications to the atmosphere and to annulus experiments. J. Atmos.Sci., 39, 14-23.

Madden, R.A.,1978: Further evidence of travelling planetary waves. J.Atmos.Sci., 35,1605-1618..

Madden, R.A. and K.Labitzke,1981: A free Rossby wave in the troposphere and stratosphere during January 1979. J.Geophys.Res., 86, 1247-1254.

Rinne,J.,V.Karhila and S.Järvenoja,1981: The EOF's of the 500 mb height in the extratropics of the Northern Hemisphere. Report No. 17, Dept. of meteorology, University of Helsinki, Finland, 16pp.+Appendix.

Ulrych, T.J. and T.N. Bishop,1975: Maximum entropy spectral analysis and autoregressive decomposition. Rev. Geophysics and Space Physics, 13, 183-200.

ON THE PROSPECTS FOR CLIMATE PREDICTION

Roland A. Madden
National Center for Atmospheric Research*
Boulder, Colorado 80307 USA

1. INTRODUCTION

We must expect a certain amount of interannual variability in time averaged meteorological data due to statistical sampling variations. That is, one January mean temperature can differ from another because they each are finite time averages of daily values which result from a continuous process of passing weather systems. These sampling variations are unpredictable at long range (say a season or year in advance) since details of weather systems cannot be predicted at long range. Therefore, we will refer to them as "climate noise". Additional variance of the January means could reflect the effects of slowly changing external conditions such as sea surface temperature.† This additional variance may be unpredictable as well, but unlike climate noise it is, at least, potentially predictable. While it is not possible to isolate climate noise with certainty, it is important to have some idea of its magnitude in order to assess the prospects for climate prediction. To this end, we first make estimates of climate noise for temperatures over the Northern Hemisphere and then compare them with actual interannual variations. The more the actual variance exceeds the noise variance the greater is the potential for long-range predictability. Of course, in practice, the value of a long-range forecast is dependent on the size of the actual predictable variability or signal (not the potential predictability) relative to all of the unpredictable variability or effective noise (not the climate noise alone). In a closing section it is assumed that we have a viable signal in hand. A heuristic discussion is presented which compares the size of the signal to that of the effective noise in the manner described by Leith (1973) in order to assess the effect of a signal on predicted probability forecasts.

2. CLIMATE NOISE

2.1 Theory

Leith (1975, 1978) has applied a statistical definition of climate in terms of ensemble averages to aid in discussions of climate predictability. We can imagine an ensemble of Earths, each with the same distribution of land and sea and each under the control of identical external conditions. The daily weather varies on each Earth independent of that on all others, but we assume the physical laws governing its evolutions allow only a single, unique set of statistics in the presence of constant external conditions. The weather from each such imagined Earth provides a realization. Averages over the ensemble of realizations defines the climate for the given land-sea distribution and set of external conditions.

Finite time averages of individual realizations would fluctuate about the ensemble values and these fluctuations are what is called climate noise. They are statistical fluctuations of unpredictable (at long range) weather and therefore unpredictable themselves. We can further imagine that new external conditions produce a different set of statistics and the ensemble of resulting realizations defines a new climate. If only finite time averages of individual realizations are available, our ability to detect the change from the original to the new climate would depend on the ratio of signal (change from original to new climate)-to-climate noise. In the present context of the prospects of climate prediction we can think of the signal as a change whose value, even if correctly predicted, is limited by the attendant climate noise.

2.2 Practice

The above provides a definition of climate noise from which we can proceed to assess the prospects for long-range prediction. However, its application is not straightforward because in practice we cannot sample from an ensemble of Earths. With real data we cannot isolate climate noise with precision but we can get reasonable estimates of its size. In the idealized case all moments and joint moments of pressure, temperature, precipitation, etc. would define climate and estimates of each based on finite time averages of individual realizations would be degraded by climate noise. In what follows we estimate climate noise associated with only the first moment or mean of a single variable.

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

† Here we conceptually view all processes whose time scales are long relative to daily weather fluctuations as external conditions.

To begin let us assume that a large sample of daily observations taken during a given season of the year approximates the idealized ensemble. If for the time being we neglect changing external conditions then, for example, the distribution of daily winter temperatures might look as in Fig. 1-a. The mean value is determined by the seasonal cycle and the standard deviation by the day-to-day variability resulting from the passage and intensification of weather systems. Fig. 1-b shows schematically the distribution of 31-day averages of the daily temperatures in winter assuming about 5.2 days for an effective independence time (T_o). The dashed curve in 1-b is the same distribution if one day's temperature were independent of the next (i.e., $T_o = 1$ and $\sigma_N = \sigma/\sqrt{31}$).

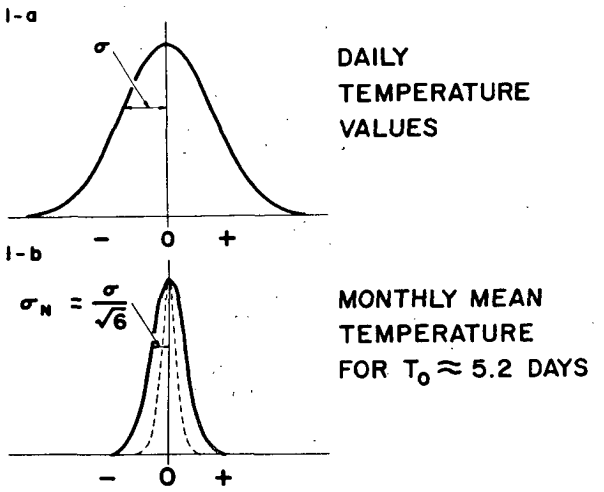


Fig. 1. Schematic of the distribution of daily temperatures with standard deviation σ (a) and of monthly mean temperature with standard deviation σ_N (b).

The standard deviation σ (Fig. 1-a) is a measure of daily temperature variability which is predictable up to deterministic predictability limits (one-two weeks). The standard deviation σ_N (Fig. 1-b) is a measure of climate noise since it represents the variability of 31-day averaged weather which would similarly be unpredictable at lead times longer than deterministic predictability limits. If we calculate σ and T_o from our ensemble of observed winter daily temperatures we can make estimates of what the variance of a time average of those temperatures would be. If the calculated σ and T_o closely approximate values associated with unpredictable weather alone then the estimated variance of time averages will closely approximate climate noise. Leith (1973, 1975) determines T_o by the integral of the autocorrelation coefficient (near-zero frequency estimate of the normalized spectrum) and σ can be estimated from the daily data. Alternatively, the variance of time-averaged data can be evaluated from the near-zero frequency estimate of the unnormalized spectrum itself (Jones, 1975). Here we use the latter approach. The data are daily, 00 and 12Z surface temperatures at grid points over the Northern Hemisphere during the period October 1965 through December 1977. These data originated at the National Meteorological Center, Washington, D.C. and are part of the NCAR data archives (Jenne, 1975). The spectra of the daily data were computed within winters (96 days beginning each Dec 1) at each grid point after the seasonal cycle in the mean was removed. Low-frequency variances (periods longer than 96 days) were modeled by a constant or low-frequency white noise extension of the computed spectra in order to estimate expected climate noise for monthly means. The procedure is discussed in detail in Madden and Shea (1978). The resulting climate noise for 31-day averages in winter is presented as standard deviations in Fig. 2. Fig. 2 represents a measure of the interannual variation we might expect as a result of sampling variations of daily weather.*.

Representativeness of Fig. 2 is dependent on how well the observed seasonal realizations approximate the idealized ensemble discussed in 2.1. We expect that there are changes in the mean of our schematic distribution of daily temperatures (Fig. 1-a) from year to year. If those changes were to occur only at very low frequencies, say interannually, they present no problems. In fact, they are the potentially predictable part discussed in the following section. It is likely that the variance and the autocorrelation of the spectrum change from year to year as well, but the analysis assumes they do not. Since weather variations are influenced by continually changing external conditions, time series of daily data even within a single season are composed, to some degree, of the equivalent of variations within our imagined single ensemble and variations between ensembles themselves. We do not know how serious these effects are on our estimates of climate noise. Numerical models offer an opportunity for controlled experiments which can provide an ensemble of realizations of model weather in the presence of constant external conditions. Studies with important implications for understanding climate noise have already been published (e.g., Chervin, 1980; Lau, 1981; Straus

* There are problems with the grid point temperature data used for Fig. 2 (personal communication, van Loon and Williams). Seasonal means have unrealistic, large variations in certain areas. These variations may not seriously degrade estimates of daily variance and autocorrelation if they occur between and not within seasons. Nevertheless, Fig. 2 is only a preliminary estimate of climate noise. Eventually we hope to make similar computations with station data. For the present we can only say that values over the U.S. from Fig. 2 are similar to ones that we have already determined from station data (Madden and Shea, 1978).

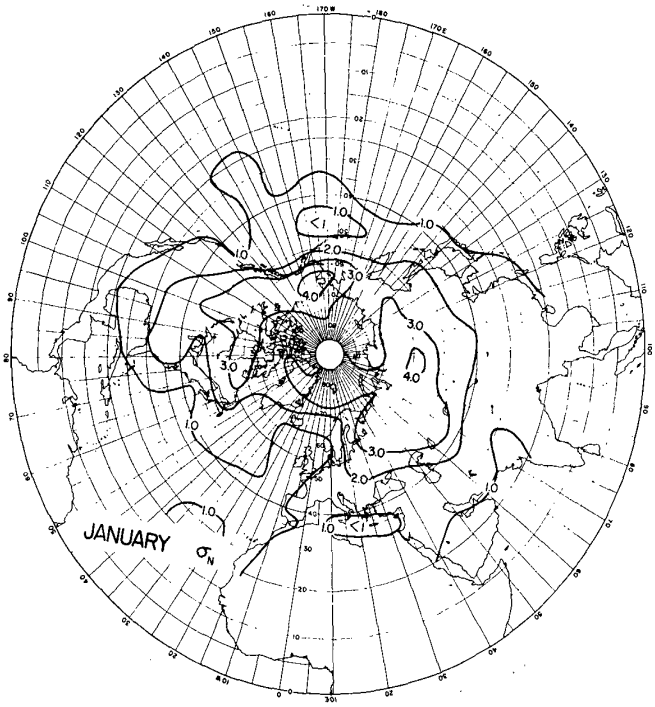


Fig. 2. Standard deviation of 31-day winter averages in °C associated with the estimated climate noise.

and Halem, 1981). Diagnoses of model data of this sort promise to provide further insight into the low frequency behavior of weather and into the quantitative reliability of estimated climate noise.

3. POTENTIAL LONG-RANGE PREDICTABILITY

Fig. 2 gives an estimate of the sort of variability of January mean temperatures that we should expect if all daily observations came from the same distribution. We might imagine that slowly changing sea temperature, for example, causes a shift in the distribution towards warmer (or colder) temperatures. If the variance and autocorrelation of the daily data remained constant, then 31-day averages would have the new, warmer ensemble mean but they would also vary with standard deviation σ_N. The shift to warmer temperatures is potentially predictable at long range since it is a variation over and above σ_N. We assume, then, that the actual variance of observed January means reflects climate noise and potentially predictable changes. The potentially predictable variations imply actual spectral values that would exceed the low-frequency white noise extension of the computed spectra of daily data that was used to determine values of σ_N in Fig. 2.

Because of the shortness of the record of grid point data, and more importantly because of inhomogeneities in it (see footnote) January means for stations published in the World Weather Records and available at NCAR (Jenne, 1975) were used to determine the actual variance (σ_A^2). Only stations with records longer than 40 years were included. A hand analysis of the σ_A^2 was made and values read at grid points matching those of Fig. 2. The resulting ratios σ_A^2/σ_N^2 are presented in Fig. 3. Areas where the ratio is greater than two, or where the potential predictability is greater than 50% of the total variance, are shaded. Ratios such as those in Fig. 3 determined with different data and/or slightly different methods vary somewhat in magnitude over the United States (Madden and Shea, 1978; Madden, 1981). All suggest that there is relatively less potential predictability (smaller ratios) over the center of North America and increasing potential nearer the coasts. From Fig. 3 we tentatively conclude that the same is true over Eurasia. Nicholls (1982) has shown that the potential predictability of temperature over Australia exceeds 50% in most areas. Highest values occur nearest the equator. Similarily, the potential predictability of pressure tends to increase from mid-northern latitudes to the subtopics (Madden, 1976). As a result, we further conclude (though not demonstrated here) that potential predictability is larger in tropical than in midlatitudes. Potential predictability varies from season to season as well (Madden and Shea, 1978; Nicholls, 1982).

We can only speculate why the pattern of potential predictability is as it is in Fig. 3. It is natural to point to the proximity of oceans for the relatively large ratios in coastal zones. The relative maxima that extend

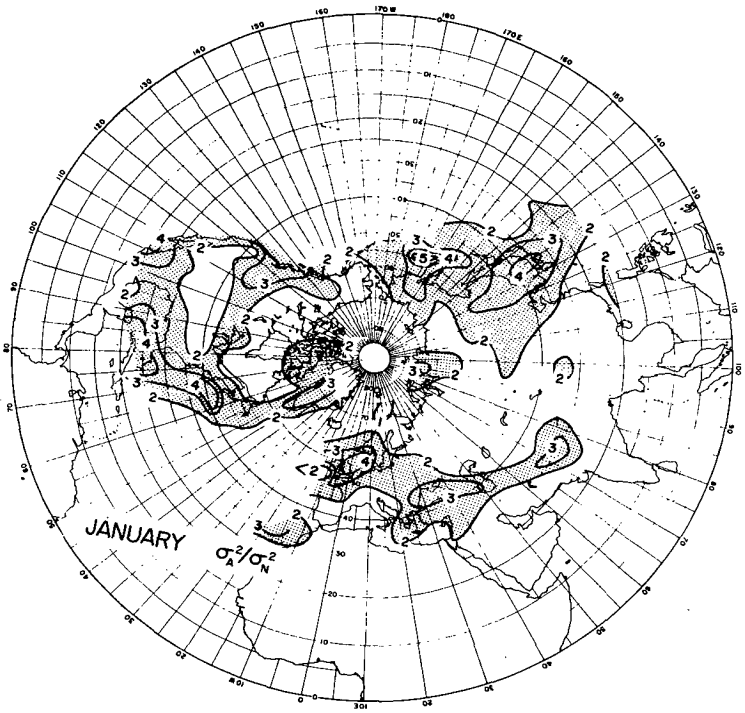


Fig. 3. Ratio of the actual variance σ_A^2 of January mean temperature to the expected climate noise σ_N^2. Areas where the ratio exceeds 2, or where potential predictability is greater than 50%, are shaded.

along the U.S.-Canada border and across eastern Europe and central Asia lie approximately along the mean, January snow line (Wiesnet and Matson, 1976). The oceans and snow provide local influences whose time scales are long relative to that of daily weather. In addition, slight changes in the direction of flow, from or toward the ocean or snow surface, that might be controlled from far away events could change the distribution of temperatures. This latter effect could occur with no change in the local ocean or snow surface.

Besides providing a measure of the potential predictability, the ratio σ_A^2/σ_N^2 is very effective in isolating inhomogeneities due to changes in instrumentation, exposure, station location, or reporting methods (see Fig. 5 of Madden, 1976). Unfortunately some of these effects that we are unaware of may be present in Fig. 3.

4. EFFECT OF A SIGNAL ON PROBABILITY FORECASTS

Our attempt to study the prospects for long range predictability considers that January means differ as a result of variability within a single imagined ensemble (climate noise) and due to shifts in ensemble means (potentially predictable signals). Leith (1973) considered the quantitative effect of signal-to-noise ratios on probability distributions. Fig. 4 is taken from this work. We can think of the distribution marked by the solid line as the distribution of monthly or seasonal mean temperatures under the influence of one set of external conditions and that marked by the dotted line under the influence of different external conditions. The shift in the mean is the signal and the standard deviation, assumed not to change, is the noise. Although the mean of the second distribution has shifted into the "above average" category of the first distribution, there is a finite probability that, due to noise, a single average from the second distribution will be in the "below average" of the first distribution. Fig. 5 from Leith (1973) shows the shifting probabilities for various signal-to-noise ratios.

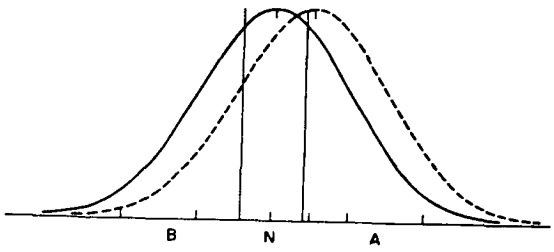


Fig. 4. Diagram showing the influence of a shift in the mean on the probability of outcomes in various categories: A, above average; N, near average; B, below average (from Leith, 1973).

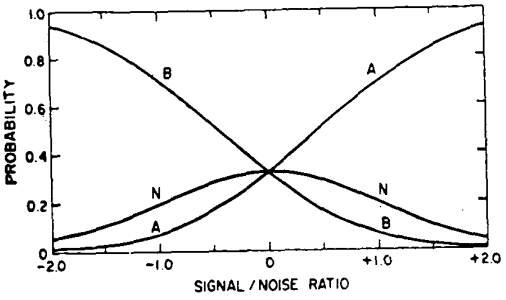


Fig. 5. The probability of outcomes in the three categories of Fig. 4 as a function of signal-to-noise ratio (from Leith, 1973).

To assess the effect of a plausible signal on a probability forecast we can measure the consequences of long-range predictions based on the Southern Oscillation. Fig. 6 shows the lag correlation between Darwin, Australia June-July-August (JJA) average pressure and following December-January-February (DJF) temperatures at North American cities. The most correlation is -.31 at New Orleans. Assuming this correlation is representative of future values, one could predict nearly 10% ($-.31^2 \times 100\%$) of the variance of DJF temperature at New Orleans from observed JJA pressures at Darwin.

We can relate the effect of correlations on shifting probability by considering a linear prediction equation of the form

$$y = a + bx \quad . \tag{1}$$

The predictand is y, the predictor x, a is a constant and the slope b is equal to $r_{y,x}\,\sigma_y/\sigma_x$. The correlation between y and x is $r_{y,x}$ and σ_y and σ_x are standard deviations of the predictand and predictor, respectively. The corresponding relation for time averaged values of y and x is

$$\overline{y} = a + b\overline{x} \tag{2}$$

From (1) and (2) we can relate signal-to-noise ratios in the predictand to observed shifts in the predictor:

$$\frac{y-\overline{y}}{\sigma_y(1-r_{y,x}^2)^{1/2}} = \frac{r_{y,x}}{(1-r_{y,x}^2)^{1/2}} \quad \frac{x-\overline{x}}{\sigma_x} \quad . \tag{3}$$

The quantity $(1-r_{y,x}^2)$ is the fractional part of σ_y^2 not explained by the linear relationship with x. The lefthand side of (3) is a signal-to-noise ratio of the predictand and the righthand side is a constant times a signal-to-noise ratio of the predictor. Fig. 7 is determined from (3) and gives predicted signal-to-noise ratios for various correlation coefficients or amounts of variance explained as a function of observed shifts in the predictor.

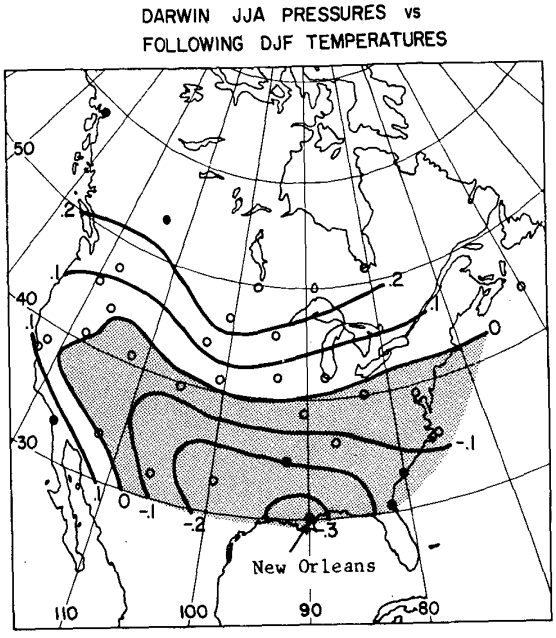


Fig. 6. Lag correlation between Darwin, Australia JJA pressure and the following DJF temperatures at stations indicated by the open and solid circles. Values based on 80 years of data from 1898-1977. Stations indicated by solid circles are locations where the sign of the lag correlation remained the same in each of four 20-year subperiods. Shaded region has negative correlations (from Madden, 1981).

The relatively large correlation at New Orleans would result in predictions with signal-to-noise ratios exceeding .5 about 10% of the time (Fig. 7). Corresponding probabilities for below, near-average, or above average values would shift from .33/.33/.33 to .53(.17)/.30(.30)/.17(.53).

The potential predictability at New Orleans is indicated to be about 50% of the total variance (Fig. 3). Assuming that, in fact, some measure of the Southern Oscillation can be used to predict 10% of the variance, it is possible that better knowledge of the quasi-biennial oscillation, persistence between season, etc. would enable us to account for additional variance. By analogy with these linear prediction arguments, it is conceivable that signal-to-noise ratios as big as .5 could be predicted more than 50% of the time in areas where 50% of the variance is potentially predictable; that is, a lag-correlation of slightly greater than .7 explains 49% of the variance.

5. SUMMARY

There is a certain percentage of variance of time-averaged meteorological variables (climate noise) that we will never be able to predict because it is simply sampling variability of unpredictable (at long range) daily

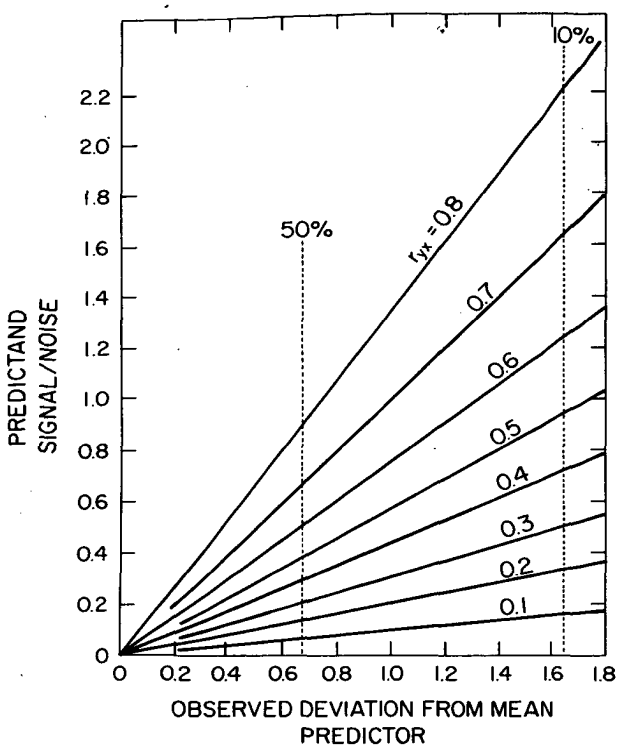


Fig. 7. Predicted signal-to-noise ratio or numbers of standard deviations that predictand is predicted to deviate from its mean (ordinate) for various correlations ($r_{y,x}$) as a function of the number of standard deviations predictor is observed to deviate from its mean (abscissa). Dashed vertical lines marked 50% and 10% are probabilities that predictor will be observed to differ by ±0.675 and ±1.65 standard deviations from its mean, respectively. These probabilities assume that the predictor is normally distributed (from Madden, 1981).

weather. Comparisons between estimates of climate noise and actual variability indicate that there is sizeable potential predictability, but that it varies from place to place and from season to season.

The value of even perfectly predicted signals is degraded by all the unpredictable variance or effective noise (climate noise plus potentially but not now actually predictable). At present, it is likely that accurate predictions of sizeable shifts in probabilities can only be made on relatively rare occasions. Forecasts of this sort can be made more frequent if we learn to predict more of the potential predictability.

Acknowledgments: Dennis Shea computed the climate noise shown in Figure 2.

6. REFERENCES

Chervin, R. M., 1980: Estimates of first- and second moment climate statistics in GCM simulated climate ensembles. J. Atmos. Sci., 37, 1889-1902.

Lau, N.-C., 1981: A diagnostic study of recurrent meteorological anomalies appearing in a 15-year simulation with a GFDL general circulation model. Mon. Wea. Rev., 10, 109, 2287-2311.

Leith, C. E., 1973: The standard error of time-averaged estimates of climatic means. J. Appl. Meteor., 12, 1066-1069.

_____, 1975: The design of a statistical-dynamical climate model and statistical constraints on the predictability of climate. Appendix 2.2, The Physical Basis of Climate and Climate Modelling. WMO No. 16, 265 pp.

_____, 1978: Predictability of climate. Nature, 276, 352-355.

Jenne, R. L., 1975: Data sets for meteorological research. NCAR Tech. Note TN/IA, 111, 172 pp.

Jones, R. H., 1975: Estimating the variance of time averages. J. Appl. Meteor., 14, 159-163.

Madden, R. A., 1976: Estimates of the natural variability of time averaged sea-level pressure. Mon. Wea. Rev., 104, 942-952.

_____, 1981: A quantitative approach to long-range prediction. J. Geophys. Res., 86, 9817-9825.

_____, and D. J. Shea, 1978: Estimates of the natural variability of time averaged temperatures over the United States. Mon. Wea. Rev., 106, 1695-1703.

Nicholls, N., 1982: Prospects for empirical long-range weather prediction. In the collection of position papers presented at WMO-CAS/JSC expert study meeting on long-range forecasting, Princeton, 1-4 December 1982.

Straus, D. M. and M. Halem, 1981: A stochastic-dynamic approach to the study of the natural variability of climate. Mon. Wea. Rev., 109, 407-421.

Wiesnet, D. R. and M. Matson, 1976: Monthly winter snow line variation in the northern hemisphere from a satellite records 1966-75. NOAA Tech. Memo Ness 74, 21 pp.

500 MBAR GEOPOTENTIAL VARIABILITY IN A 10 YEAR CLIMATE SIMULATION EXPERIMENT

Michel Déqué

Centre National de Recherches Météorologiques
Avenue Eisenhower, 31057 Toulouse Cedex, FRANCE

1. INTRODUCTION

A 10 year climate simulation has been performed at European Center for Medium range Weather Forecast (ECMWF) . The model used was a spectral T21 (i.e. truncated up to 21st zonal and meridional wavenumber) , 15 level (in σ=pressure/surface pressure coordinate) version of the ECMWF numerical prediction model . The climate generated by this model when computing monthly means of the output data has a good resemblance to the real global climate as shown by Cubasch (1981) for the general circulation and by Rilat and Thepaut (1982) for the hydrological cycle . In order to complete the analysis of these data , we investigated the variability of the simulated general circulation . In this study , we restricted ourselves to the daily values of the 500 mbar geopotential field which summarizes the barotropic aspect of the general circulation .

2. STATISTICAL TECHNICS USED

As we cannot consider 3660 charts of 32 latitudes by 64 longitudes , and as a survey on the map of variance is too restrictive , we performed a Principal Component Analysis (PCA) of the field.

Let $Z(x,t)$ be the geopotential field (x space coordinate, t time coordinate), $\bar{Z}(x)$ be the time average of $Z(x,t)$. We can expand $Z(x,t)$ in

$$Z(x,t)=\bar{Z}(x)+a_1(t)C_1(x)+a_2(t)C_2(x)+\ldots$$

$$+a_k(t)C_k(x)+R_k(x,t) \qquad (1)$$

in order to minimize the mean square of $R_k(x,t)$ for a given value of the integer k .

$a_1(t), a_2(t), \ldots, a_k(t)$ are called Principal Components (PC) and $C_1(x), C_2(x), \ldots, C_k(x)$ are called Empirical Orthogonal Functions (EOF) in the literature (Lorenz ,1956) . The EOFs can be represented by maps . The variance of $a_1(t)$, $\mathrm{Var}(a_1)$, is maximum and we have :

$$\mathrm{Var}(a_1) \geq \mathrm{Var}(a_2) \geq \ldots \geq \mathrm{Var}(a_n) \qquad (2)$$

where n is the number of degree of freedom of the field Z . We generally consider the rate $\mathrm{Var}(a_i) / \sum_{j=1}^{n} \mathrm{Var}(a_j)$ which represents the part of the ith component in the decomposition of the field . The solution of (1) needs the diagonalisation of the covariance matrix . In our study its size is 2048x2048 which cannot take place in the core storage of a Cyber 175 . The fields are in fact stored in spectral coefficients and if we calculate the covariance matrix of the spectral coefficients , if we diagonalize this matrix , and if we transform the eigenvectors from spectral coefficients to values on the Gaussian grid associated to the spectral truncation (here T21), it can be demonstrated that they are eigenvectors of the covariance matrix calculated with the fields on the Gaussian grid , with the same eigenvalues . So we have only 484 degrees of freedom and a 484x484 covariance matrix to diagonalize .In fact , as we want to focus on the large scale patterns we reduced the T21 to T13 truncation (up to 13th zonal and meridional wavenumber) . We lost 10% of the total variance in the small waves , and we get a 196x196 matrix to diagonalize .

The PCA on the data yields a first component which variance is 70% of the total variance and which time evolution is approximately a sine curve with a maximum in January and a minimum in July . The corresponding EOF shows negative values at North Pole and positive values at South Pole . This component represents clearly the seasonal cycle . The following PCs have small relative variances (2%) and present some annual cycle and some interannual variation . So we decided to deal separately with the annual cycle and with the interannual variability . We calculated thus a mean year from the 10 years of the sample (i.e. a mean 1 January , a mean 2 January , ...). We carried out a PCA on the mean year and we obtained a first PC which variance is 92% of the total. The annual cycle of the 500 mbar geopotential is essentially contained in one direction :a "seesaw" between Northern and Southern Hemisphere (the corresponding EOF is the same as that from the raw sample) .

Then we subtracted the mean year to each year of the data and we obtained a sample without annual cycle (in fact there remains a cycle in the

variances) . When we performed a PCA on the data after removing the annual cycle , we get two components of 4% of relative variance which present the same pattern in the Northern Hemisphere and opposite pattern in the Southern Hemisphere . Obviously , there is a degeneresccence of the first two eigenvectors (confirmed by Anderson's test). If we consider the correlation coefficient between $a_1(t)$ and $a_2(t+\tau)$ where $\tau=1,2,\ldots,10$ days ,we may see that it increases with τ so that $a_1(t)$ and $a_2(t)$ are uncorrelated but not independant .

Given an angle α ,we set

$$x_1(t)=a_1(t)\cos\alpha+a_2(t)\sin\alpha$$
$$x_2(t)=-a_1(t)\sin\alpha+a_2(t)\cos\alpha \qquad (3)$$

We look for α so that $x_1(t)$ and $x_2(t+\tau)$ have no correlation for $\tau=1,2,\ldots,10$ days . We find for each value of τ the value of $\alpha=\pi/4$. This technics called"Persistent Component Analysis"is detailed in Déqué (1982) . So we rotate the eigenvectors by $\pi/4$ and get a decoupling between Northern and Southern Hemisphere which seems reasonable for non-seasonal variations . The following components have a relative variance lower than 3% .

The persistency of the first PCs is very important . The autocorrelation coefficient is .9 at 2 days , .5 at 10 days , and .2 at 20 days . This effect is not a consequence of the statistical proprieties of the PCA ,but a propriety of the general circulation : the most important phenomena have a large spatial scale , and to this spatial scale corresponds a large time scale . Consequently , if we performe a PCA on 10 day averages , we get the same first EOFs with increased relative variances .

The variance of the first PC ,which is related to the Northern Hemisphere doubles from July to January , and conversely for the second PC , which shows that the annual cycle did not disappear . We should perform a PCA for a given season or a given month but if we want to compare the results of the simulation with real data we do not have sufficient global observed data samples to study the interannual variability .

3. COMPARISON WITH OBSERVED DATA

We are able to compare the simulated annual cycle with reality on the global scale . This has been done and confirms the predominance of the seasonal seesaw between Northern and Southern Hemisphere (60% for the first PC instead of 70% in the simulation) . In fact the variance of the annual cycle is good but the variance of extraseasonal events is underestimated in the simulation (about one half) because the boundary conditions such as the sea surface temperature or the sea ice extent are prescribed by the climatology and have no interannual variation .

In order to study the interannual variability we have to restrict ourselves to a domain where important data samples are available . On the Europe-Atlantic domain (see Fig. 1 to 3), we have a 20 year sample (from Météorologie Nationale , SMM/PREVI/MLE) which enables us to carry out comparisons .

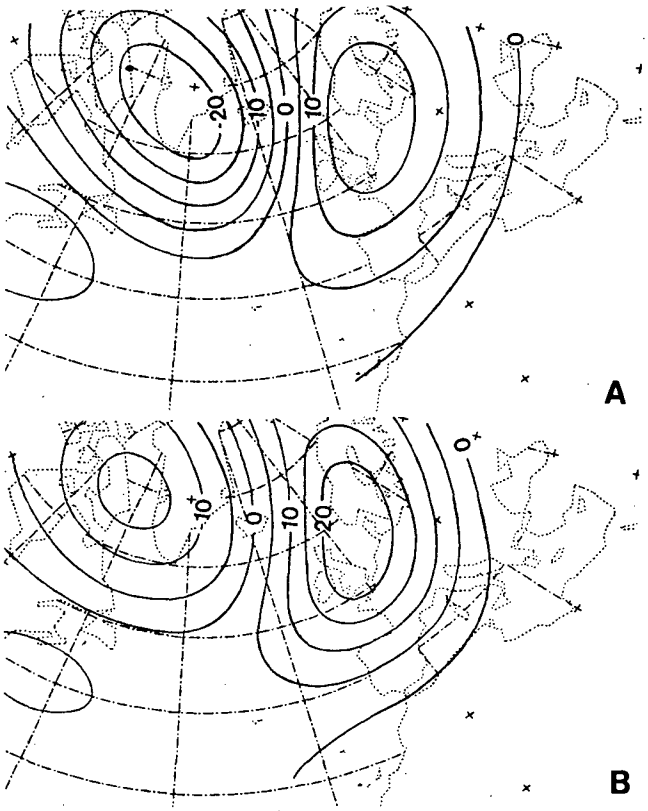


Fig. 1. Spatial distribution of the 1st EOF observed (A) and simulated (B) over the Europe-Atlantic domain . Arbitrary units .

Figure 1 shows that the first EOF (relative variance of the corresponding PC : 16%) is the same in the simulation and in the reality . The 2 day autocorrelation is about .7 . At this spatial scale , the time scale of the fluctuations is shorter than at the global scale , but we have still persistent phenomena . On the other hand , the first PCs have more important relative variance . This EOF corresponds to a phenomenon described by Van Loon and Rodgers (1978) as "Greenland seesaw" . When the PC is positive , there is a low over Greenland and a maximum over Northern Europe . The flow over Europe is SW-NE and we get warm air . When it is negative , the flow is NW-SE and we get cold air .

Figure 2 shows the 2nd EOF (relative variance of the corresponding PC 12% , 2 day autocorrelation .8) . The agreement between the simulation and the reality is not so good but the kind of phenomenon is the same . When the PC is positive , there is an intensification of the flow over the British Isles , when it is negative the flow is divided in two parts, the one at 60°N the other at 40°N , and we get a blocking situation over Western Europe (reality) or over near Atlantic (simulation) . This phenomenon is more persistant than the Greenland seesaw .

The 3rd (Fig. 3), and the 4th (not shown) simulated EOFs present a good resemblance with the reality . This is not due to a boundary effect nor any numerical effect because a control experiment which used data from a 6 year simulation with a

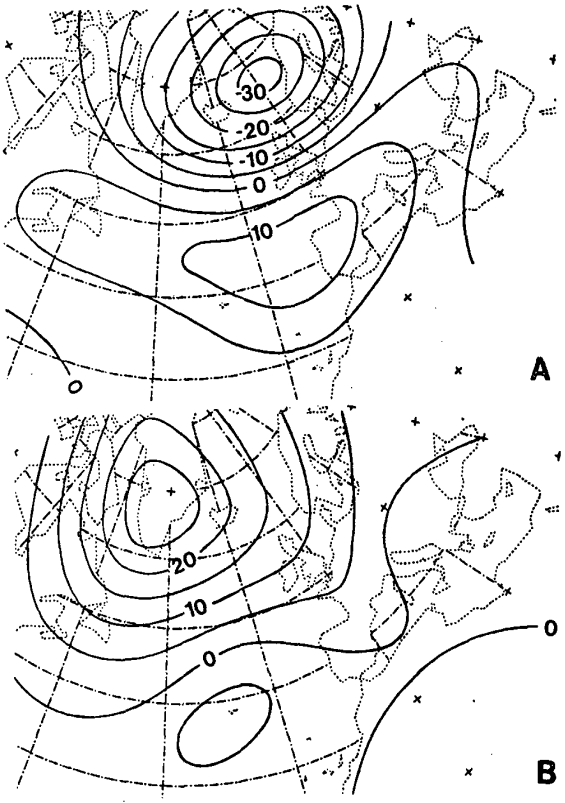


Fig. 2. As in Fig. 1 except for the 2nd EOF .

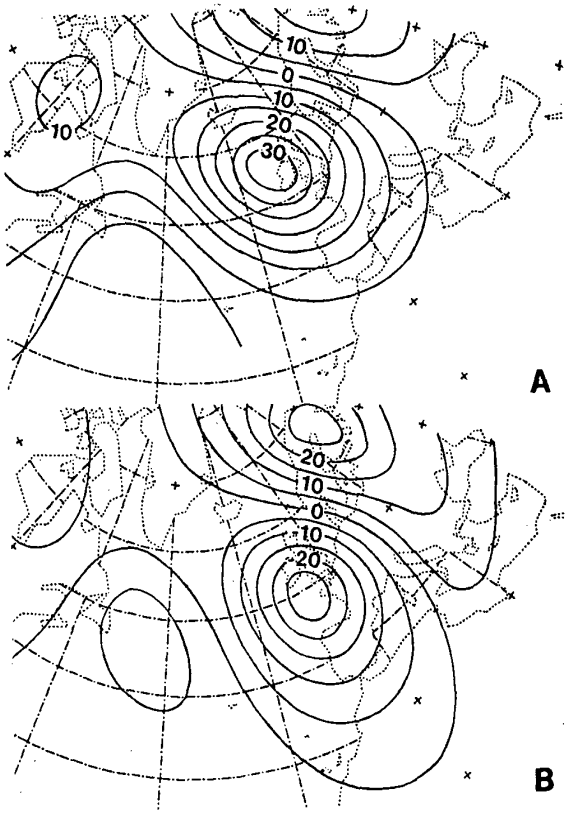


Fig. 3. As in Fig. 1 except for the 3rd EOF .

less performing model did not yield such patterns (only the 1st EOF was correctly simulated) .

We have also carried out a PCA on a 13 year sample of hemispherical data (from Météorologie Nationale , SMM/CLIM/DEV) . We can see on Fig. 4.A that the Greenland seesaw is included in a stationary wave of wavenumber 3 . The comparison with the simulated EOF (Fig. 4.B) is not so satisfactory as for the Europe-Atlantic domain . The reason is that in the model , as in most model , the troughs over the Aleutian Islands are too

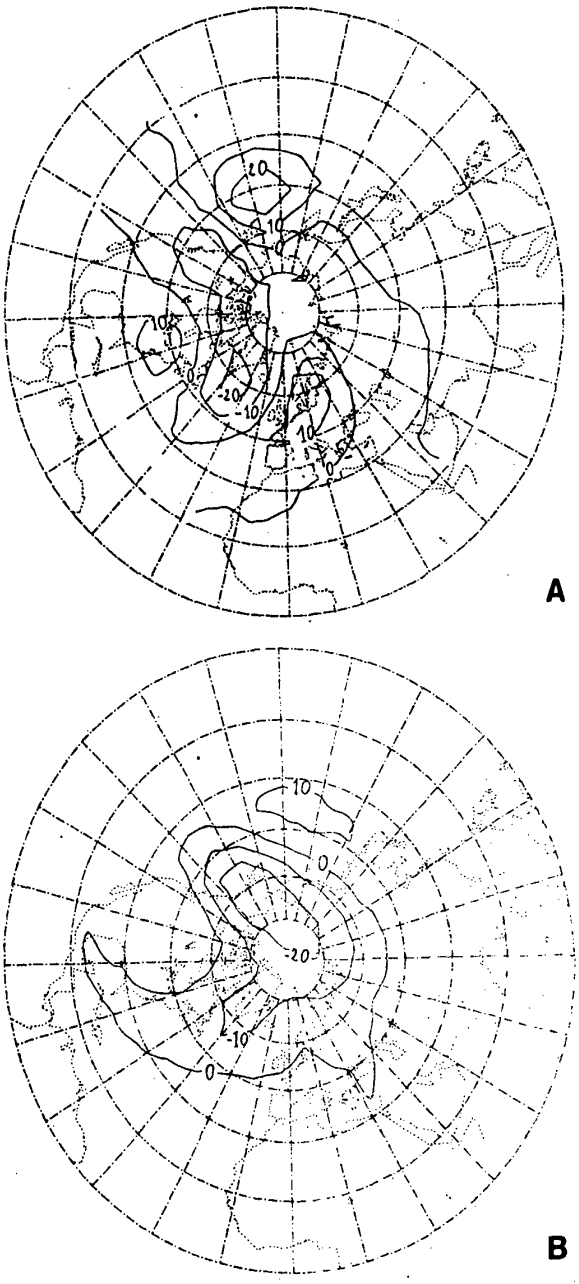


Fig. 4. Spatial distribution of the 1st EOF observed (A) and simulated (B) over the Northern Hemisphere . Arbitrary units .

developed , if compared with those over Iceland , and the variance over this region is overestimated . So the simulated EOFs are mainly located near this region . Only the first EOF (relative variance 7%) presents some resemblance with reality .

4. SUMMARY AND CONCLUDING REMARKS

Long integrations of General Circulation Models have shown that the mean state of the atmosphere can be approximately obtained by a numerical model and that its main features are induced by the boundary conditions of the model (orography , land-sea contrast , ...) .

As in the real atmosphere , the model oscillates on both sides of this mean state . The PCA is a good technics to investigate such phenomena. We have shown that these oscillations are following a certain number of modes and that the most important modes are quite similar to those of the real atmosphere .

This suggests that the modes of variation of the atmosphere are for an important part determined by the boundary conditions (orography , etc) rather than the cause of variation itself (like the tune of a string which depends only on the physical properties of the string , and not on the way you set it into vibration) . The response of the general circulation to a non-seasonal external forcing (e.g. a sea surface temperature anomaly) can be well represented with a small number of these modes .

6. REFERENCES

Cubasch , U. , 1981 ,Preliminary assessments of long range integrations performed with the ECMWF global model . ECMWF Technical Memorandum N°28 .(Available at ECMWF,Shinfield Park,Reading,Berks,RG29AX England) .

Déqué , M. ,1982 ,Analyse en Composantes Persistantes . Note de Travail de l'EERM N°38 . (Available at EERM/CNRM,Avenue Eisenhower, 31057 Toulouse Cedex,France) .

Lorenz , E. N. , 1956 , Empirical Orthogonal Functions and statistical weather prediction . Scientific Report N°1 , Dept. of Meteorology, Massachusetts Institute of Technology, Cambridge , Massachusetts , USA .

Rilat , M. F. , and Thepaut , M. A. , 1982 , Analyse du cycle hydrologique dans une experience de 10 ans de simulation du climat . Note de travail de l'ENM N°20 . (Available at EERM/CNRM,Avenue Eisenhower,31057 Toulouse Cedex , France) .

Van Loon , H. , and Rodgers , J. C. ,1978 , The seesaw in winter temperature between Greenland and Northern Europe . Mon.Wea.Rev.,106, 296-310 .

OBJECTIVE CLASSIFICATION OF SYNOPTIC PRESSURE PATTERNS OVER SOUTHEAST AUSTRALIA

J.D. Jasper and H. Stern

Bureau of Meteorology
Melbourne, Victoria, Australia.

1. INTRODUCTION

Richman (1981) describes the aim of map typing as 'to identify a subset of the data patterns which occur most frequently thereby summarizing the essential information of the larger set so that meaningful descriptions and conclusions can be drawn from this type of classification process'. There are three key methods of synoptic typing which predominate in the literature. They are:

(a) The semi-objective approach. This has been applied by many authors, an often cited application being that of Lamb (1972) who published a classification of British synoptic situations over the period 1861-1971;

(b) The linear correlation technique. This method was developed by Lund (1963) who used the simple linear correlation coefficient $r_{i,j}$ with corresponding mean sea level (MSL) pressures from maps i and j in a sequence in order to identify recurring situations;

(c) The use of principle component analysis and empirical orthogonal functions. This approach to synoptic classification was first used by Fukuoka (1951) and Lorenz (1956) and in application any individual synoptic chart may be reconstructed by linearly combining a set of spatial functions.

The purpose of the present paper is to provide an objective classification of synoptic types in southeastern Australia employing an extension of Lund's (1963) method, and to indicate that this classification may form the basis for subsequent work in the areas of climatology and weather forecasting.

2. THE DATA SET

MSL pressure data were extracted for eight stations in southeastern Australia (Fig. 1). The distribution of these data is considered, in the authors' view, adequate to describe the direction, curvature and strength of the surface flow over the region. The period 1 June 1957 to 31 May 1979 was selected. This gave 8035 separate synoptic situations, at 0900 hours local time, for the 22 year period.

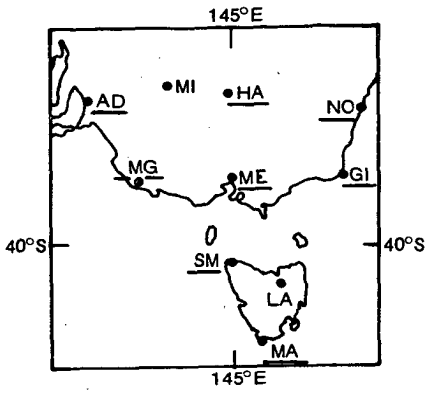
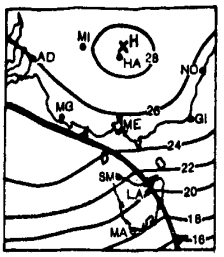


Fig. 1. Locations referred to in the text. Those underlined are the eight stations for which MSL pressure data are extracted (AD = Adelaide, GI = Gabo Island, HA = Hay, LA = Launceston, MA = Maatsuyker Island, ME = Melbourne, MG = Mount Gambier, MI = Mildura, NO = Nowra, SM = Smithton).

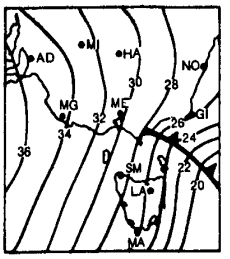
The authors were aware of errors, mainly arising from transcription, in the 'raw' data set. An editing procedure was devised which involved the development of regression equations relating the pressure at one location to that of the other seven - for each month of the year and using all the data. Large departures from values predicted by the regression equations were then investigated and many corrections were made to the original data. An illustration of the application of this technique is now presented.

An illustration of the data editing procedure.
Date - 7 September 1966.
MSL pressure station data (mb) - AD, 1016.1; GI, 1007.8; HA, 1013.8; MA, 1007.7; ME, 1011.2; MG, 1015.9; NO, 1009.5; SM, 1018.7 (see Fig. 1 for interpretation of station name abbreviations).

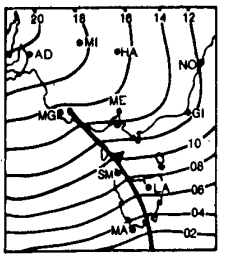
The September regression equation relating the MSL pressure at Smithton to the pressure at the other seven locations suggests that the value at SM is 1008.9 mb, not 1018.7 mb, a departure of 9.8 mb. The authors referred to the original archived analysis for that day and found that



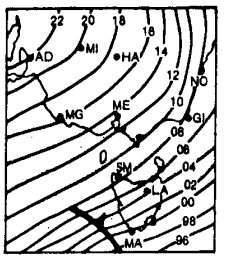
Type 1



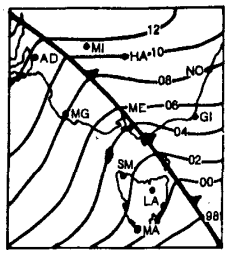
Type 2



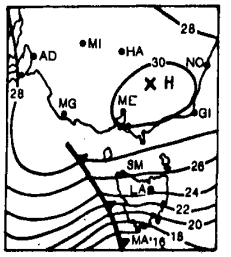
Type 1 example
Date: 29 Dec 1964
$r_{i,j}$: 0.9083



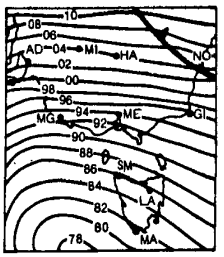
Type 2 example
Date: 29 Aug 1968
$r_{i,j}$: 0.8940



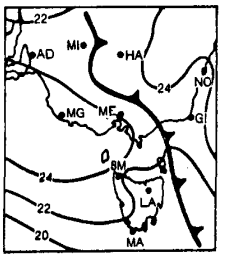
Type 3



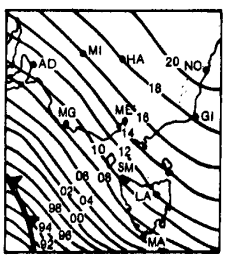
Type 4



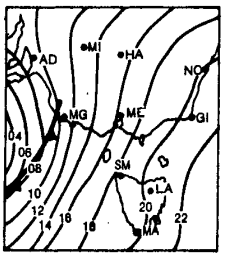
Type 3 example
Date: 18 Jul 1974
$r_{i,j}$: 0.9083



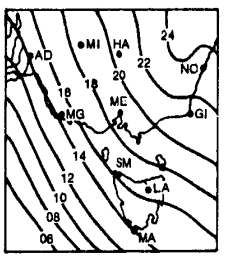
Type 4 example
Date: 12 Sep 1957
$r_{i,j}$: 0.9054



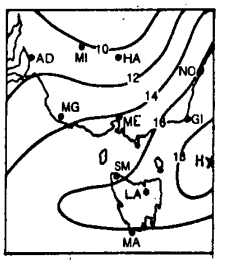
Type 5



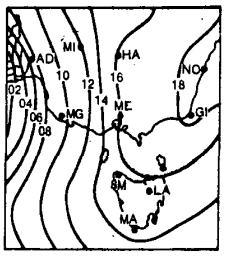
Type 6



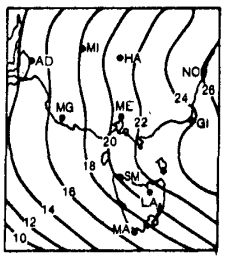
Type 5 example
Date: 8 May 1978
$r_{i,j}$: 0.9023



Type 6 example
Date: 10 Jan 1969
$r_{i,j}$: 0.8800



Type 7



Type 7 example
Date: 8 Jan 1961
$r_{i,j}$: 0.9127

Fig. 2. MSL pressure (mb) analyses of the seven most frequently occurring synoptic types. Isobars are labelled in whole mb (thousands and hundreds are not shown).

Fig. 3. MSL pressure (mb) analyses depicting randomly selected examples of the seven most frequently occurring synoptic types which have correlation coefficients ($r_{i,j}$) of about 0.9 to their respective archetypal patterns (shown in Fig. 2). Isobars are labelled in whole mb.

this pressure was, indeed, in error by about 10 mb. It was therefore concluded that in compiling the 'raw' data set a '1' had incorrectly been entered instead of a '0' in the third column and the value for Smithton was adjusted to 1008.7 mb.

In applying the data editing technique across the entire data set, the number of 'large' departures (≥5 mb) is greater for some locations (e.g. MA) than for others. This is mainly attributable to differences in MSL pressure variance between the locations. Nevertheless, the same criterion for a departure to be investigated (≥5 mb) is applied to the data set from each location.

3. THE METHOD OF CLASSIFICATION

The method of synoptic classification used is an extension of the linear correlation technique (Lund, 1963). Lund's technique may be summarised as follows:

Step 1. Correlate MSL pressures on each map with corresponding MSL pressures on each other map (using the simple linear correlation coefficient $r_{i,j}$ for each pair of situations i,j).

Step 2. The situation j associated with the most correlation coefficients $r_{i,j} \geq 0.7$ over i is designated as Type A.

Step 3. Remove all Type A maps, that is, maps whose correlation coefficients with Type A is ≥ 0.7, and select from the remainder the map associated with the most correlation coefficients ≥ 0.7 and designate it as Type B.

Step 4. Remove all Type A and Type B cases and select Type C as explained above.

Step 5. Repeat the process to derive Types D, E and so on.

The method of synoptic classification used in the present work is an extension of Lund's (1963) technique in that, following adjustments to correct apparently anomalous MSL pressure values, the data are partitioned into four seasonal blocks (the three-month periods commencing March, June, September and December) and synoptic types derived for each season. These seasonal types are then combined into an annual set of types. This is done in order that data from infrequently occurring types associated with, say, only one season, could be retrieved as a distinct pattern. Furthermore, the cutoff correlation coefficient decided upon is 0.9 in order that the synoptic character of the elements in each type would appear, on a subjective assessment, to be sufficiently similar.

In selecting the seasonal synoptic types (from some 2000 situations in each season), the data was further partitioned in order that computer time and data storage problems would be minimised. Using 'systematic sampling', with samples comprising every tenth daily situation, ten mutually exclusive samples (with starting points at the 1st, 2nd,...10th situation) were selected and Lund's (1963) technique applied. The classification technique was then applied to the ten sets of types for each season and sets of seasonal types were so derived. The technique was then applied to these four seasonal sets and an annual set of archetypal patterns was obtained. Finally, each of the 8035 situations was assigned to the archetypal pattern with which it was most highly correlated.

4. RESULTS

The application of the method as described above yielded thirty eight distinctly different patterns (note that the use of different criteria, e.g. cut off correlation coefficient = 0.95, may have given a different number of types). MSL pressure data associated with these patterns, and their frequency of occurrence, is presented in Table 1. It may be noted that the most frequently occurring seven types account for over half the sample and MSL pressure analyses for these situations are given in Fig. 2. A register of daily synoptic types over the period under consideration is presented by Jasper and Stern (1983).

An indication of the high degree of correspondence between the members of each type is given by Table 2, which lists mean correlation coefficients between the members and the archetypal situation. For most types, these are about 0.9. An illustration of the similarity of a randomly selected situation (with correlation coefficient of about 0.9) to its respective archetypal pattern, may be obtained by comparing corresponding maps of Fig. 3 with those of Fig. 2. Only the Type 3 example, which shows cyclonic westerly, rather than southwesterly, flow 'appears' to differ substantially from the archetypal pattern.

An examination of the frequency of occurrence of the patterns throughout the year indicates that some types are strongly associated with particular seasons. Table 3, which presents the seasonal distribution of the first seven types, underlines the tendency for Types 1 and 5 to occur mainly in the winter half of the year and for Type 6 to occur most frequently during summer. Spring and autumn are seen to be favoured seasons for Type 4.

A test of how well the classification technique discriminates between the types is to establish the differences between the 'average weather' associated with each circulation pattern. For the purposes of this paper, maximum temperature is examined. Maximum temperature is a consequence of a number of weather parameters such as cloud cover, precipitation, wind directions and speed, and mean temperature of the airmass to a considerable depth. Therefore, if real differences in maximum temperature are observed between the types it may be concluded that the technique has made adequate distinction between the types. This statement is made notwithstanding the realization that such a test may not necessarily show these differences at all locations or for other weather elements. Table 4 presents mean monthly maximum temperature data at Launceston, Melbourne and Mildura for each of the first seven types. Table 4 illustrates such real differences, particularly during the summer half of the year when maximum temperature variances are large, Types 2 and 3 being associated with

relatively cool temperatures and types 5, 6 and 7 being associated with relatively warm temperatures.

5. CONCLUDING REMARKS

A method of synoptic typing has been described and applied to the classification of 8035 southeast Australian MSL pressure patterns. Thirty eight distinctly different patterns were isolated, the differences being established through the method of averaging the data associated with those situations highly correlated with each other. Seasonal variations in the frequency of type occurrence and associated mean maximum temperature at several locations were also noted.

Finally, the work presented is seen as forming the basis for a great deal of subsequent research, namely:

a study of the occurrence of weather phenomena associated with each type for various localities, and, arising out of this:

(i) a study of fluctuations in the frequency of occurrence of each type as an indicator of short-term climatological fluctuations; and

(ii) the development of objective forecasting aids using the type as a discriminator.

An investigation into the relationship between forecasting errors and synoptic type.

Table 1. MSL pressure (mb) data and date (0900 local time) of the archetypal situation associated with the thirty eight synoptic types, and their frequency (%) of occurrence (AD = Adelaide, GI = Gabo Island, HA = Hay, MA = Maatsuyker Island, ME = Melbourne, MG = Mount Gambier, NO = Nowra, SM = Smithton).

Type	Frequency	AD	GI	HA	MA	ME	MG	NO	SM	DATE
1	10.4	1025.8	1025.4	1028.1	1018.7	1025.8	1024.4	1025.2	1022.0	16. 5.1969
2	9.0	1035.3	1026.9	1031.4	1026.7	1031.5	1034.6	1026.9	1028.3	26. 5.1972
3	7.6	1011.8	1005.6	1010.2	1000.7	1005.4	1008.3	1008.4	1001.9	7. 9.1969
4	7.5	1028.2	1028.9	1029.4	1016.7	1030.2	1029.1	1029.6	1025.9	19. 5.1961
5	7.3	1010.8	1018.3	1018.1	1006.1	1015.2	1009.6	1020.5	1010.7	6. 5.1973
6	6.4	1008.1	1020.8	1015.3	1019.9	1016.5	1010.6	1020.3	1019.7	20.11.1974
7	4.4	1005.6	1018.3	1016.8	1014.3	1015.1	1009.5	1018.3	1015.0	25.10.1971
8	4.0	1027.5	1026.5	1026.5	1024.2	1028.3	1028.5	1026.7	1026.6	5.10.1973
9	3.9	1020.3	1023.1	1020.2	1026.2	1022.6	1022.2	1022.6	1025.6	10. 5.1963
10	3.0	1029.0	1020.9	1023.6	1027.4	1026.5	1030.0	1019.3	1027.2	9. 4.1977
11	2.8	1018.2	1025.9	1022.5	1020.1	1021.2	1014.8	1028.8	1020.6	23. 5.1979
12	2.1	1019.4	1015.5	1014.0	1010.4	1017.1	1018.0	1012.4	1015.3	20.10.1970
13	2.1	1011.1	1011.7	1012.0	1004.7	1008.1	1006.8	1016.0	1005.7	11. 3.1977
14	2.0	1014.8	997.3	1007.2	1007.2	1007.6	1013.8	1001.7	1004.3	2. 3.1961
15	2.0	1018.1	1022.6	1020.7	1017.4	1021.5	1020.3	1022.8	1021.2	6.11.1977
16	1.8	1016.6	1020.7	1017.2	1019.8	1020.2	1019.2	1019.0	1020.4	17.10.1966
17	1.8	1014.9	1013.8	1013.2	1018.8	1016.4	1017.5	1010.3	1019.8	19.10.1976
18	1.8	1008.9	1016.4	1010.7	1018.6	1010.2	1010.1	1017.2	1015.4	20.12.1971
19	1.7	1019.3	1017.5	1016.5	1017.2	1019.3	1020.7	1015.8	1017.8	25.11.1969
20	1.7	1016.5	1013.4	1009.4	1026.7	1013.7	1015.7	1010.1	1022.0	15. 3.1979
21	1.7	1008.6	1008.2	1006.1	998.0	1004.5	1007.0	1011.0	999.8	4. 1.1974
22	1.3	1018.3	1020.5	1018.7	1016.0	1020.9	1020.7	1020.0	1019.1	21. 3.1971
23	1.3	1031.3	1026.9	1031.9	1030.8	1032.9	1033.5	1026.3	1031.9	14. 8.1976
24	1.2	1017.9	1021.1	1018.5	1023.0	1021.7	1021.0	1019.7	1020.7	7. 1.1963
25	1.1	1016.8	1016.3	1015.0	1019.9	1020.5	1020.4	1014.5	1019.3	10. 3.1962
26	1.1	1008.8	1015.1	1009.6	1012.9	1009.1	1004.7	1016.3	1011.0	9.10.1975
27	0.9	1017.4	1018.4	1014.1	1018.5	1018.9	1019.6	1017.9	1019.1	15. 2.1977
28	0.9	1019.7	1021.1	1025.0	1017.7	1021.2	1015.8	1023.1	1017.8	31. 7.1977
29	0.9	1031.6	1036.6	1033.2	1032.7	1036.1	1034.4	1035.5	1035.0	28. 6.1973
30	0.9	1018.7	1018.7	1016.2	1016.2	1018.8	1019.8	1017.5	1018.0	20. 3.1971
31	0.9	1012.5	1004.6	1007.7	1008.3	1006.1	1009.2	1005.0	1008.3	29.11.1973
32	0.8	1015.0	1012.1	1011.7	1009.8	1011.9	1014.2	1014.2	1009.8	15.12.1973
33	0.8	1015.9	1019.0	1020.1	1019.7	1020.3	1016.8	1019.0	1019.7	21. 8.1974
34	0.8	1030.4	1031.0	1027.6	1030.3	1032.4	1033.3	1029.8	1033.3	21. 3.1969
35	0.7	1020.1	1024.3	1020.7	1022.1	1024.2	1024.3	1022.7	1023.4	22.11.1965
36	0.6	1006.9	1011.5	1006.3	1008.0	1010.4	1008.9	1012.5	1011.0	8. 1.1979
37	0.5	1011.3	1015.5	1018.2	1014.2	1014.7	1010.9	1016.7	1015.3	9. 4.1969
38	0.5	1014.2	1015.3	1012.9	1014.1	1016.3	1016.2	1014.3	1014.2	4.12.1965

6. ACKNOWLEDGMENTS

The authors gratefully acknowledge the assistance and encouragement of their colleagues, in particular that of Miss A. Forfar and Messrs. P. Parker and R. Seecamp for their computer support. Helpful reviews by S. Barrell, R. del Beato and D.J. Linforth of the Bureau of Meteorology are also gratefully acknowledged. This paper is published with the permission of the Director of Meteorology, Australia.

7. REFERENCES

Fukuoka, A. 1951. 'A study of 10-day forecast (A synthetic report)', Geophys. Mag., 22, 177-208.

Jasper, J. D. and Stern, H. 1983. 'Objective classification of synoptic pressure patterns over southeast Australia and a register of daily types (1957-79)', to be submitted to Jnl. Climatology for publication.

Lamb, H. H. 1972. 'British Isles weather types and a register of the daily sequence of circulation patterns, 1861-1971', Geophys. Mem., London, 16, No.116. 85 pp.

Lorenz, E. N. 1956. 'Empirical orthogonal functions and statistical weather prediction', Sci. Rep. 1, Statistical Forecasting Project, Dept. of Meteorology, MIT, 48 pp.

Lund, I. A. 1963. 'Map-pattern classification by statistical methods', J. Appl. Meteor., 2, 56-65.

Table 2. Mean correlation coefficient between members of each synoptic type and the archetypal situation itself. This computation excludes the archetypal situation itself, for which the correlation coefficient is, of course, equal to 1.0000 .

Type	Number of members (excluding archetypal situation)	Mean correlation coefficient (excluding archetypal situation)
1	831	0.9390
2	719	0.9328
3	612	0.9188
4	601	0.9438
5	583	0.9444
6	515	0.9423
7	349	0.9418
8	322	0.9143
9	310	0.9245
10	244	0.9341
11	221	0.9332
12	169	0.9213
13	169	0.9191
14	159	0.9324
15	158	0.9153
16	145	0.9073
17	141	0.9036
18	140	0.8931
19	136	0.9118
20	134	0.8782
21	132	0.9128
22	103	0.9096
23	100	0.8406
24	92	0.8898
25	88	0.8839
26	86	0.8658
27	75	0.8576
28	73	0.8718
29	71	0.8893
30	70	0.9075
31	70	0.8243
32	67	0.8412
33	65	0.8238
34	60	0.8990
35	55	0.8564
36	50	0.8645
37	42	0.8596
38	40	0.8698

Richman, M. B. 1981. Obliquely rotated principle components: an improved meteorological map-typing technique?', J. Appl. Meteor., 20, 1145-59.

Table 3. Mean monthly number of occurrences for each of the first seven types.

Type	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	0.5	0.3	0.9	3.9	5.1	4.9	6.3	4.9	4.3	3.0	2.0	1.8
2	1.6	1.8	2.1	2.3	3.1	2.3	3.1	4.0	3.0	3.1	3.5	2.8
3	1.0	1.0	1.6	1.4	2.4	2.2	3.9	3.4	3.1	2.9	2.6	2.4
4	2.0	1.8	2.2	3.5	2.6	2.1	2.0	1.3	2.4	2.7	2.3	2.4
5	0.9	1.3	2.0	2.6	2.5	3.4	3.3	3.6	3.0	2.0	1.1	0.9
6	3.6	2.6	2.6	1.6	1.5	1.4	0.9	1.5	1.3	1.7	2.0	2.7
7	1.2	0.8	1.3	1.3	1.7	1.3	1.5	1.3	1.8	1.8	1.0	0.9

Table 4. Mean monthly maximum temperature (^{o}C) at Launceston (LA), Melbourne (ME) and Mildura (MI) for each of the first seven types.

Type	Location	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	LA	20.9	20.8	18.6	16.1	13.6	12.1	11.3	11.7	13.6	14.7	16.7	18.6
1	ME	24.8	24.4	24.0	20.3	16.9	15.2	14.7	15.2	17.6	18.7	23.2	24.0
1	MI	30.6	28.3	27.3	22.4	19.0	17.0	16.3	17.3	20.1	22.8	27.5	27.8
2	LA	20.2	20.5	18.8	15.9	12.7	10.4	10.2	11.6	13.1	15.7	17.2	18.2
2	ME	20.2	20.0	19.1	16.7	14.3	12.6	12.0	12.7	13.5	15.2	16.8	18.5
2	MI	26.1	25.7	23.9	19.7	15.9	14.0	13.7	14.9	16.5	20.2	23.0	24.7
3	LA	19.2	20.7	18.3	15.4	12.7	10.9	10.5	11.3	12.8	14.1	16.2	17.7
3	ME	22.5	23.9	20.6	17.8	15.2	14.1	13.1	13.5	15.0	17.3	19.2	21.0
3	MI	26.9	27.3	23.7	20.5	17.1	15.6	14.5	15.5	17.0	20.6	23.4	25.3
4	LA	21.6	22.1	20.2	17.3	14.8	12.4	12.0	12.5	14.4	16.7	18.0	20.0
4	ME	25.0	25.7	23.3	21.0	18.0	14.6	14.7	15.6	17.6	20.6	22.2	24.1
4	MI	31.3	30.7	27.8	24.4	20.3	16.5	16.0	18.4	21.1	25.7	28.0	29.7
5	LA	25.8	27.6	22.9	19.7	14.8	12.8	11.7	12.5	14.8	18.2	20.2	24.3
5	ME	35.7	33.8	30.7	24.9	18.6	15.7	14.9	16.7	20.5	25.1	29.0	34.6
5	MI	37.5	35.5	32.9	27.7	21.1	18.6	17.6	19.7	24.2	29.2	33.3	36.8
6	LA	24.9	25.7	22.7	19.1	14.0	11.1	10.7	12.7	14.5	17.8	19.9	23.4
6	ME	31.8	32.9	28.7	24.6	19.4	15.2	14.5	18.0	19.6	24.0	26.7	30.1
6	MI	34.0	34.4	30.8	26.4	21.5	17.1	15.9	19.2	21.2	27.2	31.3	34.2
7	LA	28.2	28.6	23.6	19.2	14.7	11.9	10.3	12.1	14.7	17.3	19.7	23.5
7	ME	34.2	34.4	30.2	24.7	19.4	15.8	14.5	16.5	20.4	23.4	27.7	33.2
7	MI	36.5	36.5	31.8	27.3	21.6	17.1	16.6	19.2	23.5	27.0	31.5	36.9

THE SYNOPTIC CLIMATOLOGY OF WINTER SURFACE TEMPERATURE IN THE UNITED STATES IN RELATION TO THE MONTHLY MEAN 700 mb CIRCULATION

William H. Klein
and
Joann M. Kline

Department of Meteorology
University of Maryland
College Park, Maryland 20742
USA

1. INTRODUCTION

The relationship between surface temperature and 700 mb heights has been the subject of considerable study since World War II. For example, Martin and Hawkins (1950) derived a graphical system of specifying 5-day mean temperature anomalies from 5-day mean 700 mb height anomalies at two points of maximum correlation, opposite in sign. This technique was extended by Stidd (1954), who showed that, if we disregard spatial variations in standard deviation of height, a correlation field is strictly analogous to an anomalous flow pattern. Correlation fields between 5-day mean anomalies of surface temperature and 700 mb height 2 days earlier during the winter season were constructed for 30 stations across the United States by Klein et al. (1959). Similar fields for simultaneous anomalies of 5-day mean temperature and height were published by Klein (1965) for a network of 15 cities. These studies revealed the existence of two widely separated centers of maximum correlation, one positive and within 1200 km of the reference station; the other negative and about 3000 km to the northwest.

The object of the present paper is to increase our knowledge of synoptic climatology by extending the work described above from 5-day mean to monthly mean data, and to do so for a longer time period, larger number of stations, finer grid of heights, and larger area. By analyzing and summarizing numerous correlation fields and height anomaly maps, regional differences in the relations between temperature and circulation will be highlighted, and schematic models of 700 mb height for opposite extremes of surface temperature will be constructed. However, this is primarily a fact finding study; a complete explanation of the results will not be attempted here.

2. CORRELATION FIELDS AND COMPOSITE MAPS

The basic data for this study consisted of monthly mean values of surface temperature at the 109 stations plotted in Fig. 1 and 700 mb height at 133 standard intersections of latitude and longitude shown in Fig. 2. The period of record extended from Jan. 1948 through Dec. 1981; this gave 102 months to represent the winter season, defined as Dec., Jan., and Feb.

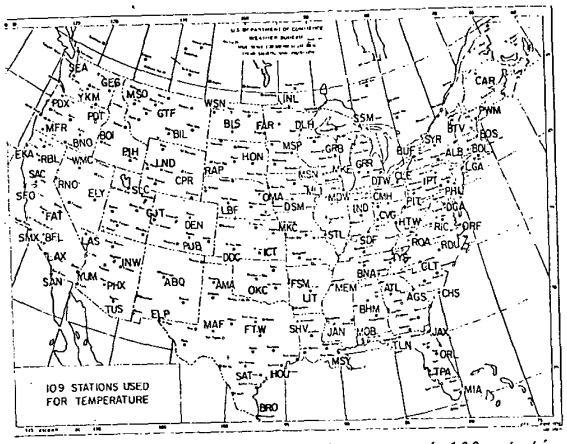


Fig. 1 Locations and call letters of 109 stations for which monthly mean surface temperatures were available.

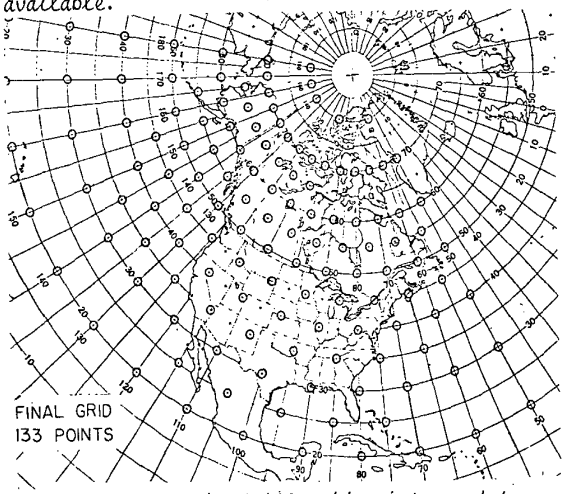


Fig. 2 Network of 133 grid points used to delineate the field of 700 mb height.

Anomalies or departures from local normal were computed for both temperature and height during each month of the sample by subtracting individual values from the 34-year monthly mean.

The first step was to compute simple linear correlation coefficients between the temperature anomaly at each city in Fig. 1 and the simultaneous monthly mean 700 mb height anomaly at each grid point of Fig. 2. Fig. 3 presents the monthly mean correlation field for Tampa, FL.

This chart was selected for illustration because its maximum positive correlation of .92 between the surface temperature anomaly at Tampa and the 700 mb height anomaly off the coast of Georgia is the highest single correlation yielded by any of the 109 cities studied. The sign and magnitude of this center, plus its proximity to Tampa, show that the lower troposphere in this area tends to be equivalent barotropic, with 700 mb ridges (or positive height anomalies) accompanied by warm 700 - 1000 mb thicknesses and surface temperatures, and troughs (or negative anomalies) by cold.

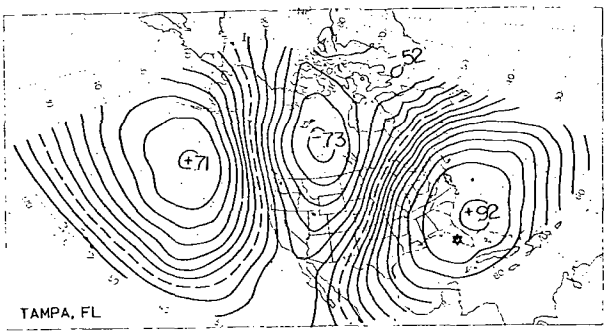


Fig. 3 Correlation field between concurrent winter anomalies of monthly mean surface temperature at Tampa (located by the star) and 700 mb height on grid of 133 points. The isopleths are drawn at intervals of .10, with zero dashed and central values labeled.

Fig. 3 also contains a pronounced center of negative correlation about 3800 km to the northwest of Tampa, with largest value (-.73) in western Canada. This center reflects a large scale advective effect, whereby the amplitude of the ridge upstream controls the deployment of cold air southeastward, as explained by Martin and Hawkins (1950) and Namias (1953). As a result, above normal heights in western Canada usually go with cold temperatures at Tampa, while below normal heights in this key region tend to be accompanied by warm weather at Tampa. A third correlation center (+.71) is found in Fig. 3 near 45°N/153°W, about a full wave length upstream from Tampa. This center resembles the well known teleconnection between the eastern Pacific and eastern United States (Namias, 1981; Wallace and Gutzler, 1981). The weakest center of correlation in Fig. 3 (-.52) is located at the eastern tip of Labrador.

Previous interpretations of correlation fields in terms of geostrophic anomalous flow (Stidd, 1954; Klein, 1965) were based on the assumption that the variability of 700 mb height does not change with location. As shown in Fig. 4, however, this variability actually exhibits well defined gradients in both zonal and meridional directions. The assumption of negligible spatial gradient can be removed by multiplying the correlation coefficient at each point by the standard deviation of monthly mean 700 mb height anomaly at the same point. This is equivalent to computing a map of the weighted correlation coefficients between anomalies of height and temperature. This map gives the optimum distribution of height anomalies for abnormally warm or cold weather at the reference station. It is therefore a sort of composite or height anomaly map but it is based on all available data, instead of only a few extreme cases as customarily used for composite maps. It can therefore be interpreted more easily and directly than can a correlation field.

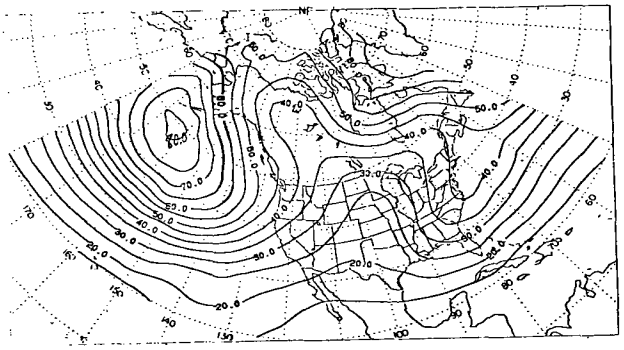


Fig. 4 Standard deviation of monthly mean 700 mb height anomalies (in meters) for 102 winter months.

An example of the new composite map is given in Fig. 5 for Tampa, FL for the same data used to construct Fig. 3. It shows that above normal temperatures at Tampa are favored by high 700 mb heights at the station and by strong southeasterly anomalous flow around a center of positive height anomaly near Norfolk, VA. Additional features contributing to Tampa's warmth are centers of negative height anomaly in central Canada and Davis Strait and a strong positive anomaly center in the eastern Pacific. Since the correlation coefficients are linear, the optimum height pattern for cold weather is the same but with signs reversed. Thus, below normal temperatures at Tampa go with below normal heights locally, northwesterly anomalous flow around a negative height anomaly center near Norfolk, a classical blocking pattern over North America, and a strong Pacific-North American (PNA) teleconnection.

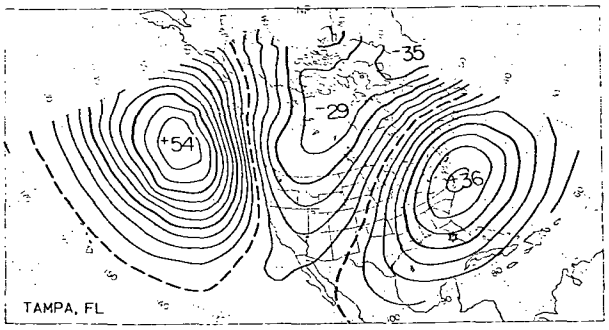


Fig. 5 Composite map showing the pattern of monthly mean winter 700 mb height anomalies conducive to warmer than normal months in winter at Tampa. The isopleths are drawn every 5 meters, with zero line dashed and central values labeled. They were obtained by multiplying the correlation coefficients of Fig. 3 by the standard deviations of Fig. 4.

Although Figs. 3 and 5 are similar in broad outline, they exhibit several differences. The correlation centers of Fig. 3 are displaced from

300 to 1000 km in Fig. 5 towards regions of greater height variance (Fig. 4). As a result, the anomalous flow at Tampa is more striking in the composite map than it is in the correlation field, and the direction for warm weather is southeasterly rather than southerly. The two centers of negative correlation in Canada in Fig. 3 are reversed in relative magnitude on the composite map (Fig. 5), where the western center becomes quite weak and the eastern center stronger. This reflects the classic North Atlantic Oscillation (Walker and Bliss 1932) and emphasizes the importance of blocking on temperatures in Tampa. Finally, the eastern Pacific center on the composite map is the most intense of any on the chart, even though it was next to the weakest on the correlation field. This may reflect the large variability of 700 mb heights (Fig. 4) in this area rather than any direct effect on Tampa temperatures.

3. LOCAL POSITIVE CORRELATION

In an effort to summarize the climatological features of the correlation fields and composite maps for all 109 cities, numerous quantities were tabulated, plotted, and analyzed. For example, Fig. 6 gives the magnitude of the maximum positive correlation associated with each station. It ranges from a high of .92 at Tampa to a low of .43 at Kansas City. Values are considerably higher over the Southeast, Northern Rockies, and Far Southwest than over the central United States, southeastern Plateau, West Coast States, and Northeast.

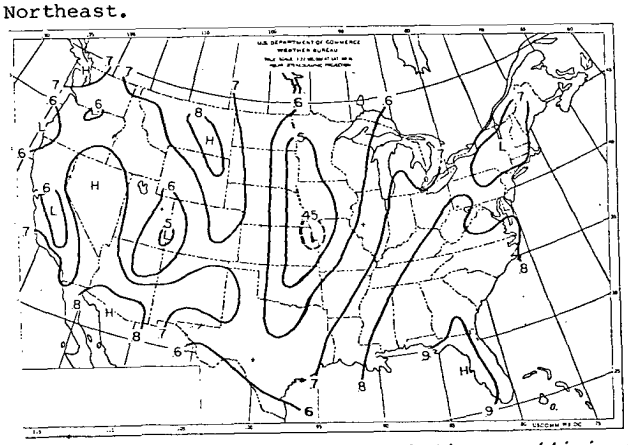


Fig. 6 Maximum positive correlation coefficient between concurrent anomalies of monthly mean surface temperature and 700 mb height at center of positive correlation, as analyzed on each of 109 correlation fields.

Fig. 7 shows the correlation between the surface temperature anomaly and the 700 mb height anomaly at the same point, rather than at the center of positive correlation as in Fig. 6. The general patterns of the two maps are quite similar as expected, but with magnitudes smaller in Fig. 7. Note the very low values on the mid-Pacific Coast. The amount by which the maximum positive correlation (Fig. 6) exceeds the strictly local correlation (Fig. 7) is given in Fig. 8. This difference is generally less than .15 in the eastern two-thirds of the country, but greater than .2 in much of the western third, with a maximum of .6 at Medford, OR.

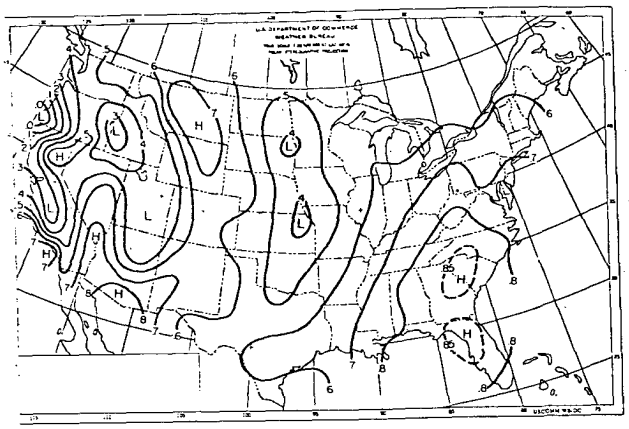


Fig. 7 Simple linear correlation coefficient between concurrent anomalies of winter monthly mean surface temperature and 700 mb height at the same point, as interpolated from isopleths on each of the 109 correlation fields.

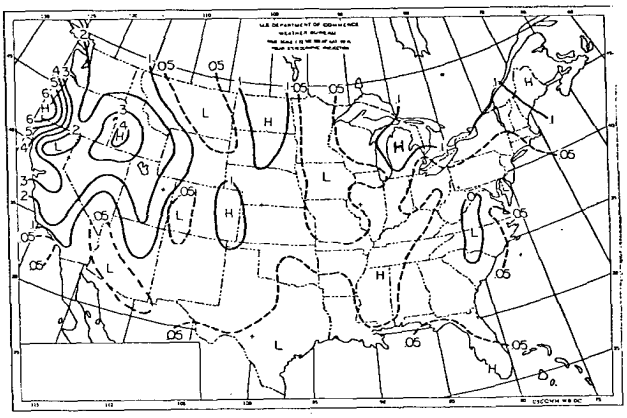


Fig. 8 The maximum positive correlation of Fig. 6 minus the strictly local correlation of Fig. 7.

If we plot the curvature of the anomaly isopleths on the composite maps at each reference station (not shown), we find that above normal temperatures are accompanied by anticylonic flow (as expected) in nearly all parts of the United States except the Pacific Northwest. The opposite curvature prevails for below normal temperatures. The direction of local anomalous flow can be inferred in a similar manner. The composite maps then indicate that warm weather is associated with southerly wind components in all portions of the nation except the Great Plains and Rocky Mountains, where northerly or westerly anomalous flow tends to bring above normal temperatures.

A similar feature is indicated in Fig. 9 which shows that the ridge of local positive anomaly usually is located east of the reference station, as expected. However, in the Northern Plains and Rocky Mountain region, the shaded negative values indicate that the ridge is located as much as 900 km west of the city. The distance from each station to its local positive anomaly center, regardless of direction, is

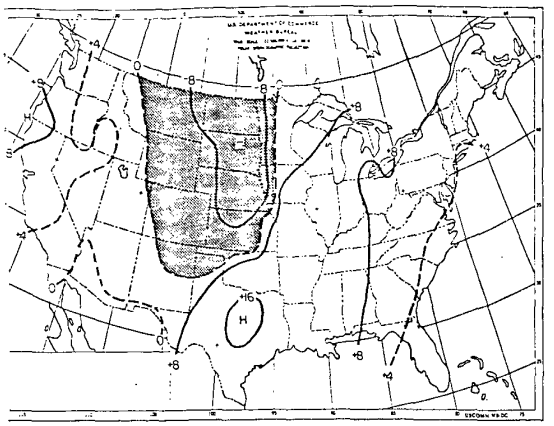


Fig. 9 Zonal distance in hundreds of km from station to ridge of local positive winter 700 mb height anomaly. All data are based on 109 composite maps. Values are shaded (negative) in areas where ridge is west of station.

depicted in Fig. 10. This distance is less than 1000 km in most of the country with minima below 400 km in parts of the Plateau and Middle Atlantic States. However, it exceeds 1200 km in parts of southern, central, and Pacific Coast regions, with a maximum of 3200 km in Texas.

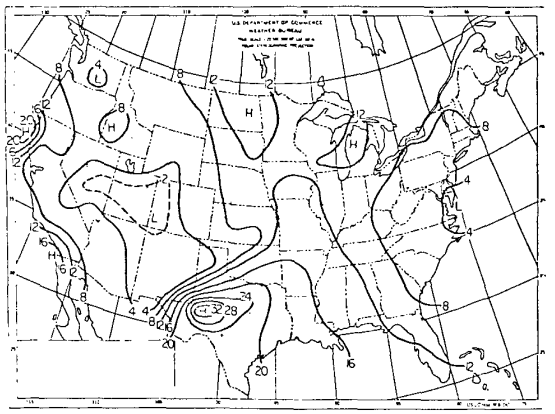


Fig. 10 Distance in hundreds of km from station to center of local positive winter 700 mb height anomaly. All data are based on 109 composite maps.

When the location of each local positive anomaly center is plotted on a map, three distinct areas of clustering emerge, as depicted in Fig. 11 for the reference stations grouped by region in the inset. Most pronounced is a small area of large height variability (Fig. 4) near the Middle Atlantic States that contains a positive anomaly center on the composite map for warm weather at every station in the southeast quadrant of the United States. Another area of clustering is located over the Rocky Mountain region, where positive anomaly centers are associated with warm weather throughout the western half of the country. A third grouping of anomaly centers is found for stations in the northeast quadrant of the nation, but here the anomaly centers are more widely dispersed.

When the centers of the local positive anomaly are plotted on a polar coordinate diagram with the reference city at the origin,

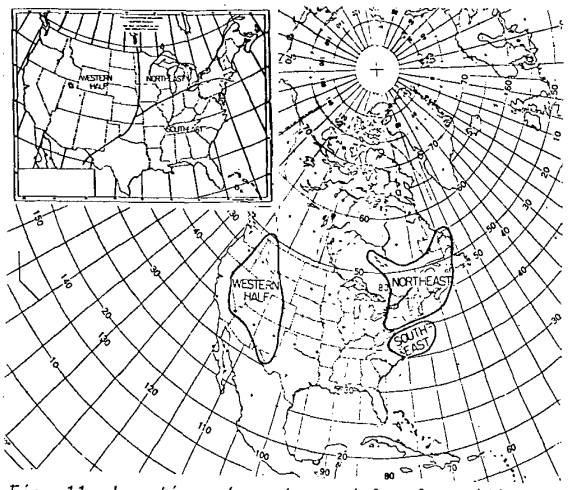


Fig. 11 Location of centers of local positive winter 700 mb height anomaly. All data are based on 109 composite maps, with separate presentation for cities in the Southeast, Northeast, and Western half of the United States.

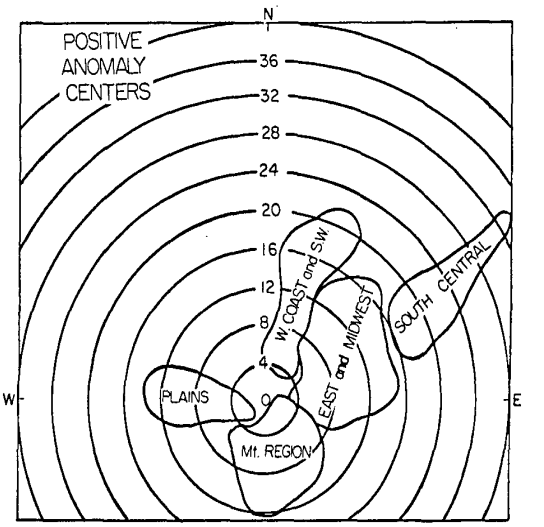


Fig. 12 Distance (km x 10^2) and direction (north, east, south, west) from reference station (at origin) to center of local positive winter 700 mb height anomaly grouped by five areas. All data are based on 109 composite maps.

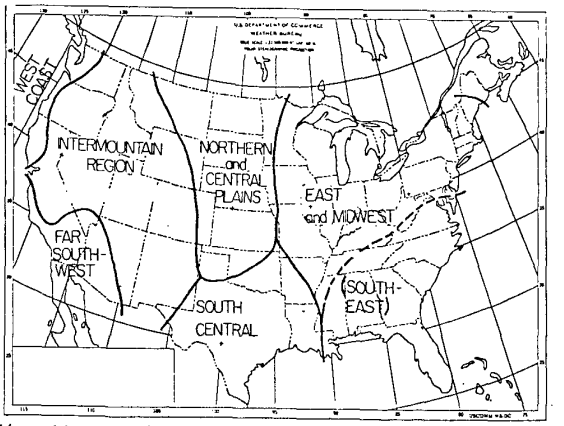


Fig. 13 Boundaries of homogeneous zones in the United States relative to location in Fig. 12 of local positive 700 mb height anomaly during winter.

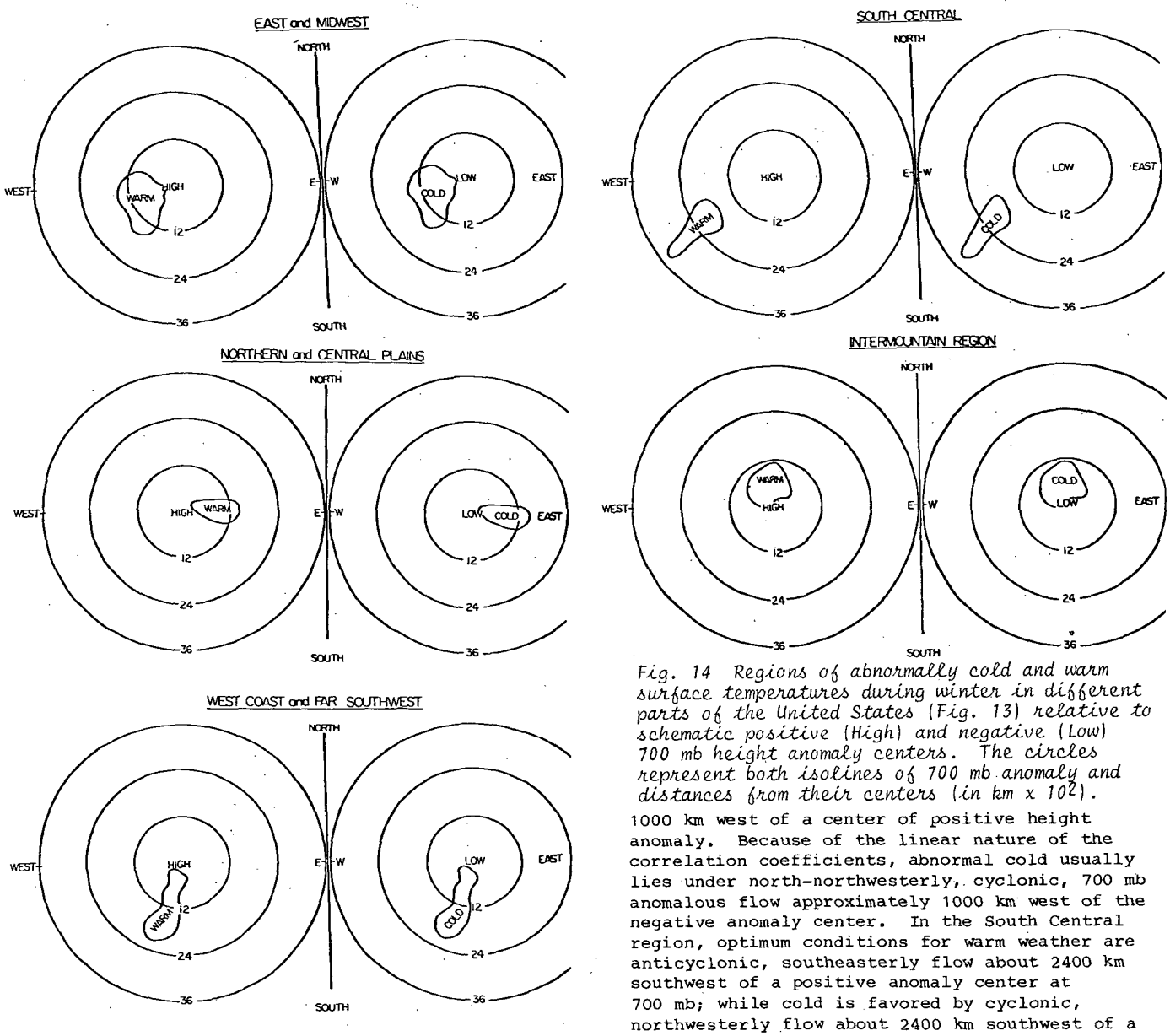


Fig. 14 Regions of abnormally cold and warm surface temperatures during winter in different parts of the United States (Fig. 13) relative to schematic positive (High) and negative (Low) 700 mb height anomaly centers. The circles represent both isolines of 700 mb anomaly and distances from their centers (in km x 10^2).

the United States can be divided into the five climatologically homogeneous areas depicted in Fig. 12. The positive anomaly center is located from 300 to 1700 km east-northeast of stations in the eastern third of the nation, from 1600 to 3300 km northeast of stations in the South Central region, from 300 to 2200 km north-northeast of cities on the Pacific Coast and in the desert Southwest, from 100 to 1200 km south of stations in the Rocky Mountains and Plateau Regions, and from 200 to 1400 km west of cities in the Great Plains. The centers for all but 1 of the 109 cities fell into these homogeneous areas, which are delineated in Fig. 13.

The relation between temperature and circulation in each area is illustrated schematically in Fig. 14. For example, above normal temperatures in the East and Midwest are likely to occur under anticyclonic, south-southeasterly, 700 mb anomalous flow about 1000 km west of a center of positive height anomaly. Because of the linear nature of the correlation coefficients, abnormal cold usually lies under north-northwesterly, cyclonic, 700 mb anomalous flow approximately 1000 km west of the negative anomaly center. In the South Central region, optimum conditions for warm weather are anticyclonic, southeasterly flow about 2400 km southwest of a positive anomaly center at 700 mb; while cold is favored by cyclonic, northwesterly flow about 2400 km southwest of a negative 700 mb anomaly center. In the Northern and Central Plains, above normal temperatures are favored by northerly, anticyclonic flow about 800 km east of the positive 700 mb anomaly center, while cold tends to occur under southerly, cyclonic, flow approximately 800 km east of the negative center. This reflects a tendency for cold air outbreaks to occur ahead of the trough at 700 mb in the surface upslope flow, and vice versa for warm downslope flow. In the Rocky Mountain and Plateau regions, abnormal warmth generally goes with westerly anomalous flow of mild Pacific air about 600 km north of the positive 700 mb center; cold is usually found an equal distance north of the negative center under easterly flow of continental air. Finally, the Pacific Coast and Far Southwest tend to experience their warmest weather under southeasterly anomalous flow about 1200 km southwest of the positive center, and their coldest weather under northwesterly flow about an equal distance and direction from the negative center.

4. UPSTREAM NEGATIVE CORRELATION

Of course, the positive correlation centers are not the only factors affecting temperature. The centers of negative correlation upstream are almost as important. Fig. 15 gives the maximum negative correlation coefficient for each station. Values range from -.87 along the coast of northern California to -.33 in Maine. The correlation magnitudes are

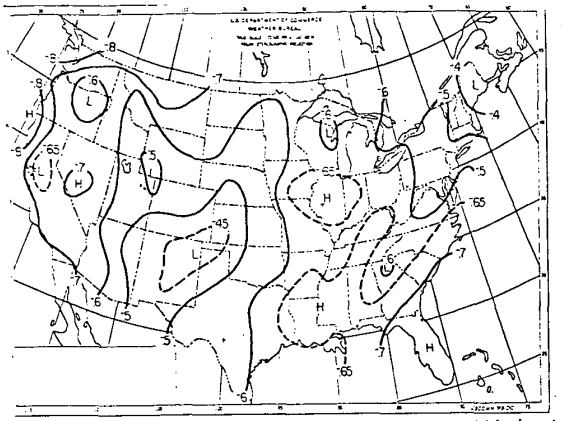


Fig. 15 Maximum negative correlation coefficient between concurrent anomalies of monthly mean winter surface temperature and 700 mb height at upstream negative center, as analyzed on each of 109 correlation fields.

generally larger in the northwest and southeast quadrants of the country than they are in the Northeast or Southwest. In the latter region, a large area with absolute values below -.45 stretches from southern New Mexico to southwestern Kansas.

The difference between the maximum positive correlation (Fig. 6) and the maximum negative correlation (Fig. 15) is presented in Fig. 16. When the former exceeds the latter, the numbers in Fig. 16 are positive and the surface temperature is affected more by the local 700 mb height than by heights upstream. This situation prevails throughout the eastern third of the United States as well as in the Rocky Mountain and Plateau regions, with maximum values of .38

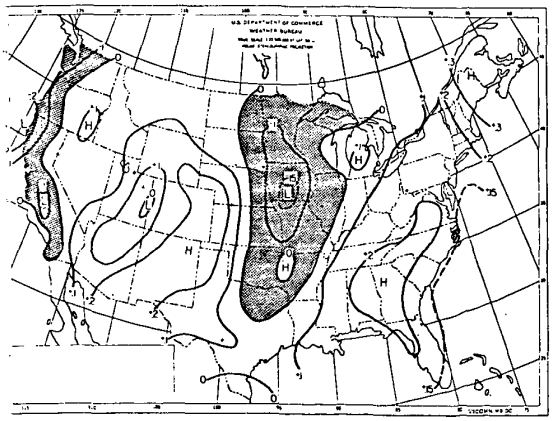


Fig. 16 Difference between maximum positive correlation of Fig. 6 and maximum negative correlation of Fig. 15. Values are shaded (negative) in areas where negative exceeds positive correlation in absolute value.

at Caribou and .28 at Albuquerque. However, the opposite condition is found in the central part of the country and along the Pacific Coast. Here the negative values (shaded) delineate areas where the circulation in a key region upstream is more important than the local circulation. This situation prevails in about 25% of all stations in the country.

If we plot the magnitudes of the largest correlation coefficients in Fig. 6 and 15, regardless of sign, we obtain the results shown in Fig 17. Here the maximum absolute correlation ranges from .92 at Tampa to .48 at Des Moines. Coefficients exceed .8 in most of

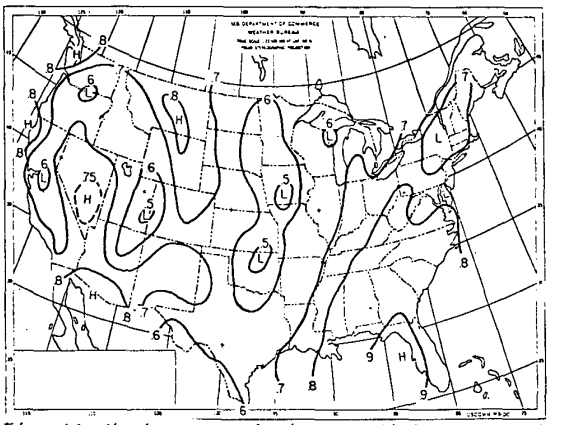


Fig. 17 Maximum correlation coefficient, regardless of sign, between concurrent anomalies of monthly mean winter surface temperature and 700 mb height, as analyzed on each of 109 correlation fields.

the Southeast and in small parts of the Far Southwest, Northern Rockies, and Pacific Coast. However, values are less than .6 in much of the central United States and southern Plateau Region, as well as parts of the interior valleys of the West Coast States. The correlation analysis of Fig. 17 indicates how well surface temperature can be specified from a single 700 mb height, regardless of location.

When the location of each negative anomaly center on the composite maps is plotted relative to the reference station, a grouping by four regions of the country can be made (Fig. 18).

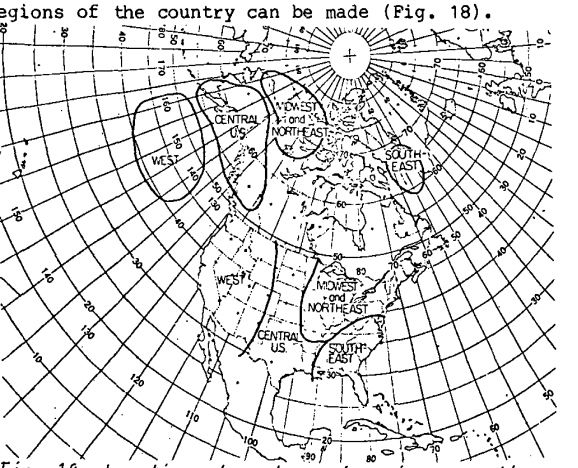


Fig. 18 Location of centers of maximum negative 700 mb height anomaly for 102 winter months and 109 composite maps. Heavy lines delineate areas where centers cluster for stations in four parts of the United States.

All negative centers cluster in well defined areas, located approximately in the Davis Strait for all cities in the southeast United States, in the Alaskan and western Canadian Arctic for cities in the Midwest and Northeast, in the Gulf of Alaska and surrounding land masses for staions in the central United States, and in the northeast Pacific for cities in the western third of the nation.

When the locations of the negative centers are plotted on polar coordinate paper with the reference station at the origin, only three homogeneous areas emerge (Fig. 19). All stations in the Southeast have centers from 3000 to 4600 km northeast of the reference station, all cities in central and northeastern portions (except Caribou) have centers from 3000 to 4700 km northwest of the reference station, and all cities in and west of the Rockies have centers from 2200 to 3800 km to the west-northwest. The precise location of these areas on a map of the United States is shown in Fig. 20.

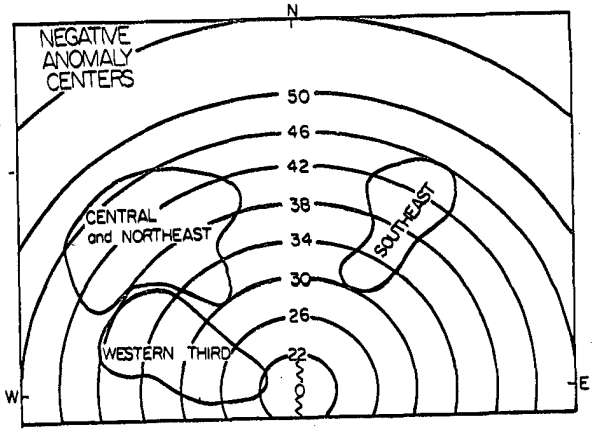


Fig. 19 Distance (km x 10^2) and direction (north, east, south, or west) from reference station (at origin) to centers of maximum negative 700 mb height anomaly in three different parts of the United States on 109 composite maps.

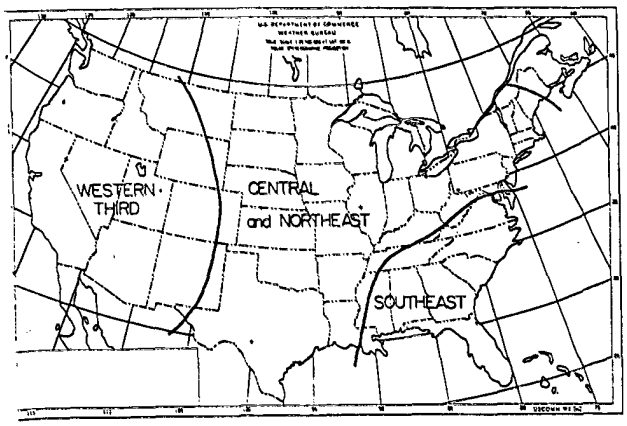


Fig. 20 Boundaries of three homogeneous zones in the United States relative to location in Fig. 19 of maximum 700 mb height anomaly during winter.

The joint effect of the positive and negative anomaly centers is depicted in Fig. 21 (next page). This schematic diagram was obtained by combining Fig. 12 and 19 for each of six homogeneous areas into which the United States may be divided (Fig. 13 with Southeast as separate area). For example, in the Southeast, above normal surface temperatures usually occur with a center of negative 700 mb height anomaly about 3800 km northeast of the reference station and a positive anomaly center about 1000 km to the east. In the Northeast and Midwest, above normal temperatures go with below normal heights centered about 4000 km to the northwest and above normal heights about 1000 km to the east. In the South Central region, the warmest weather is generally associated with a negative height anomaly center about 4000 km to the northwest of the station and above normal heights about 2400 km to the northeast. In the Northern and Central Plains, warmest temperatures go with a negative anomaly center about 4000 km to the northwest and above normal heights about 800 km to the west. In the Rocky Mountain and Plateau Regions, the preferred configuration for warm weather is a negative anomaly center about 3000 km west-northwest of the reference station and above normal heights about 600 km to the south. Finally, in the Far Southwest and along the Pacific Coast, warmest conditions are accompanied by below normal heights about 3000 km to the west-northwest and above normal heights about 1200 km to the northeast. Because of the linear nature of the correlation coefficients, optimum conditions for cold weather are exactly opposite to those described above for warmth.

5. CONCLUDING REMARKS

Thus far we have discussed the influence of only two 700 mb height anomaly centers on winter surface temperature. In order to assess the effects of the entire height field in an objective fashion, multiple regression equations were derived by applying a stepwise selection program to the entire grid of 133 points. The development of these specification equations, their performance on independent data, and their operational application have been described in a companion paper (Klein, 1983). In addition, the relations between surface temperature and 700 mb height during summer are currently being investigated by means of correlation fields and height anomaly maps; results and seasonal differences will be discussed in future papers.

ACKNOWLEDGEMENT

This work was sponsored by the U.S. Department of Energy, Grant DE-AS05-81EV10539, as part of the National Climate Program. Valuable help was received from Bob Kaylor, Charlene Mann, Caren Klarman Grim, and Trish Palasik at the University of Maryland; data were supplied by Bob Taubensee and Dave Durdahl at the Climate Analysis Center of NOAA.

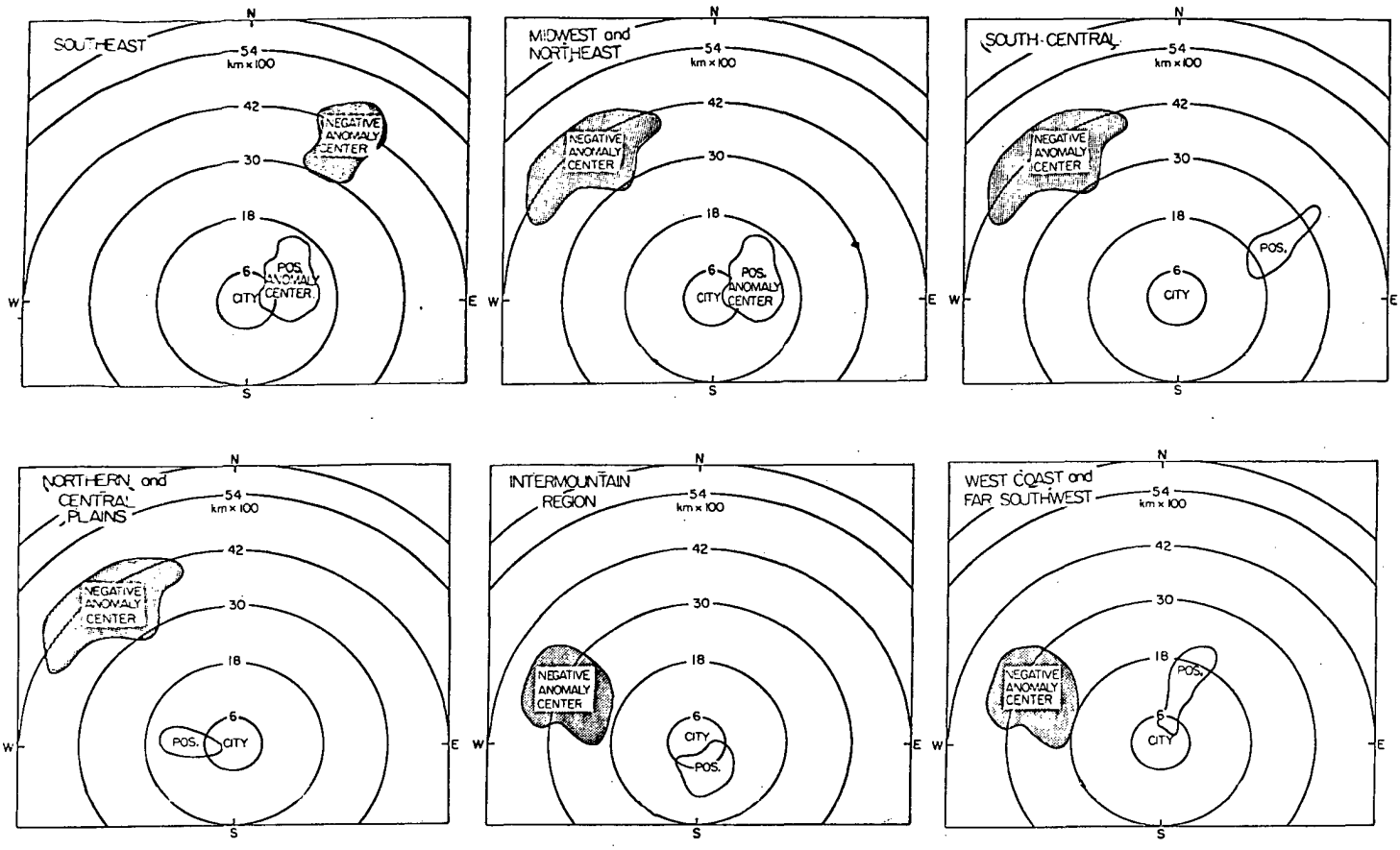


Fig. 21 Schematic diagram showing optimum locations of positive and negative (shaded) 700 mb height anomaly centers for warm months during winter in six parts of the United States (Fig. 13), with reference station at origin.

REFERENCES

Klein, W.H., 1965: Application of synoptic climatology and short-range numerical prediction to five-day forecasting. Weather Bureau Research Paper No. 46, U.S. Dep't. of Commerce, Washington, DC, 109 pp.

_____, 1983: Objective specification of monthly mean surface temperature from mean 700 mb heights in winter. Mon. Wea. Rev., 111, April issue.

_____, B.M. Lewis, and I. Enger, 1959: Objective prediction of five-day mean temperatures during winter. J. Meteor., 16, 672-682.

Martin, D.E. and H.F. Hawkins, Jr., 1950: Forecasting the weather - the relationship of temperature and precipitation over the United States to the circulation aloft. Weatherwise, 3, pp. 16-19, 40-43, and 65-67.

Namias, J., 1953: Thirty-day forecasting: a review of a ten year experiment. Meteor. Monogr., 2, 83 pp.

_____, 1981: Teleconnections of 700 mb height anomalies for the Northern Hemisphere. CALCOFI Atlas NO. 29, Scripps Institution of Oceanography, University of California, San Diego, LaJolla, CA 92093, 265 pp.

Stidd, C.K., 1954: The use of correlation fields in relating precipitation to circulation. J. Meteor., 11, 202-213.

Walker, G.T. and E.W. Bliss, 1932: World Weather V. Mem. Roy. Meteor. Soc., 4, 53-84.

Wallace, J.M. and D.S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. Mon. Wea. Rev., 109, 784-812.

SPECIFICATIONS OF MONTHLY PRECIPITATION IN THE WESTERN UNITED STATES FROM MONTHLY MEAN CIRCULATION

Bryan C. Weare and Marc A. Hoeschele

Department of Land, Air and Water Resources
University of California
Davis CA 95616

1. INTRODUCTION

Experimental attempts at forecasting monthly precipitation one or more months in advance have often relied upon forecasting mean circulation patterns and then inferring the associated precipitation from those patterns (Namias, 1980). The circulation forecasts have depended on a combination of methods including analogs, trends, teleconnections and air-sea interactions (Namias, 1978). The inferences of associated rainfall have relied upon (Namias, personal communication) the earlier statistical studies of Stidd (1954) and Klein (1963). A number of temperature and precipitation forecast schemes have been recently evaluated by Preisendorfer and Mobley (1982). They find that the quality of predictors is rather low and differs quite substantially for different geographic regions and time periods.

Most researchers have attempted to improve the veracity of forecasts by investigating new methods of forecasting the circulation patterns. Namias (1975) in a series of papers has discussed and utilized North Pacific sea surface temperatures in this regard. Relatively little recent work has been carried out on understanding and improving the quality of the inferences of monthly rainfall from current monthly mean circulation indices (Harnack, 1979). This seems especially useful at this time because ideas concerning methods of multiple regression model specification have been altered by the work of Davis (1976, 1978), Barnett and Preisendorfer (1978), Barnett and Hasselman (1979), and others.

For instance, Davis (1976) emphasizes the importance of an a' priori choice of a small number of predictor variables. He does this in his analysis of expected "hindcast" and "forecast" skills, i.e., those derived from use of a regression model on the data utilized to develop it and from use of the model on independent data. Davis shows that the hindcast skill is inflated by "artificial" skill which is directly proportioned to the number of predictors (M) and inversely proportioned to the number of degrees of freedom (N). Furthermore, since many geophysical data are autocorrelated, the number of degrees of freedom is often far less than the number of samples. Davis (1976) summarizes these inferences in the form

$$\langle S_H \rangle \cong S + \frac{M}{N} \qquad (1)$$

$$\langle S_F \rangle \cong S - \frac{M}{N} \qquad (2)$$

where $\langle S_H \rangle$ and $\langle S_F \rangle$ are the expected hindcast and forecast skills, S is the "true" skill, and M/N is the artificial skill. Thus, while additional predictors may appear to improve a model in terms of its hindcast skill they may actually be leading to a deterioration of forecast skill. Finally, Davis (1978) argues that "screening" procedures that reduce the number of predictors a' posteriori give rise to much larger artificial predictability than equivalent models whose preditors were chosen a' priori.

The earlier correlation studies of Stidd (1954) and Klein (1963) preceded these developments. Furthermore, reexamination of the specification problem at this time seems worthwhile because circulation data is now available for many more years. Finally, the earlier studies had attempted to specify average precipitation over state-size regions, whereas precipitation in smaller regions important to agriculture or water management problems is perhaps more relevant.

Therefore we have carried out a pilot study exploring our ability to specify monthly precipitation for five "winter" months at 36 stations in Washington, Oregon and California using indices of circulation variables for the same months. It should be emphasized that the following comparisons are all of the ability to specify rather than forecast monthly precipitation. The circulation indices used in developing the specification models were the time coefficients of the three most important eigenvectors (Kutzbach, 1967) of monthly mean circulation variables for the period 1951-76. The eigenvector analysis provides an a'priori method of reducing the large number of potential predictors (i.e., grid values over the Northern Hemisphere) to a very few predictors to be used in the statistically derived specification equations. Data for 1977-80 were used to test the models on new data.

The series of "experiments" to be described compare the ability of various models to specify precipitation relative to the ability of a reference model. In each case the differences in specification ability in both space and time are compared to the reference hindcast. In this way it is hoped that a better understanding may be gained of the limits of our ability to specify monthly precipitation with monthly mean circulation variables.

2. DATA

There are a large number of potential atmospheric circulation variables at many levels that might be of use in specifying rainfall. 700 mb geopotential height has been chosen as the primary variable because it has been traditionally used in long-range forecasting and because earlier studies indicate that it is moderately linked to precipitation amounts. Monthly mean 700 mb geopotential heights (Z_{700}) at 5° latitude by 10° longitude intersections, north of 20°N, for the period December 1946-December 1980 were kindly supplied by Professor Jerome Namias. From this data we also calculated finite difference estimates of the monthly relative vorticity (ζ_{700}) at the same intersection points (except ing the northern and southernmost perimeters).

Rainfall data for the "cooperative" stations in Washington, Oregon and California for 1951-80 were obtained from the National Climatic Center. These data, which are used to derive climatological normals, are serially complete, since the few missing values were estimated from the records of nearby stations. From the stations available we chose a group of 36 stations which had as few estimated values as possible and represented as well as possible most of the climatic and topographic regions in these three states. The locations of these stations are shown in Fig. 1.

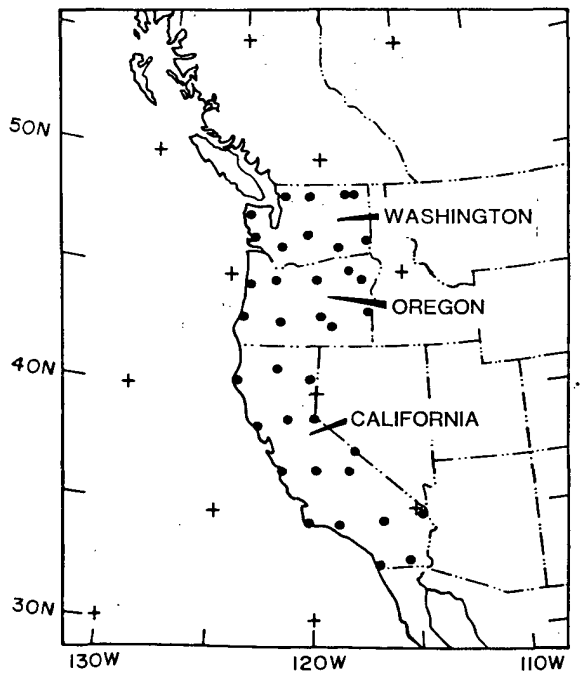


Fig. 1. Locations of "cooperative" precipitation stations used in this study (dots) and of 12 points at which circulation data are derived in the "reference" experiment (plusses).

The circulation variables were subjected to eigenvector analysis (Kutzbach, 1967) in order to define a few indices which describe most of the variability. It seemed to be an open question as to how large a spatial domain one should use in the eigenvector analysis. If one chose too small a domain, one might exclude a region which had strong associations with the rainfall in question. If, on the other hand, too large a region were chosen, the dominant eigenvectors, explaining less of the overall variance, may largely be describing the variability of a region unimportant to the specification. To explore this problem we chose to look at three successively larger domains. The twelve points of the smallest domain, for which most of the studies were carried out, are indicated by plusses on Fig. 1. The other regions, whose data were subjected to the eigenvector analyses, were simply defined by expanding the longitudinal extent in both directions until twenty four or forty eight 5° x 10° grid points were included. The time coefficients of the eigenvectors of the data sets will be designated by their type, level and number of grid points so that, for instance, $Z_{700}(12)$ refers to the time coefficients of the eigenvector analysis of the 700 mb height at the twelve locations illustrated in Fig. 1.

Since the significance of the derived specifications relations depends strongly on the number of predictors, the following examples use only the time coefficients of the three most important eigenvectors of the circulation variables to specify the rainfall. This is done despite the fact there are moderate differences in the cumulative variance explained for the different variables over different domains. For instance, the first three eigenvectors of $Z_{700}(12)$ explain a cumulative percent of the variance at those twelve points of between 94.2 and 92.6 for different calendar months from November through March. The first three eigenvectors of $Z_{700}(24)$, $Z_{700}(48)$ and $\zeta_{700}(12)$ explain together approximately 84, 72, and 85 percent of the variance, respectively.

3. RESULTS

All of the statistical regression models were developed in a similar manner. The model using 700 mb heights at 12 points was considered the reference to which the other results were compared. In this case, as with the other experiments, linear multiple regression equations for each rainfall station for each month from November through March were derived from the available circulation and rainfall data for the period 1951-76. The specifier variables in this reference experiment are the time coefficients of the three most important eigenvectors in the analysis of the 700 mb heights at the 12 points shown in Fig. 1 for the sub-period 1951-76. Thus for a month m, year y and station i,

precipitation P_{myi} may be specified using the 700 mb heights for the same month by

$$P_{myi} = \alpha_{0mi} + \alpha_{1mi} Z(12)_{1my} + \alpha_{2mi} Z(12)_{2my} + \alpha_{2mi} Z(12)_{3my} + E_{myi} \quad (3)$$

where the α's are regression coefficients derived from the 1951-76 data set (Weisberg, 1980), the Z(12)'s are the time coefficients of the eigenvector analysis, and the E's are the errors resulting from imperfect specification.

Variable	Fraction of Variance					
	Nov.	Dec.	Jan.	Feb.	Mar.	Ave.
$Z_{700}(12)$	.59	.47	.48	.39	.46	.48
$Z_{700}(24)$	.57	.45	.46	.37	.45	.46
$Z_{700}(48)$	.42	.34	.46	.29	.41	.38
$\zeta_{700}(12)$	.54	.47	.48	.42	.42	.47

Table 1. Average fractions of the variance of rainfall explained by multiple linear regres sion equations with the independent variables indicated for the same month in the years 1951-76. The average is over all the stations in Fig. 1.

Fig. 2 shows the fraction of the monthly rainfall variance specified by the three time coefficients of the 700 mb height averaged over the 5 calender months for 1951-76. Table 1 shows the fraction of explained variance averaged over all stations for each calender month and the 5-month average. The spatial patterns of the specification maps for the individual months is quite similar to the average map in Fig. 2. This figure suggests the best specification is possible along the coast, especially from central California through Oregon. The poorest specifications are along the ridge of the Sierra Nevadas in California and in the high, relatively dry, plains of Washington and Oregon. This general pattern is probably indicative of the fact that most coastal rainfall is due to frontal systems which are enhanced by orographic uplift, whereas the rainfall in the eastern areas is partially influence by convective activity. Furthermore, the high mountains of much of the eastern region have considerable amounts of snow, which is probably not measured with the accuracy of rainfall.

This reference experiment was modified using the time coefficients of two or four eigenvectors. The results in both cases were quite similar, except that the two parameter case resulted in slightly lower explained variance.

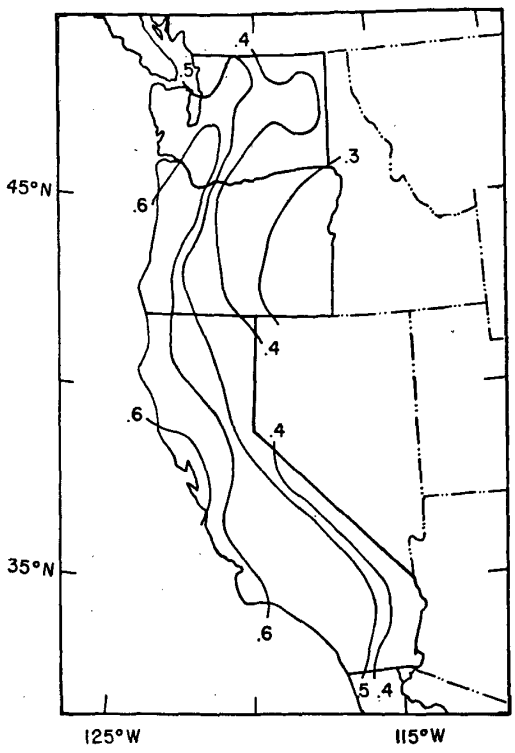


Fig. 2. Subjective analysis of the fraction of the variance in monthly precipitation at 36 sites (Fig. 1) which is specified by three time coefficients of an eigenvector analysis of 700 mb geopotential heights at 12 points (Fig. 1) for 1951-76. The values are the averages of the results for November, December, January, February and March.

The input data of the reference experiment was altered in two other ways. In the first case the time coefficients were derived using the 700 mb data at the 24 points; in the second the coefficients correspond to the 48 point eigenvector analyses. Table 1 also indicates the average results for these two cases. Overall, the use of the 24 point data set results in a slight deterioration of specification. The use of the 48 point set results in a substantial reduction in explained variance for all months but January. Fig. 3 illustrates the average fraction of explained variance for all five months (corresponding to Fig. 2) when the 48 point data set is used. Comparisons of Fig. 2 and 3 suggest that the use of the 48 point set results in a deterioration of specification ability at all stations with the greatest losses of up to 0.27 in western Washington. This strongly suggests that the size of the spatial domain of the specifier field is very important.

The reference experiment was again altered. This time the first three time coefficients of the 700 mb relative vorticity at the twelve points shown in Fig. 1 were used to specify the same month's precipitation for

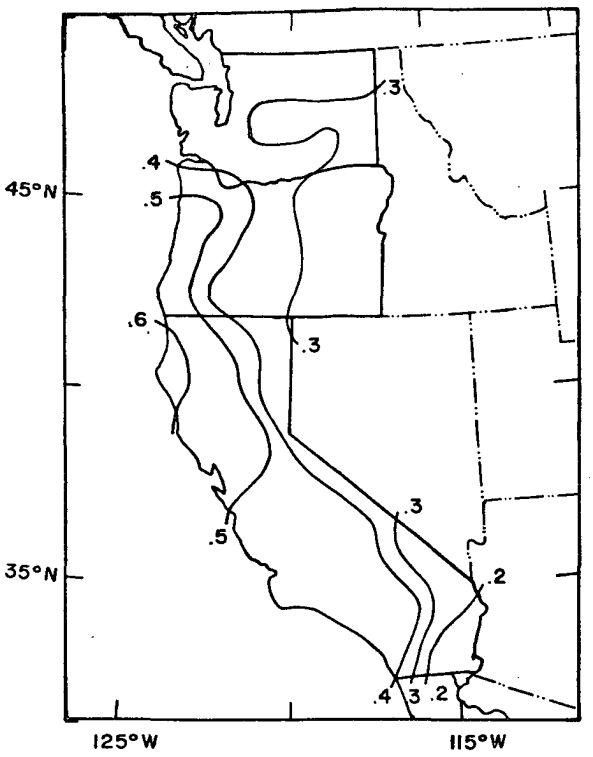


Fig. 3. Fraction of the variance specified by three time coefficients of 700 mb height at 48 points (see text) as in Fig. 2.

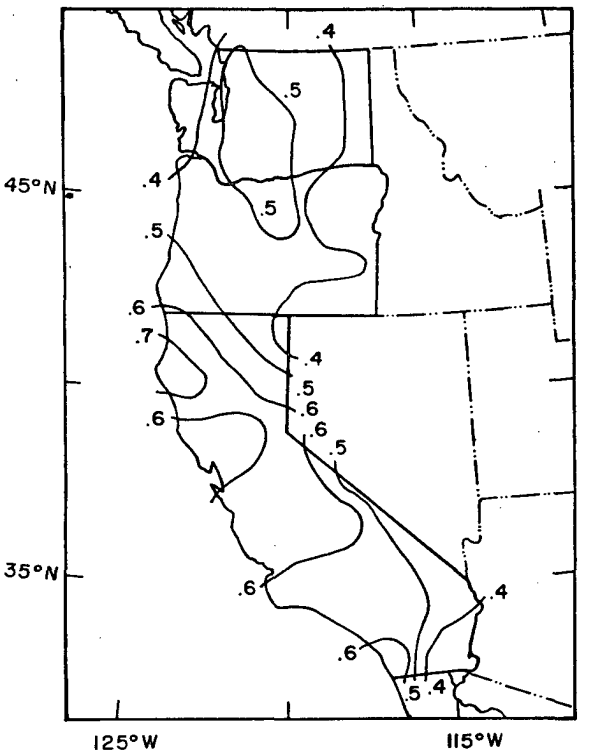


Fig. 4. Fraction of the variance specified by three times coefficients of 700 mb relative vorticity at 12 points as in Fig. 2.

each of the five calendar months. Fig. 4 illustrates the average fraction of the total precipitation variance at each station for the same 1951-76 period for all five months. Table 1 indicates the average over all stations for each calendar month. As may be seen from Table 1 the overall ability of the relative vorticity to specify the precipitation is quite comparable to the geopotential height. Comparison of Figs. 2 and 4, however, indicate that the patterns of the ability to specify differ somewhat for the two data sets. In western Washington and Oregon the height data explains from 0.04-0.14 more of the precipitation variance than do the vorticities. In the remainder of the region, however, differences between the two results are quite small with the vorticities being slightly superior.

In an attempt to better understand the specification abilities of these various models using "new" data, they were used to predict precipitation amounts for 1977-80 using the appropriate 1977-80 circulation data indices. Table 2 shows the fraction of the measured precipitation variance explained by the model specifications averaged over all 36 stations for each calendar month and the 5-month mean. As would be expected from Eqn. (2) nearly all of the values are smaller than the equivalent hindcast results in Table 1. In fact some values in Table 2 are negative indicating that the average specification was worse than simply "guessing" the 30-year mean precipitation amount. It should also be noted that the total values of the explained variances are quite variable from month to month. This is probably not surprising if one considers that these statistics are generated from only four years of "new" data, one of which 1977-78 was a major drought year in California.

Variable	Fraction of Variance					
	Nov.	Dec.	Jan.	Feb.	Mar.	Ave.
$Z_{700}(12)$	.32	.42	.52	.27	.06	.32
$Z_{700}(24)$	-.24	.05	.07	-.09	-.26	-.10
$Z_{700}(48)$	.35	-.09	.10	.13	-.25	.05

Table 2. Average fractions of the variance of rainfall explained by the multiple linear regression equations derived from data for 1951-76 when applied to "new" data for 1977-80. The average is over all the stations in Fig. 1.

Nevertheless, the results shown in Table 2 tend to confirm the previous conclusion that the $Z_{700}(12)$ is the best overall specifier of precipitation with the greatest skills in December and January. The models using the 24 and 48 point circulation data sets give consistently poorer results with the only

exception being $Z_{700}(48)$ in November. The larger errors using the 24 and 48 point data probably develop from the fact that those eigenvectors reflect changes far from the region of the precipitation stations.

Fig. 5 shows the fraction of explained variance specified by $Z_{700}(12)$ for 1977-80 averaged over the five calendar months November through March. The pattern is somewhat noisier than for the hindcast shown in Fig. 2. Nevertheless, broad features in Fig. 2 and 5 are quite similar. In both, the highest skills are in coastal northern California-Oregon, and the lowest are in southeastern California and eastern Oregon and Washington. The average skills in Fig. 2 and 5 are given in the right-hand columns of Tables 1 and 2. These indicate that on average the specifications on "new" data explain about 0.16 less of the variance than was possible on the data used to develop the models. A comparison of Fig. 2 and 5 suggest that the differences range from less than 0.1 in some coastal regions to greater than 0.4 in the poor specification locales. Maps of average skills (not shown) using 1977-80 height data at the 24 and 48 points are similar to Fig. 5 except that they show overall lower values and higher variability.

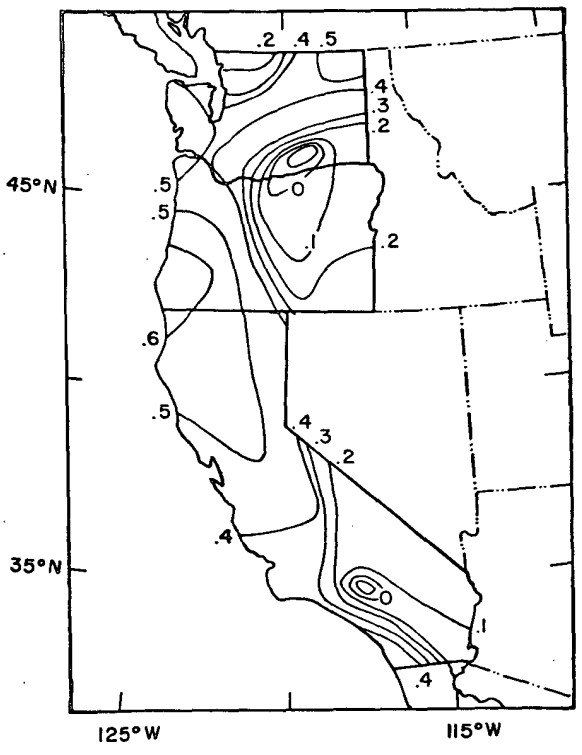


Fig. 5. Subjective analysis of the fraction of variance in monthly precipitation specified by 700 mb heights at the 12 points (Fig. 1) for "new" data during 1977-80 as in Fig. 2.

In general, the tests with "new" data confirm that regions of good skill in the hindcasts tend to remain good when the models are applied to new data. Furthermore, those models using $Z_{700}(12)$ which give the best skills in the hindcasts tended to do far better when tested with "new" data than the models based on the $Z_{700}(24)$ or $Z_{700}(48)$. However, these "new" data results must be interpreted quite cautiously since they involve tests of only four years and therefore cannot be seen as true expected values.

4. DISCUSSION

The purpose of this research was to better understand the factors which control the ability to specify monthly precipitation from monthly mean circulation parameters. The data in Table 1 suggest that a maximum of between ~35 and 60 percent of the precipitation variance may be specified by three 700 circulation variables. In general the ability to specify coastal stations is appreciably greater than that for stations in the interior mountains or high plains of eastern Oregon and Washington. This points out the possible danger of specifying precipitation for a coarse grid or large area averages and assuming that those specifications are equally reliable over the entire region. The abilities of the geopotential heights or relative vorticities seem comparable.

Since we have chosen to follow a single methodology (using the time coefficients of the three most important eigenvectors as the dependent data in the multiple regression equations), we have not fully explored the optimization of the rainfall specification even for the chosen variables. For instance one might circumvent the eigenvector analyses and calculate the regression coefficients for circulation variables chosen at sites dictated by a "screening" analysis (Klein, 1981). We have chosen not to do so because we believe, following the work of Davis (1976) and others, that the present method is likely to give more statistically significant results even though the fraction of explained precipitation variance for dependent data may be somewhat higher from the screening analyses. The question of significance is especially important for the proposed use of any derived specification relations, i.e., interpretation of month ahead forecasts of monthly mean circulation parameters. In such cases one would hope for a correct overall pattern of the circulation but expect relatively large errors at individual points. In this case it would seem likely the use of the eigenvector analysis, which provides coefficients based upon weightings of the entire circulation field, would be more reliable than using "screening" equations which focus on specific locations which may have considerable errors.

Finally our analysis suggests a result which may be important to water management decisions in the western United States. Many reservoirs in this region derive their water

from the mountainous regions of eastern California or central Washington. The 700 mb geopotential heights are relatively poor specifiers of the precipitation in these regions compared with the coastal regions. This would imply that even if accurate month ahead forecasts of the 700 mb height were possible, there may be severe limitations in inferring predicted precipitation from those heights. On the other hand Fig. 4 suggests that in the above mentioned region the 700 mb relative vorticity may be a better specifier. Thus the variable of choice may be a function of the requirements of the forecast.

Acknowledgments. This work was supported by ATM-8241336 from the National Science Foundation and funds from the Department of Agriculture.

REFERENCES

Barnett, T. P. and K. Hasselmann, 1979; Techniques of linear prediction, with application to oceanic and atmospheric fields in the tropical Pacific. Rev. Geophys. Space Phys., 17, 949-968.

_____and R. W. Preisendorfer, 1978; Multifield analog prediction of short-term climate fluctuations using a climate state vector. J. Atmos. Sci., 35, 1771-1787.

Davis, R. E., 1976; Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. J. Phys. Oceanogr., 6, 249-266.

_____, 1978; Predictability of sea level pressure anomalies over the North Pacific Ocean. J. Phys. Oceanogr., 8, 233-246.

Harnack, R., 1979; The statistical relationship between mean summer precipitation in the United States and the concurrent pressure distribution. Proceedings of the Forth Annual Climate Diagnostics Workshop, NOAA, Washington [NTIS, PB-201130].

Horel, J. D. and J. M. Wallace, 1981; Planetary-scale atmospheric phenomena associated with the Southern Oscillation. Mon. Wea. Rev., 109, 813-829.

Hoskins, B. J. and D. J. Karoly, 1981; The steady linear response of a spherical atmosphere to thermal and orographic forcing. J. Atmos. Sci., 38, 1179-1196.

Klein, W. H., 1963; Specification of precipitation from the 700-millibar circulation. Mon. Wea. Rev., 92, 527-536.

_____, 1981; Objective specification of monthly mean surface temperatures in the contiguous United States from 700-mb mean heights over North America and adjacent oceans during winter. Proceedings of the Sixth Annual Climate Diagnostics Workshop, NOAA, Washington [NTIS PB82-219486].

Kutzbach, J. E., 1967; Empirical eigenvectors of sea-level pressure, surface temperature and precipitation complexes over North America. J. Appl. Meteor., 6, 791-802.

Namias, J., 1975; Short period climate variations, collected works of J.Namias, 1934 through 1974, U. of Calif., San Diego, 905 pp.

_____, 1978; Long-range weather and climate predictions. Geophysical Predictions, National Academy of Science, Washington, D.C.

_____, 1980; The art and science of long-range forecasting. EOS, 61, 449-450.

Preisendorfer, R. W. and C. D. Mobley, 1982; Climate forecast verifications, U.S. Main-land, 1974-82. NOAA Technical Memorandum ERL PMEL-36.

Stidd, C. K., 1954; The use of correlation fields in relating precipitation to circulation. J. of Meteor., 11, 202-213.

Weisberg, S., 1980; Applied Linear Regression. Wiley, New York.

THE STOCHASTIC-DYNAMIC THEORY OF AIR-SEA INTERACTION APPLIED TO THE QUASIBIENNIAL AND SEMI-ANNUAL OSCILLATION OF SEA SURFACE TEMPERATURE

Mai-Tsun Li

Institute of Atmospheric Physics
Academia Sinica
Beijing, People's Republic of China

Jia-You Huang

Department of Geophysics
Peking University
Peking, People's Republic of China

1. INTRODUCTION

It is well known that there is interaction in the sea-air system. What is the mechanism of the interaction? How does it proceed? Many meteorologists have discussed these problems. For example, Frankignoul and Hasselman (1977) considered a stochastic model in which the interaction in sea-air was explained as the response of the oceanic surface layers to short-time-scale atmospheric forcing that was a white noise process. Later, Wright (1979) and Nicolis (1979) had a similar idea. They suggested that it might be necessary to add some random force to some systems of equations describing the interaction in the sea-air system.

In fact, it seems that the sea-air interaction can not be regarded as resulting only from the random force from the atmosphere. It is necessary to consider forces both in the atmosphere and in the ocean. The process of the interaction may be a complicated feedback one that is adjustable and conformable (Li, 1981). In order to simulate the processes, we will propose a second-order stochastic fluctuation model, which is a system of differential equations of the sea-air couple which is developed from the atmosphere-ocean motion and energy balance equations.

2. STOCHASTIC MODEL

According to the thermal balance equations in the atmosphere and not considering level transfer, we can write the local air temperature variation with time as

$$\frac{\partial T}{\partial t} = (H_S + H_L + H_R)/\rho C_p H \qquad (1)$$

where T is air temperature, H_S, H_L, and H_R are the sensible, latent, and radiative heat flux in air, ρ and C_p are the density and specific heat of air, and H is the isogeopotential height of the barotropic atmosphere.

Similarly, without considering the effects on salt content, diffusion, and advection in the ocean, the local variation of SST can be written as

$$\frac{\partial T_S}{\partial t} = (H_S^W + H_L^W + H_R^W)/\rho^W C_p^W h \qquad (2)$$

where the variables are the same as in (1) except that the right shoulder sign ("W") denotes SST's and h is the depth of the mixed layer of ocean.

First we will discuss the input and output of the heat flux of the ocean.

According to the discussion about the mixed layer of water, the sensible and latent heat flux of water have a relationship with SST and sea surface wind velocity (see Frankignoul and Hasselmann, 1977). It is expressed as

$$H_L^W + H_S^W = C_H(1 + B)\rho C_p(T - T_S)|V| \qquad (3)$$

where C_H is the bulk transfer coefficient for the sensible heat flux, B is the Bowen ratio of the latent heat flux to the sensible heat flux, and $|V|$ is the absolute value of sea surface wind velocity. The radiation balance can be expressed as

$$H_R^W = I^W - F^W \qquad (4)$$

where I^W and F^W denote the short-wave radiation being absorbed and long-wave radiation being

given off. I^W can be expressed as

$$I^W = \frac{1}{4}\mu I_o(1 - \alpha_p{}^W)(1 - \eta_1(T)) \quad (5)$$

where I_o is the solar constant, $\alpha_p{}^W$ is the albedo of water, μ is the attenuation coefficient, and $\eta_1(T)$ denotes the attenuation of short-wave radiation by the clouds. $\eta_1(T)$ represents the capability for absorption and dispersion of water in clouds. The process happens in the atmosphere, so it is given as a function of air temperature.

The albedo of water can be described by SST,

$$\alpha_p{}^W = a^W - b^W T_S$$

where a^W and b^W are constants (larger than zero). Hence the short-wave radiation reaching the ocean surface is

$$I^W = [\frac{1}{4}\mu I_o(1 - a^W) + \frac{1}{4}\mu I_o b^W T_S](1 - \eta_1(T)). \quad (6)$$

In the same way, the long-wave radiation given off from water can be denoted as

$$F^W = \varepsilon^W \sigma T_S{}^4(1 - \eta_2(T)) - \varepsilon\sigma T^4 \quad (7)$$

where ε^W and ε are the effective emissivities of water and air, σ is Boltzmann's constant, and $\eta_2(T)$ is the same as $\eta_1(T)$ except that it is for long-wave radiation.

Combining (6) and (7) with (4), we have the radiation balance of the ocean,

$$H_R{}^W = \frac{1}{4}\mu I_o(1 - a^W + b^W T_S)(1 - \eta_1(T)) - \sigma[\varepsilon^W T_S{}^4(1 - \eta_2(T) - \varepsilon T^4]. \quad (8)$$

As for the input and the output of the atmosphere, they are similar to those of the ocean. The transfer term of the sensible and latent heat can be written as

$$H_S + H_L = \rho C_p C_D(T_S - T)|V_o| + [L\rho K_T \bar{\gamma} \frac{\partial \ln \tilde{e}_S}{\partial \bar{T}} \frac{\partial \tilde{q}_S}{\partial T}]_o T_S \quad (9)$$

where C_p is the viscous drag, $|V_o|$ is the absolute value of sea surface wind velocity, $\tilde{q}_S$ and $\tilde{e}_S$ are the saturation-specific humidity and the saturation vapor pressure of moist air with respect to water, $\bar{\gamma}$ is the average temperature lapse rate, K_T is the thermal conductivity of turbulence, and L is the latent evaporation.

The short-wave radiation in the atmosphere is

$$I = \frac{1}{4}\mu I_o(1 - \eta_3(T))$$

where $\eta_3(T)$ is the attenuation of radiation in the air by the clouds, like $\eta_1(T)$ for the ocean. The long-wave radiation given off in the atmosphere is denoted as

$$F = 2\varepsilon\sigma T^4 - \varepsilon^W \sigma T_S{}^4(1 - \eta_4(T))$$

where $\eta_4(T)$ is the attenuation of long-wave radiation by the clouds for the ocean. Thus the radiation balance is written as

$$H_R = \frac{1}{4}\mu I_o(1 - \eta_3(T)) - 2\varepsilon\sigma T^4 + \varepsilon^W \sigma T_S{}^4(1 - \eta_4(T)). \quad (11)$$

The heat input and output per unit thickness of water layer can be expressed as Q^W. It is expressed as

$$Q^W = (H_L{}^W + H_S{}^W + H_R{}^W)/hC_p\rho^W, \quad (12)$$

meaning it is considered a function of SST, air temperature, and the other variables in the atmosphere (e.g., wind velocity and relative humidity). Thus

$$Q^W = Q^W(T_S, T, X_i)$$

where X_i denotes the variables in the atmosphere besides those for SST and air temperature. These variables can be resolved into a climate average part (normal) and a perturbation part (anomaly). We have

$$T_S = \bar{T}_S + T_S', \; T = \bar{T} + T', \text{ and } X_i = \bar{X}_i + X_i'.$$

Q^W is similarly expressed as

$$Q^W = \bar{Q}^W + Q^{W'}.$$

It can be expanded as a Taylor series with respect to the balance state ($T = T_S = T_o$). Omitting high-order terms, we have

$$Q^W = \bar{Q}^W + \frac{\partial Q^W}{\partial T_S}|_{T_S=T_o} T_S' + \frac{\partial Q^W}{\partial T}|_{T=T_o} T' + \Sigma \frac{\partial Q^W}{\partial X_i}|_{T_o} X_i'.$$

The variable part of anomalies of SST with time is owing to the anomalies of the heat flux of water. Hence we have

$$\frac{\partial T_S'}{\partial t} = \frac{\partial Q^W}{\partial T_S}\Big|_{T_S=0} T_S' + \frac{\partial Q^W}{\partial T}\Big|_{T=T_o} T' + \Sigma \frac{Q^W}{X_i}\Big|_{T_o} X_i'.$$

Combining (3), (8), and (12) with the above equation, we have

$$\frac{\partial T_S'}{\partial t} = -\lambda_3 T_S' + \lambda_4 T' + W_2(t) \tag{13}$$

where

$$\lambda_3 = \frac{1}{h\rho^W C_p^W}\left\{C_H(1+B)\rho C_p|V| - \frac{1}{4}\mu I_o b^W(1 - \eta_1(T)) + 8\sigma\varepsilon^W(1-\eta_2(T_o))T_o^3\right\},$$

$$\lambda_4 = \frac{1}{hC_p^W\rho^W}\left\{C_H(1+B)\rho C_p|V| - \frac{1}{4}\mu I_o(1 - a^W + b^W T_o)\frac{\partial \eta_1(T)}{\partial T}\Big|_{T=T_o} + \sigma\varepsilon^W T_o^4 \frac{\partial \eta_2(T)}{\partial T}\Big|_{T=T_o} + 4\sigma\varepsilon T_o^3\right\},$$

and

$$W_2(t) = \Sigma \frac{\partial Q^W}{\partial X_i}\Big|_{T_o} X_i'.$$

Similarly, the anomalous variable part of air temperature with time is expressed as

$$\frac{\partial T'}{\partial t} = -\lambda_2 T' + \lambda_1 T_S' + W_1(t) \tag{14}$$

where

$$\lambda_2 = \frac{1}{\rho C_p C_H}\left\{\rho C_p C_D|V_o| + \frac{1}{4}\mu I_o \frac{\partial \eta_3(T)}{\partial T}\Big|_{T=T_o} + 4\varepsilon\sigma T_o^3 + \varepsilon^W\sigma T_o^4 \frac{\partial \eta_4(T)}{\partial T}\Big|_{T=T_o}\right\},$$

$$\lambda_1 = \frac{1}{\rho C_p C_H}\left\{\rho C_p C_D|V_o| + \nu + 4\varepsilon^W\sigma T_o^3(1 - \eta_3(T))\right\},$$

and

$$W_1(t) = \Sigma \frac{\partial Q}{\partial X_i}\Big|_{T_o} X_i'$$

where ν in the λ_1 is

$$\nu = [L\rho K_T \bar{\gamma} \frac{\partial \ln \tilde{e}_S}{\partial \bar{T}} \frac{\partial \tilde{q}_S}{\partial \bar{T}}]_{T=T_o}.$$

The systems of the differential equations describing the air-sea interaction are formed by (13) and (14), i.e.,

$$\left.\begin{aligned}\frac{dT'}{dt} &= -\lambda_2 T' + \lambda_1 T_S' + W_1(t)\\ \frac{dT_S'}{dt} &= -\lambda_3 T_S' + \lambda_4 T' + W_2(t).\end{aligned}\right\} \tag{15}$$

From (15) we obtain a second-order differential equation of SST. It is

$$\frac{d^2 T_S'}{dt^2} + (\lambda_2 + \lambda_3)\frac{dT_S'}{dt} + \lambda_2\lambda_4\left(\frac{\lambda_3}{\lambda_4} - \frac{\lambda_1}{\lambda_2}\right)T_S' = W \tag{16}$$

where

$$W = \frac{dW_2(t)}{dt} + \lambda_4 W_2(t) + \lambda_2 W_1(t).$$

Because X_i is regarded as a random variable which has little effect on the sea-air system, $W_1(t)$ and $W_2(t)$ in (15) can be considered to be random forces having a white noise process.

Without considering the effect of random forces in the ocean, i.e., $W_2(t) = 0$, the variation of SST with time is owing only to the random forces in the atmosphere. The SST response spectrum can be expressed as

$$y(\omega) = \lambda_2^2 F(0)/\left[(\lambda_2\lambda_3 - \lambda_1\lambda_4 - \omega^2)^2 + (\lambda_2 + \lambda_3)^2\omega^2\right] \tag{17}$$

where $F(0)$ is the white noise spectrum and ω is the angular frequency.

Considering the random forces both in the ocean and in the atmosphere, we obtain the response spectrum of SST,

$$y(\omega) = [(\lambda_2 + \lambda_4)^2 + \omega^2]F(0) \div \left[(\lambda_2\lambda_3 - \lambda_1\lambda_4 - \omega^2)^2 + (\lambda_2 + \lambda_3)^2\omega^2\right] \tag{18}$$

where $F(0)$ is the same as in (17).

3. THE EXPLANATION FOR PERIODIC OSCILLATION OF SST

We carried out a factor analysis for SST in the Eastern Pacific tropical region (180°-90°W, 10°N-10°S), choosing 40 mesh points in it and using the monthly means during 1951-75. The spectrum analysis of the first factor of monthly SST in this region showed that there was a very significant quasibiennial oscillation (period of about 40 months) in that region (see third row in Table 1).

In order to research the physical mechanism of the quasibiennial oscillation of SST, we will simulate the observed spectrum of SST by means of its response (i.e., (18)). If the stochastic model (15) is consistent with the objective reality, the response spectrum curve of SST (i.e., (18)) must have only one maximum, which is present at $\omega = 0.16$.

Consistent with the above condition, what relationship should those feedback coefficients have in the model (15)?

In fact, let

$$\alpha = \lambda_2\lambda_3 - \lambda_1\lambda_4, \quad \beta = (\lambda_2 + \lambda_3)^2,$$

and

$$\gamma = (\lambda_2 + \lambda_4)^2.$$

Differentiating (18), we obtain ω corresponding to $y(\omega)$. It is

$$\omega = \left\{\frac{1}{2}\left[-2\gamma \pm \left(4\gamma^2 - 4(\beta\gamma - \alpha^2 - 2\alpha\gamma)\right)^{\frac{1}{2}}\right]\right\}^{\frac{1}{2}}. \tag{19}$$

Obviously, the condition necessary for a solution with real ω is

$$-2\gamma \pm \left(4\gamma^2 - 4(\beta\gamma - \alpha^2 - 2\alpha\gamma)\right)^{\frac{1}{2}} > 0,$$

i.e.,

$$4\gamma^2 - 4(\beta\gamma - \alpha^2 - 2\alpha\gamma) > 4\gamma^2$$

or

$$\beta\gamma - \alpha^2 - 2\alpha\gamma < 0. \tag{20}$$

We apply the relationship of feedback coefficients to the above inequality, and obtain

$$\frac{(\lambda_2 + \lambda_3)^2}{2} < \lambda_2\lambda_3(1 - q) + \frac{\lambda_2^2\lambda_3^2(1 - q)^2}{2(\lambda_2 + \lambda_4)^2} \tag{21}$$

where

$$q = \frac{\lambda_1\lambda_4}{\lambda_2\lambda_3}.$$

Oerlemans (1979) pointed out that the necessary condition for the stable solution in (15) is $0 < q < 1$. Hence, if the above conditions are satisfied and we make a hypothesis that these feedback coefficients are positive real numbers, the following inequality will be true.

$$\lambda_2\lambda_3 + \frac{\lambda_2^2\lambda_3^2}{2(\lambda_2 + \lambda_4)^2} > \lambda_2\lambda_3(1 - q) + \frac{\lambda_2^2\lambda_3^2(1 - q)}{2(\lambda_2 + \lambda_4)^2}.$$

Comparing (21) with the above inequality, we have

$$(1 - q)\left[\lambda_2\lambda_3 + \frac{\lambda_2^2\lambda_3^2}{2(\lambda_2 + \lambda_4)^2}\right] > \frac{(\lambda_2 + \lambda_3)^2}{2},$$

i.e.,

$$\frac{(\lambda_2 + \lambda_3)^2}{2} < \lambda_2\lambda_3 + \frac{\lambda_2^2\lambda_3^2}{2(\lambda_2 + \lambda_4)^2}. \tag{22}$$

After (22) is simplified the following inequality

Table 1. Comparison between Two Power Spectra

period (months)	∞	40.0	20.0	13.3	10.0	8.0	6.7	5.7	5.0	4.4
angular frequency	0	0.16	0.31	0.47	0.63	0.79	0.94	1.0	1.26	1.41
observed spectrum	0.294	0.378	0.205	0.042	0.027	0.020	0.018	0.019	0.019	0.010
simulated spectrum	0.170	0.378	0.154	0.063	0.033	0.020	0.014	0.010	0.008	0.006

period (months)	4.0	3.6	3.3	3.1	2.9	2.7	2.5	2.4	2.2	2.1	2.0
angular frequency	1.57	1.73	1.89	2.04	2.17	2.33	2.51	2.67	2.83	2.99	3.14
observed spectrum	0.008	0.012	0.012	0.008	0.005	0.005	0.007	0.008	0.006	0.007	0.004
simulated spectrum	0.005	0.004	0.003	0.003	0.003	0.002	0.002	0.002	0.002	0.001	0.001

may be obtained.

$$\lambda_2^2 < \lambda_3^2\left[\frac{\lambda_2^2}{(\lambda_2 + \lambda_4)^2} - 1\right]. \qquad (23)$$

However, if the hypothetical condition were true, the following inequality would hold.

$$\frac{\lambda_2^2}{(\lambda_2 + \lambda_4)^2} < 1.$$

Obviously it is not consistent with the hypothesis mentioned above, because by applying the above inequality to (23) we can obtain $\lambda_2^2 < 0$. In other words, the condition for a maximum in the SST response spectrum is that all feedback coefficients are not positive. If (22) is true, either λ_2 or λ_4 must be negative. On the other hand, λ_2 must be positive, because the attenuation coefficients of clouds, $\eta_3(T)$ and $\eta_4(T)$, are usually regarded to increase with temperature, and it is taken to be positive in most models (see Frankignoul and Hasselmann, 1977; Wright, 1979; Nicolis, 1979), owing to significant feedback by itself to the sea or air system. Therefore, λ_4 has to be taken negative.

In fact, as far as the expression of λ_4 is concerned, λ_4 may be negative when the attenuation action of the clouds on short-wave radiation is getting more significant with increasing temperature. The process may be the following one. The more the amount of clouds increases, the more the clouds weaken the short-wave radiation. The process lowers the SST. Wetherald and Manabe (1980) found that the surface radiation balance owing to increasing of clouds in middle latitudes and tropical belt regions is negative. Hunt (1981) pointed out that the negative feedback for cloud surface temperature is very obvious.

We compute the SST response spectrum, (18), by means of the estimation of the white noise spectrum according to its relationship with the diffusion coefficients, i.e.,

$$F(0) = \frac{D}{\pi} \qquad (24)$$

where D is $1.5(°C)^2$/yr and the coefficients are $\lambda_1 = 1/1.6$ rad mon^{-1}, $\lambda_2 = 1/5.8$ rad mon^{-1}, $\lambda_3 = 1/1.95$ rad mon^{-1}, and $\lambda_4 = -1/24.0$ rad mon^{-1}. The spectrum multiplied by 0.194 is on the fourth row in Table 1. Comparing the figures between the third and the fourth rows, we find that the observed and simulated results are very consistent.

We will discuss the simulation process as follows. As far as the figures of feedback coefficients are concerned, there is $|\lambda_3| > |\lambda_4|$. It seems to show that the feedback actions in the atmosphere and in the ocean are larger than their interaction with each other. From the case that $|\lambda_1|$ of these feedback coefficients is a little greater than the others, it shows that SST is playing an important role in quasibiennial oscillation. However, the interaction of air-sea in the process mentioned above can not be omitted. It can be found from comparing between Frankignoul and Hasselmann's model and the model in this paper. The feedback coefficient itself is 1/4.5 mon^{-1} in the former (corresponding to λ_3 (= 1/0.72 rad mon^{-1}) in this paper). It is obviously much larger than one in the latter, and yet the phenomenon of periodic oscillation can be simulated in the former. The conclusion is that the phenomenon of SST periodic oscillation depends on the interaction in the sea-air system in which the action of the feedback of SST is weaker than the interaction between the ocean and the atmosphere.

We carried out a factor analysis for middle latitudes in the Pacific. Considering the different features between east and west, we separately computed the power spectra of the monthly SST first factors in the east (160°W-110°W, 10°N-30°N) and west (120°E-170°W, 20°N-40°N). Fig. 1 shows the power spectrum curves; it displays roughly semi-annual oscillation (solid curves). We computed the simulated power spectrum by means of (18), taking D = $1.5(°C)^2$/yr, $\lambda_1 = 1/15.9$ rad mon^{-1}, $\lambda_2 = 1/1.4$ rad mon^{-1}, $\lambda_3 = 1/6.2$ rad mon^{-1}, and $\lambda_4 = -1/33.0$ rad mon^{-1}; it is shown by the dashed curve in Fig. 1 (Huang and Li, 1983).

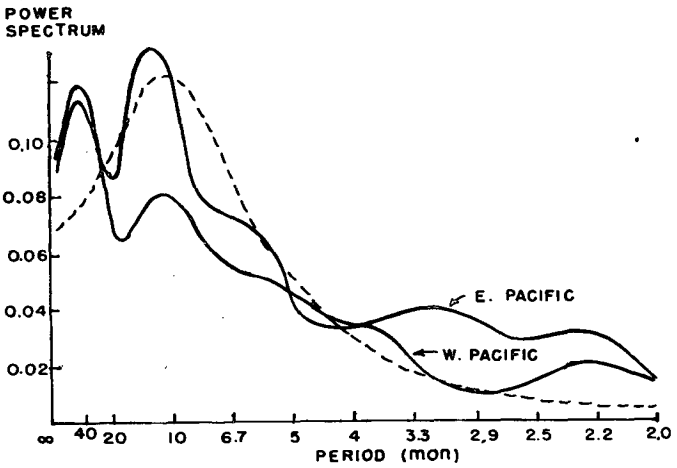


Fig. 1. Observed power spectra for the east and west Pacific SST in middle latitudes (solid lines) and general power spectrum with period of one-half of a year as simulated by (18) (dashed line).

Comparing the feedback coefficients of the simulated spectra of the quasibiennial oscillation and the semiannual one, we find that the feedback action of the atmosphere to the ocean in the process of the latter is weaker than in the former (from comparing the two λ_4). It seems to cause a little cloudiness in the middle latitudes. From comparing the two λ_3, we find that the action of feedback itself in the ocean strengthens and the action of the atmospehre to the ocean feedback weakens.

4. CONCLUSION AND DISCUSSION

In this paper, we suggest a second-order stochastic climate model which has explained the phenomena of SST periodic oscillation in the Pacific, which results from some random forces and the interaction of sea-air in the Pacific. Clouds are the premier factor in the processes. The negative feedback action of clouds plays an important role in quasibiennial oscillation of

SST in tropical regions. As concerns semi-annual oscillation of SST in middle latitudes, the actions of feedback to itself and feedback to the air-sea system are major.

As is well known, quasibiennial and semi-annual oscillation in the tropical belt and middle latitudes exist not only in SST but also in other meteorological elements. It is important to simulate and explain them. Our study makes a beginning in this respect. We expect that it will be improved in the near future.

5. REFERENCES

Frankignoul, C. and K. Hasselmann, 1977: Stochastic climate models, part II. Application to sea-surface temperature anomalies and thermocline variability. Tellus, 29, 289-305.

Huang, Jia-you and Mai-Tsun Li, 1983: Factor analysis of sea-surface temperature in the Pacific. To be published in Collected Papers of Atmospheric Science (in Chinese). Beijing, People's Republic of China, Scientific Press.

Hunt, B.G., 1981: An examination of some feedback mechanisms in the carbon dioxide climate problem. Tellus, 33, 78-88.

Li, Mai-Tsun, 1981: The stochastic-dynamic theory of air-sea interaction. Acta Oceanologica Sinica, 3, 382-389.

Nicolis, N., 1979: A simple air-sea interaction model. Quart. J. R. Met. Soc., 105, 93-105.

Oerlemans, J., 1979: A model of a stochastically driven ice sheet with planetary wave feedback. Tellus, 31, 469-477.

Wetherald, R.T. and S. Manabe, 1980: Cloud cover and climate sensitivity. J. Atmos. Sci., 37, 1485-1509.

Wright, P.B., 1979: A simple model for simulating regional short-term climate changes. Mon. Wea. Rev., 107, 1567-1580.

STATISTICAL CLIMATE PREDICTION : THE STATE OF THE ART

N. Nicholls

Australian Numerical Meteorology Research Centre
Melbourne, Australia.

1. INTRODUCTION

Statistics has played a major role in climate prediction (i.e. prediction of monthly, seasonal, annual or multi-annual average climate anomalies) since the start of scientific efforts in this field. The bulk of research into methods for climate prediction has involved searches for statistical techniques and the methods currently used in operational climate prediction are predominately statistical. Only recently have dynamical methods begun to play a role in research and their use in operational prediction is likely to be very limited for quite some time. Physical reasoning about the causes and likely development of climate fluctuations has also played a role in climate prediction, however our present, limited knowledge of the causes of climate fluctuations has meant that physical reasoning has generally played a supportive role to empirical, statistical analysis. The techniques used in statistical climate prediction include simple or multiple linear regression, analogue selection, and the use of apparent periodicities in the weather and climate.

Research into these methods for the prediction of climate fluctuations has resulted in a number of contributions to the theory and practice of statistics. For instance, Sir Gilbert Walker, a meteorologist who carried out an extensive search for empirical methods of use in climate prediction in the early decades of this century developed equations, now known as the Yule-Walker equations, to determine the order of an autoregressive process. Grant (1954), an Australian involved in climate prediction research during the early 1950's, recognised that biassed regression and correlation coefficients may be produced by sub-set selection in multiple linear regression. Several other studies in statistical climate prediction have also resulted in contributions to the field of statistics.

The effects of all this work, this long search for statistical methods for climate prediction, has not been so fruitful, however, in actually producing methods for successful climate prediction. In most countries where the meteorological service provides operational long-range forecasts the skill achieved is rather low. Baker (1982) has published an evaluation of the monthly forecasts issued by the British Meteorological Office from 1971 to 1980, for Central Southern England. Only 19.2% of temperature forecasts were correct, compared with an expected percentage of 20% from random forecasts. Baker described the results of his evaluation as "disappointing". Similar results are found in many verifications of long-range forecasts. An even more discouraging aspect is that these operational long-range forecasts do not seem to be improving (Nicholls, 1980; 1983).

This general lack of success has produced considerable sceptism, among meteorologists, about the feasibility of long-range prediction, at least by statistical techniques. Berlage(1957) noted that "From the definite optimism of the twenties students of the subject have been sliding down towards the pessimistic attitude of the fifties. We have arrived now again at the very bottom of the problem of seasonal forecasting, asking whether there is any hope for some recovery from this almost"deadly fall". Berlage was wrong. The "deadly fall" has since continued and doubt about the feasibility of climate prediction is wide-spread. Thus, W.J. Gibbs, an ex-Director of the Australian Bureau of Meteorology and an ex-Vice President of WMO noted recently that "despite intensive efforts ..... no significantly accurate method of making long-range forecasts has so far been discovered in Australia or elsewhere" (Gibbs, 1982). Thus the century or more research into statistical climate prediction has failed to convince many, perhaps most, meteorologists that statistical climate prediction is possible. This paper examines some possible explanations of this failure and attempts to suggest approaches to climate prediction research and practice which might overcome the general mood of sceptism. Throughout this paper "climate prediction" and "long-range forecasting" are used interchangeably. Both terms refer to the prediction, from a month to several years ahead, of time-averages of deviations (anomalies) of the weather from the usual. In practice these time-averages are taken over a month, season, a single year, or several years.

2. STABILITY AND INSTABILITY OF STATISTICAL CLIMATE PREDICTION TECHNIQUES

Possibly the aspect of recent work in climate prediction that has been most damaging has been the revelation that many of the statistical techniques proposed for climate prediction in the early decades of this century have failed to live up to expectations in subsequent validation on independent data. The field of sun-weather relationships is strewn with such cases where a

technique showing potential for climate prediction has been developed on limited data but has failed to produce useful forecasts when tested on later data.

Similar problems have been noted with other climate forecast techniques. One well-known case is the prediction of Indian monsoon rain. Walker, in a series of studies early this century, developed multiple regression techniques for predicting June to September rainfall over India. A typical formulae, for the prediction of rainfall over Peninsula (south-west) India is quoted by Khandekar (1979) as:

Rainfall = -0.067(Southern Rhodesia rain, October - April)
+0.913(South American pressure, April-May)
-0.0183(Java rain, October-February)
+0.912(Bangalore northerly winds,April)
-0.559(Calcutta easterly winds, May)
-4.307.

The choice of predictors in the above equation was made solely on the basis of the magnitude (and statistical significance) of the correlation coefficient between the monsoon rainfall and each of the predictors. Unfortunately, these correlations have not remained constant. Table 1 shows the decade correlations for three of the predictors with Penisula rainfall, up to the decade 1951-1960. The correlations for all the factors become statistically insignificant in some decade and even abrupt reversals of sign are displayed.

TABLE 1
DECADE CORRELATIONS BETWEEN VARIOUS PREDICTORS AND MONSOON RAINFALL OVER PENINSULA INDIA. (From Khandekar, 1979).

Decade	South Rhodesia Rain	South American Pressure	Java Rain
1881-90		0.48	-0.11
1891-00		0.78	-0.78
1901-10		0.42	-0.48
1911-20	-0.72	0.27	-0.19
1921-30	-0.24	0.02	-0.24
1931-40	-0.19	0.50	-0.07
1941-50	0.31	0.58	0.04
1951-60		-0.73	-0.11
Entire Period	-0.31	0.34	-0.21

Some of the variation in the decadal coefficients in Table 1 will be the result of sampling variations, even if a real, non-zero correlation exists between monsoon rainfall and the predictors. The magnitude of the changes (e.g. 0.58 to -0.73 for South American pressure) do, however, seem to be larger than might be expected from sampling variations, suggesting that either the "real" correlation is closer to zero than initially believed or climate changes on a timescale of decades occur which substantially alter the relationships between the variables. Whatever the case, the substantial variations shown in Table 1 do not provide one with much confidence in the use of the regression formula to predict monsoon rainfall.

Not all validations of promising statistical methods for climate prediction have produced such discouraging results. Nicholls (1981) verified on independent data the suggestion, first published over 70 years earlier, that Darwin pressure during the Southern Hemisphere winter could be used to predict rainfall during the period around the onset of the wet-season over Indonesia (the Southern Hemisphere spring). Correlations between Darwin June-August pressure and September-November rainfall totals for several stations in the Indonesian region are shown in Figure 1 for two periods: up to 1940, and 1941 to 1965. As expected, there is some variation in the magnitude of the correlations between the two periods. Nevertheless, the correlations in the two periods show considerable similarity with strong negative correlations occuring in the same areas.

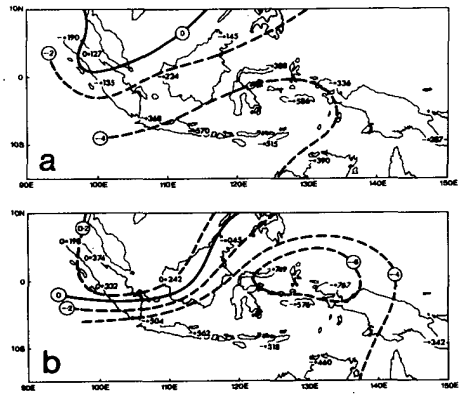


Figure 1a. Correlations between Darwin June - August pressure and September-November station rainfall. Data from 1883-1940.

Figure 1b. As in Figure 1a but with data from 1941-1965.

Further evidence of the stability of this relationship between Darwin pressure and subsequent Indonesian rainfall is provided in Figure 2 which is a scatter diagram of observed versus predicted September-November Jakarta rainfall for the period 1970-1980. The predictions were made using observed Darwin August pressure and a linear regression developed using data from 1951-1969. The correlation between the observed and predicted rainfall was 0.66 and the technique clearly discriminates between wet years (1973, 1975, 1978) and dry years (1972, 1976, 1977). Thus, unlike the case of the Indian rainfall, the lag relationship illustrated in Figures 1 and 2 has remained strong and stable over a considerable period.

Similar studies have demonstrated that rainfall, at this time of year, over much of eastern and northern Australia can also be predicted (e.g. Nicholls and Woodcock, 1981; Nicholls et al., 1982).

Another example where predictive relationships have remained relatively stable over a considerable period is provided by Yamamoto(1978) who demonstrated that December-February rainfall over the Philippines and southern Japan is significantly correlated with the preceding September-November pressure over tropical Asia and Australia. For the Japanese rainfalls, correlations were calculated for two periods: 1891-1940 and 1951-1970. The correlations for the two periods were of the order of 0.5, with only small reductions in magnitude in the second period.

A final example of apparently stable statistical climate prediction relationships is given by Dennett et. al. (1978) who showed, initially only on data from 1938-1972, that an index of the Southern Oscillation in August-October was strongly related to subsequent November-April rainfall in Fiji. The time series of data was then extended back to 1910 with no apparent detrimental effect on the relationship.

The stability of the predictive relationship in these cases strongly supports the proposition that long-range prediction in at least some areas and seasons is feasible, using simple statistical techniques. Why then have operational long-range forecasts failed, in general, to produce predictions showing a useful level of accuracy, and what is the cause of the instability of promising predictive relationships such as that for the Indian monsoon? The remainder of this paper discusses possible answers to these questions.

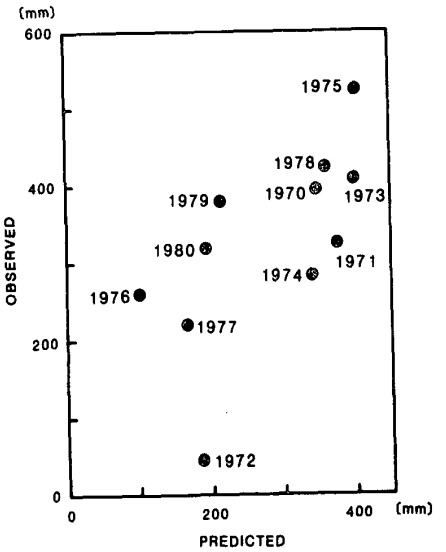


Figure 2. Predicted and observed September - November Jakarta rainfall. Predictions made from observed Darwin August pressure using linear regression derived on 1951-1969 data.

3. POSSIBLE CAUSES FOR THE FAILURE OF OPERATIONAL CLIMATE PREDICTIONS.

3.1 Climate predictability may be restricted to certain areas and seasons.

There is evidence from a number of sources which suggests that accurate statistical climate prediction may only be feasible in certain areas and seasons. The most convincing evidence of this is provided by studies of "climate noise", a topic which Madden addresses elsewhere in these proceedings. However other evidence is available which supports the results of these climate noise studies (Nicholls,1983). For instance, operational climate forecasts show seasonal and spatial variations in accuracy, possibly because of real variations in predictability. Also, studies with numerical models of the atmosphere have demonstrated that these models can more adequately represent the inter-annual variability of the mid-latitude atmosphere than of the tropics (Charney and Shukla, 1981), suggesting that some signal exists in the tropical atmosphere other than the synoptic weather fluctuations explicitly included in the model. This extra signal is presumably the result of ocean-atmosphere interaction and may be predictable at long lead times. The overall impression one gains from these various studies is that the prospects for tropical climate prediction appear better than for mid-latitudes, and the prospects for prediction over the oceans and coasts appear better than for the continental areas. If so, this factor alone could explain the poor performance of many operational climate forecasts, since these have generally concentrated on predicting climate for the mid-latitude continents.

3.2 The pursuit of lost causes

One popular approach to statistical climate prediction is to search for periods or "quasi-periodicities" in meteorological data. Especially popular has been the idea that the 11- or 22- year sunspot cycles might be related to climate variations and thus the predictability of the sunspots would therefore guarantee that climate variations could also be predicted. This search for atmospheric periodicities and sunspot-weather relationships has absorbed a considerable proportion of resources available in the field of statistical climate prediction research. Thus Pittock (1978) discusses over 150 papers on the subject of sun-weather relationships, and in an update (Pittock, 1983), cites another 100 papers published on the subject subsequently. In a number of countries, sunspots or apparent atmospheric periodicities are used in the preparation of climate forecasts.

This is the situation despite numerous studies which indicate the futility of this approach. Pittock (1978), after his extensive review, concluded that "... despite a massive literature on the subject, there is at present little or no convincing evidence of statistically significant or practically useful correlations between

sunspot cycles and the weather or climate", and further that "the evidence to hand suggests that if in the future more data and better analysis enable the detection of statistically significant relationships, these will account for so little of the total variance in the climatic record as to be of little practical value".

In respect of searches for atmospheric periodicities in general, the USA National Academy of Sciences (1975) concluded that "... it has become clear that almost all the alleged climatic cycles are either (1) artifacts of statistical sampling, (2) associated with such small fractions of the total variances that they are virtually useless for prediction purposes, or (3) a combination of both.

3.3 Assumption of stationarity

Most methods of linear prediction used in statistical climate prediction assume the statistics of the data fields to be time invariant. Many parameters that measure climatic variability, however, have nonstationary statistics, ie. they depend on the phase of the annual cycle. In such cases the usual statistical techniques based on time-invariant models are inappropriate and may lead to erroneous conclusions (Hasselman and Barnett, 1981).

An example of the misleading conclusions that may be reached from an inappropriate assumption of stationarity over the annual cycle is provided by two studies of the predictability of monthly anomalies of sea level pressure over the North Pacific, from monthly anomalies of North Pacific sea-surface-temperature (Davis; 1976, 1978). In the first study carried out by Davis, time invariance of the statistics was assumed and the study concluded that the pressure anomalies were not predictable from prior sea-surface-temperatures. The second study (Davis, 1978) repeated the analysis after seasonal stratification of the data. This time it was concluded that autumn and winter pressure anomalies, but not spring or summer anomalies, could be predicted.

Another example of a strong seasonal cycle in predictability is provided in Figure 3 which shows lag correlations between Darwin mean seasonal air temperature and an index of the Southern Oscillation. Very strong variations in the lag relations are evident between seasons. Thus December-February and March-May Darwin temperatures are strongly, negatively, correlated with the Southern Oscillation up to two seasons earlier. Contrast this with the small, non-significant, correlations of September-November temperature with the Southern Oscillation. If time invariance had been assumed, and all seasons used together in a single analysis, the strong lag correlations for December-February and March-May would not have been observed.

One problem with introducing seasonally varying statistical analysis is that the amount of data available is less than would be the case in an analysis assuming time invariance. This may make it more difficult to construct statistically significant prediction models. A possible way of overcoming this might be to use models where the coefficients are expanded in a harmonic series and statistical averages estimated by continuous time averages over the entire series (Hasselman and Barnett, 1981).

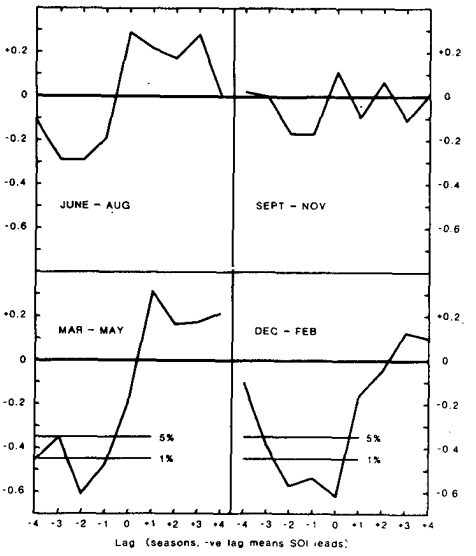


Figure 3. Lag correlations between seasonal averages of Darwin air temperature and the Southern Oscillation. The season of the Darwin temperature is indicated. Data from 1950-1979.

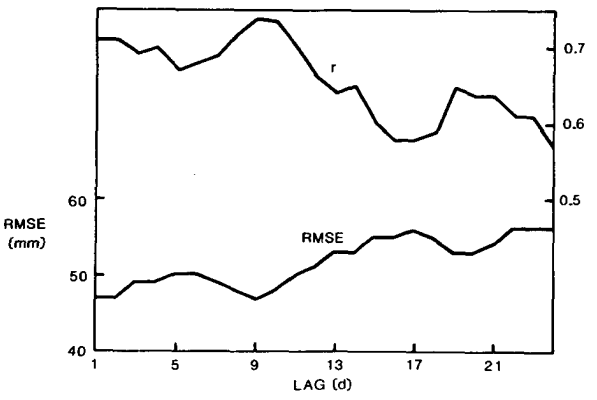


Figure 4. Root-mean-square errors (RMSE) of predictions of Darwin September-November rainfall, and correlations (r) between observed and predicted rainfall. Predictions made using linear regression, with Darwin June-August pressure as predictor, developed using 15 years of data ending d years prior to the year for which a prediction is made. For each value of d, predictions were made for each year from 1951-1980.

It may also be dangerous to assume stationarity of climate statistics, and thus of statistical climate prediction models, over several decades. Figure 4 shows root-mean-square errors of predictions of September-November rainfall in the Darwin area and correlations between the observed and predicted rainfall. The predictions were made using linear regression with Darwin June-August pressure as the only predictor (Figure 1 indicated that rainfall at Darwin, at this time of year is predictable in this way). Predictions were made for September-November in each year from 1951-1980, using 15 years of data to develop the regression equation. The 15 years used to develop the equation were varied from the 15 years ending in the year prior to the year for which a prediction was made up to the 15 years ending 25 years before the prediction was required. All the statistics in Figure 3 were calculated for predictions for each of the 30 years from 1951-1980. Thus the figure indicates how the accuracy of the predictions changed as the delay between the period used to develop the prediction equation and the year in which the equation was used to predict rainfall, increased. As the delay increases the predictions progressively deteriorate, suggesting that the statistics are not time invariant over the several decades used. The implications of this are that predictive equations may need regular updating and that simply including more data, as it becomes available, may not improve the predictions.

These examples indicate that the assumption of stationarity may lead to erroneous conclusions about predictability of climate. In general, one would expect that the stationarity assumption would lead to overly pessimistic conclusions (cf. the studies of Davis). It would seem reasonable to assume non-stationarity, at least over the seasonal cycle, rather than the reverse.

3.4 Bias in regression

Many statistical climate prediction schemes include the use of linear regression equations derived on relatively small data sets, often with the predictors selected from a large number of candidate predictors. The equations quoted above for the prediction of rainfall in India are typical examples. Unfortunately, when such regression equations are used in prediction, the fit to the new data is nearly always worse than was the fit to the original data and the regression coefficients tend to be of too large a magnitude. This situation is the result of "selection bias". A simple example, due to A.J. Miller of the CSIRO Division of Mathematics and Statistics in Australia, of the effect of selection bias in linear regression, is given below.

Consider a situation in which only two candidate predictor variables, X_1 and X_2 are available and it has been decided a priori to select only one of them, that one being the one which gives the smaller residual sum of squares when fitted to a set of data. Let us define -

$$X_1 = Z_1$$
$$X_2 = -3Z_1 + 4Z_2$$
$$Y = Z_1 + 2Z_2 + Z_3$$

where Z_1, Z_2 and Z_3 are independently sampled from a normal distribution and Y is the predictand. If we take a sample of n independent observations and fit the two alternative models,

$$Y = aX_1 + \text{residual}$$
$$Y = bX_2 + \text{residual} \quad ,$$

by least squares, then the expected values of the sample regression coefficients a and b are 1.0 and 0.2 respectively and the expected residual sum of squares is 5(n-1) for both models.

Using samples of size n=12, 200 sets of artificial data were generated, and the average value of "a" calculated for those cases when X_1 was selected as the predictor, and the average value of "b" calculated for those cases when X_2 was selected. The average values of a and b were 1.288 and 0.285, considerably larger than the expected values. The average residual sum of squares for the selected models was 45.7, considerably smaller than the expected value of 55. Thus the selection procedure has introduced a bias. The regression coefficients, under selection, are likely to be greater in magnitude than they should be, and the fit to the data will tend to suggest stronger relationships than really exist.

Selection bias can be completely eliminated by using separate, independent data sets for the selection of the model and for estimating the regression coefficients, however it is rarely sensible to use this method as it is inefficient in not using information from the data set used to select the model in estimating the coefficients (A.J. Miller, personal communication). This inefficiency is particularly important in statistical climate prediction because of the limited amounts of data available. One alternative approach to reduce the bias is through the use of some kind of shrinkage of the coefficients, eg ridge regression or James-Stein shrinkage.

A related, but better known, source of bias in prediction from regression equations is the "artificial skill" that tends to arise in multiple linear regression (Davis, 1976). In a finite number of realisations of a predictand and set of candidate predictors, there will always appear to be a correlation between the predictand and at least some of the "predictors" even if no true correlation exists. As more predictors are added to the multiple linear regression equation, the apparent total correlation between predictand and "predictor" set will always increase regardless of whether the added variables are really related to the predictand. When used in prediction on independent

data, equations developed without due consideration of this problem can actually produce predictions showing negative skill, i.e. worse than climatology. Again, some sort of shrinkage can alleviate this problem to a degree. Restricting the number of predictors can also minimize this bias.

These two sources of bias in linear regression can result in serious problems if not accounted for before the regression equations are used in prediction. Ignorance of these problems has undoubtedly contributed to the deterioration observed in regression equations when tested on dependent data. For example, Grant (1954) demonstrated that selection bias accounted for much of the deterioration of Walker's equations for predicting Indian rainfall described above. Unfortunately, it is clear that many researchers and practitioners of statistical climate prediction remain unaware of these problems and continue to produce regression equations with many predictors, selected from large numbers of competing variables, without regard for the bias such an approach will produce. The effects of these sources of bias in producing apparent deterioration of prediction formulae performance has undoubtedly contributed to the widespread scepticism of climate prediction amongst meteorologists.

3.5 Data problems

Possibly the most serious restriction to statistical climate prediction is the data problem. Over most of the globe not much more than about 100 years data is available; in some key areas even less. Even this limited data-set is plagued by missing observations, changes in observing practices, changes in station location and possibly climate changes, natural or man-induced. For instance, historical records of ship-observed air-sea temperature differences appear to show substantial secular variations over a period of decades (P.B. Wright, personal communication). Whether the air temperatures are wrong, or the sea-surface temperatures, or both, we are unable to say. This is a problem of considerable importance since sea-surface-temperatures are prominent in many proposed climate prediction techniques.

As a result of these many data inadequacies even the most significant short-term climate fluctuation, the Southern Oscillation, is still inadequately observed and described. Our knowledge of the possibilities of statistical climate prediction will remain limited until far more accurate, comprehensive and compatible time series of data become available. Improving future data, and careful examination of past data to overcome some of these deficiencies must remain a high priority if we are to advance very far in statistical climate prediction. Specifically, until we can improve our data base it will remain extremely difficult to determine how much of the apparent long-term variations in prediction equations (e.g. Table 1) is due to changes in observational practices, how much to inadequate statistical techniques (e.g. not taking account of the sources of bias discussed above), how much to sampling variations, and how much to real climate variations. Until we can answer this question we may be unable to convince many atmospheric scientists of the value of statistical climate prediction.

4. CONCLUDING REMARKS

Statistical climate prediction is one of the most challenging problems in applied statistics. The field is extremely multi-variate with literally hundreds of thousands of candidate predictors and usually no more than about 100 realisations of predictand and predictors. The available data are plagued by problems, such as changes in observing practices or secular climate changes. Stationarity cannot be assumed a priori, nor can even the feasibility of climate prediction. In these circumstances it is not surprising that statistical climate prediction, despite considerable efforts, has not yet been demonstrated to be feasible, at least to the satisfaction of many meteorologists. Yet in at least some areas and seasons there is strong evidence of its feasibility based on successful verification, on independent data, of prediction schemes proposed early this century. Some of these prediction techniques have now been verified on more than 50 years of independent data. This suggests that the important question to be answered is not whether climate prediction is feasible but why we have so far found strong evidence of predictability in only a few areas. Some possible explanations have been discussed above; other contributory factors can also be advanced. If the field of statistical climate prediction is to be successfully advanced, it would appear that general recognition of some of the potential problems is necessary. These include:

The probability that successful climate prediction may be spatially and seasonally limited,

That certain popular approaches (e.g. sunspot-weather relationships) have been demonstrated to be of little value, and that continued preoccupation with these approaches only causes scepticism about the entire climate prediction field,

That the assumption of stationarity, over the annual cycle or over a period of decades is probably unwarranted,

That bias can be expected in statistical climate prediction techniques, as they are at present derived, resulting in substantial degradation of the performance of these techniques when used in prediction on new data, and that

Considerable efforts are required to improve and extend the present data base.

Other efforts are also necessary if advances are to be made in this field. One example is the verification of operational climate forecasts. At present different organisations use different verification strategies, often producing conflicting results. The setting of some standards in

verification of climate prediction would appear to be justified. Increased interaction between those involved in statistical climate prediction, theoreticians, and numerical modellers of the atmosphere and ocean also appears to be indicated at this stage in the development of the subject. The analysis and understanding of the Southern Oscillation has, for instance, benefitted from such joint efforts in recent years. Extension of such collaborative work can only improve our knowledge of the predictability of climate fluctuations.

Finally, is it worth continuing the search for statistical climate prediction methods in view of the rather limited success so far achieved? The serious effects of climate fluctuations on society suggest that it is. For instance the 1982 Australian drought reduced average Australian farm income to about 20% of the previous year. Even more devastating effects occur in the developing countries, e.g. the reported death by starvation of 350 persons in drought affected Irian Jaya in 1982. In the face of such effects of climate fluctuations, continued efforts to develop widely-applicable statistical climate prediction techniques would appear to be justified.

5. REFERENCES

Baker, T., 1982: An evaluation of long-range forecasts. Weather, 37, 262-265.

Berlage, H.P., 1957: Fluctuations of the general atmospheric circulation of more than one year, their nature and prognostic value. Medelingen en Verhandelingen, 69, Koninklijk Nederlands Meteorologisch Institut.

Charney, J.G. and J. Shukla, 1981: Predictability of monsoons. Monsoon Dynamics, eds. J Lighthill and R. Pearce, Cambridge Univ. Press, 736pp.

Davis, R.E., 1976: Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. J.Phys.Ocean., 6, 249-266.

Davis, R.E., 1978: Predictability of sea level pressure anomalies over the North Pacific Ocean. J. Phys. Ocean., 8, 233-246.

Dennett, M.D., J. Elston and P.C. Prasad, 1978: Seasonal rainfall forecasting in Fiji and the Southern Oscillation. Agric. Meteorol., 19, 11-22.

Gibbs, W.J., 1982: A perspective of Australian meteorology. Aust. Meteorol. Mag., 30,3-17

Grant, A.M., 1954: The application of correlation and regression to forecasting. Meteorological Study 7, Bureau of Meteorology, Melbourne, Australia.

Hasselman, K. and T.P. Barnett, 1981: Techniques of linear prediction for systems with periodic statistics. J. Atmos. Sci., 38, 2275-2283.

Khandekar, M.L., 1979: Climatic teleconnections from the equatorial Pacific to the Indian Monsoon - Analysis and implications. Arch. Met. Geoph. Biokl. Ser.A, 28, 159-168.

Nicholls, N., 1980: Long range weather forecasting: value, status and prospects. Rev. Geophys. Space Phys., 18, 771-788.

Nicholls, N., 1981: Air-sea interaction and the possibility of long-range weather prediction in the Indonesian Archipelago. Mon. Weath. Rev., 109, 2435-2443.

Nicholls, N., 1983: Prospects for empirical long-range weather prediction. Proc. WMO-CAS/JSC Expert Study Meeting on Long-Range Forecasting, Princeton,December 1982.

Nicholls, N. and F. Woodcock, 1981: Verification of an empirical long-range weather forecasting technique, Quart. J. Roy. Meteorol. Soc., 107, 973-975.

Nicholls, N., J.L. McBride and R.J. Ormerod,1982: On predicting the onset of the Australian wet season at Darwin. Mon. Weath. Rev., 110, 14-17.

Pittock, A.B., 1978: A critical look at long-term sun-weather relationships. Rev. Geophys. Space Phys., 16, 400-420.

Pittock, A.B., 1983: Solar variability, weather and climate: an update. Quart. J. Roy. Meteorol. Soc., 109, 23-55

USA Academy of Sciences, 1975: Understanding Climate Change, Washington.

Yamamoto, T., 1978: Some remarks on the relationship between the climate of Japan and that of Tropical Asia, Australia and East Africa. Climate Change and Food Production, eds. K. Takahashi and M.M. Yoshino, Univ.Tokyo Press, 433pp.

BAYESIAN PREDICTION APPLIED TO SIMPLE LINEAR REGRESSION: MONTHLY MEAN NORTH AMERICAN SNOW COVER

Edward S. Epstein

National Environmental Satellite, Data, and Information Service
National Oceanic and Atmospheric Administration
Washington, D.C. 20233

1. INTRODUCTION

Since the winter of 1967/68, imagery from NOAA polar orbiting satellites has been analyzed to estimate the areal extent of continental snowcover in the Northern Hemisphere. Analyses are made weekly and these have been summarized into monthly and seasonal means for North America (plus Greenland) and Eurasia (Dewey and Heim, 1982). This information has been of considerable interest to climatologists, especially those interested in short term climate prediction (i.e. monthly, seasonal) or long term weather prediction (one to several weeks). The prediction of snow cover, from one month to the next, would contribute importantly to the general forecast process. It is natural to seek out a statistical prediction scheme.

Most climatologists would expect, for several reasons, that there is a positive autocorrelation between snowcover from one month to the next. Reasons cited might be thermal and latent heat inertia, albedo feedback, and month-to-month autocorrelation of the large-scale atmospheric flow patterns. Thus a prediction model of the form $y=\beta_1+\beta_2 x+\epsilon$, where x is the areal extent of snowcover in the predictor month and y is the snow cover in the immediately following predictand month, would appear to be reasonable. β_1 and β_2 are unknown coefficients and ϵ is a random variable with zero mean, presumed to be normal, and having a variance σ^2.

Normally one estimates the unknown coefficients from a lengthy record of observations. In the present case there is no lengthly set of observations. The climatologist's knowledge may even now, after accumulating data for 16 years, be worth considering. In other words, the climatologist would not rely entirely on a short record, but would incorporate experience and judgment in his or her own prediction. In the early years of satellite observations the climatologist's judgment was necessarily the principle basis for making monthly predictions of snowcover. I will describe how, in a situation involving simple linear regression, the expert's prior knowledge can be usefully codified, and, with the help of Bayesian methods, predictions can be made that systematically combine the prior judgment and the available data. A comparison of a set of predictions made with and without prior knowledge will demonstrate the validity of the approach.

2. BAYES' THEOREM AND SIMPLE LINEAR REGRESSION

It is first necessary to introduce some mathematical formalism. I have chosen to codify prior beliefs concerning the parameters of the process (β_1,β_2 and σ^2) in terms of a probability density that is conjugate to the normal simple linear regression process by which the data is presumably generated. This means that prior and posterior beliefs take the same functional form. Specifically they are expressed as a joint bivariate normal-gamma distribution (Raiffa and Schlaifer, 1961). This implies that the marginal distribution of the precision ($h=1/\sigma^2$) is gamma with two adjustable parameters, ν' and s'^2. (The product $h\nu's'^2=\nu's'^2/\sigma^2$ has a chi-squared distribution with ν' degrees of freedom.) The conditional distribution of $\underline{\beta}=(\beta_1\ \beta_2)^T$, when h is known is bivariate normal with expectation $\underline{b}=(b_1\ b_2)^T$ and covariance matrix $h\underline{n}'$. $\underline{n}'$ is a 2x2 positive definite symmetric matrix. In all, there are seven scalar parameters that must be quantified to define the prior or posterior distribution.

Given the prior parameters and one or more new observations $\underline{y}=(y_1 y_2 \ldots y_n)^T$ and

$$\underline{x}^T=\begin{pmatrix}1 & 1 & \ldots & 1\\ x_1 & x_2 & \ldots & x_n\end{pmatrix}$$

then the posterior probability density on $\underline{\beta}$ and h is again bivariate normal-gamma, but with revised parameters

$$\underline{n}=\underline{n}'+\underline{n}=\underline{n}'+\underline{x}^T\underline{x}$$

$$\underline{b}'' = (\underline{n}' + \underline{n})^{-1}(\underline{n}'\underline{b}' + \underline{x}^T\underline{y}) = (\underline{n}' + \underline{n})^{-1}(\underline{n}'\underline{b}' + \underline{n}\,\underline{b})$$

$$\nu'' = \nu' + n$$

$$\nu''s''^2 = \nu's'^2 + \underline{b}'^T\underline{n}'\underline{b}' + \underline{y}^T\underline{y} - \underline{b}''^T\underline{n}''\underline{b}''$$

$$= \nu's'^2 + \underline{b}'^T\underline{n}'\underline{b}' + \underline{b}^T\underline{n}\underline{b} - \underline{b}''^T\underline{n}''\underline{b}'' + (n-2)s^2$$

The matrix $\underline{n}=\underline{x}^T\underline{x}$ is to be distinguished from the scalar n which is the number of observations (and also one of the elements of the matrix $\underline{n}$.) Note that $\underline{b}$ and s^2, as defined above, are the least squares estimates of the parameters determined from the data alone. They are also the parameters one obtains if the prior is assumed to contain negligible information relative to the data (in which case $\nu''=n-2$).

Given the posterior parameters $\underline{b}''$, $\underline{n}''$, ν'' and s''^2, and a new observation of the predictor, x, then the expected value of the predictand is $E(y|x)=b''_1+b''_2x$, and the variance of the prediction is

$$\mathrm{Var}(y|x) = \frac{\nu''}{\nu''-2}\,s''^2\left[1+\frac{1}{n''_{11}}+\frac{(x-n''_{12}/n''_{11})^2}{n''_{22}-n''^2_{12}/n''_{11}}\right]$$

where n''_{11}, n''_{12}, and n''_{22} are the three distinct elements of the matrix $\underline{n}''$ ($n''_{12} = n''_{21}$).

3. ESTIMATING PRIOR PARAMETERS

In order to define the prior density, it is necessary to specify seven scalar quantities: b'_1, b'_2, s'^2, ν', n'_{11}, n'_{12}, and n'_{22}. I will explain how I have done this specifically when x and y refer, respectively, to the December and January areal extent of snowcover over North America. We proceed as though we were in complete ignorance of the dta accumulated since 1966. I do not contend that there are unique correct values for these parameters; I expect that different people, indeed different experts, will make different judgments. What I do contend, and hope to illustrate, is that the codification of reasonable beliefs, and the incorporation of those beliefs in the prediction, will lead to better, more credible predictions than one can arrive at using wholly objective methods.

First I estimate the mean snowcover we might expect in mid-winter. A general knowledge of the geography of North America and Greenland might suggest that the total area that might be covered by snow is perhaps 18 to 20 X 10^6 km^2. (Hereafter we will express snowcover amount in units of 10^6 km^2.) I assign 17 as the mean extent of snowcover expected in January; further I assert that my prior mean for December snowcover is slightly smaller, or 16. I need also to express a judgment as to the year-to-year variability of these amounts. I believe that the variability will be about the same in January as in February. I assert the prior view that var(x)=var(y)=1.0. (We will later have the opportunity to express the degree of certainty in these values.)

I have already indicated a belief that the January snowcover will be positively correlated with the December snowcover. I quantify this by stating the belief that (a) a departure from its mean of the December snowcover leads me to expect a departure from the mean of the January snowcover 1/2 as large in the same direction, and (b) about one quarter of the variance of the January snowcover is "explained" by variations in the December snowcover. These beliefs mean that appropriate prior parameters are $b'_1=9$, $b'_2=0.5$, and $s'^2=0.75$ (since $s'^2 = E[\mathrm{var}(\epsilon)] = (1-r^2)\,E[\mathrm{var}(y)]$).

The remaining prior parameters quantify the uncertainty in the values of the coefficients we have just stated. For example, ν' together with s'^2, determines the variance of the marginal distribution of σ^2. Specifically, $\mathrm{var}[\sigma^2] = 2s'^2/\nu'$. One could estimate the variance of one's prior distribution on σ or σ^2, and use that to fix the value of ν'. I have adopted a somewhat different appoach, one that uses the fact that the prior beliefs, as stated, imply that $\nu's'^2/\sigma^2$ has a chi-squared marginal distribution. I calculate, for various values of ν', 80% and 90% credible intervals, and seek to determine, by inspection of those intervals, the value of ν' that gives intervals that correspond best with my beliefs. Some pertinent results of these calculations are shown in Table 1. On the basis of these calculations, and the belief that σ is almost certainly not as large as 2.0 (which means $\nu'>4$ according to the entries in Table 1) but that a value for σ of 1.5 is credible but unlikely, I have selected $\nu'=5$.

Table 1

Credible Intervals on σ for Various Values of ν'

ν'	.05	.10	.90	.95
3	0.54	0.60	1.97	2.53
4	0.56	0.62	1.68	2.06
5	0.58	0.64	1.63	1.81
6	0.60	0.65	1.43	1.66
7	0.61	0.66	1.36	1.56

The remaining prior parameters are n'_{11}, n'_{12}, and n'_{22}. They define, in mathematical terms, the uncertainty in my knowledge of β_1 and β_2. My approach to determining the values of these parameters that best express my views is to start with a first guess, examine the uncertainty in β_1 and β_2 implied by that guess, and revise the parameters until I am satisfied the implications are consistent with my beliefs. To start the process we note that, in fixing the value of the posterior expected values of the regression coefficients, $\underline{n}'$ plays the role of a weight given to the prior parameters, while the weight given to the data is $\underline{x}^T\underline{x}$. In effect choosing the matrix $\underline{n}'$ asserts that one's confidence in the regression coefficients is equivalent to having observed n'_{11} hypothetical observations, that the sum of these hypothetical observations is n'_{12} and the sum of their squares is n'_{22}. For our first guess at $\underline{n}'$, then, let us hypothesize that my confidence in my prior beliefs is equivalent to having three observations, one at x=15, one at x=16, and one at x=17. Then $n'_{11}=3$, $n'_{12}=48$, and $n'_{22}=770$.

It follows from the codification of our beliefs as a bivariate normal-gamma distribution that the joint marginal distribution of β_1 and β_2 is a bivariate Student's distribution with ν' degrees of freedom, and with a covariance matrix proportional to $\underline{n}'^{-1}$. One can calculate the implied marginal credible intervals for each. With the values given for n', the marginal 90% intervals on β_1 and β_2 are 9.0 ± 18.8 and 0.5 ± 1.2, respectively.

These intervals are too wide, by at least a factor of two, to express my beliefs. I do not, for example, believe that there is even a 5% chance that the slope is a negative as -.5. I might accept a 5% chance of $\beta_2 < 0.0$, but not a 5% chance of $\beta_2 < -0.7$. Therefore I will seek another matrix $\underline{n}'$ that will reduce the implied credible intervals by a factor or two or more.

Since the covariance matrix for the coefficients is proportional to the inverse of $\underline{n}'$, multiplying $\underline{n}'$ by a scalar factor, say c, will have the effect of dividing the variance by c, or dividing the standard deviation by $\sqrt{c}$. Thus if we want to make the credible intervals on β_1 and β_2 half as large all that is needed is to multiply $\underline{n}'$ by 4.

However, this is not the only way to decrease the credible intervals. One can also decrease the width of the credible intervals on β_1 and β_2 by increasing n'_{22} while holding the other elements of $\underline{n}'$ fixed. (This has the effect of spreading further along the x-axis the hypothetical observations on the basis of which the prior judgments are based, and thus determining the slope--but not the mean--more accurately. Knowing the slope better, and knowing the mean only slightly less well, implies in this case that the intercept, too, is better determined.) The next step is to determine which of many possible values of $\underline{n}'$ best corresponds to my belief. The most satisfactory means for doing this, I have discovered, is to examine the consequences of hypothetical future data.

I suppose that two hypothetical observations become availble to me after the prior has been determined, and examine the extent to which the two observations cause me to change my beliefs. Figure 1 presents the prior regression line ($y=9+.5x$) and five additional regression lines that represent posterior beliefs after observations at (15,15) and (15,19), corresponding to the following five matrices $\underline{n}'$, keeping all other prior parameters as already determined.

(a) $\underline{n}' = \begin{pmatrix} 3 & 48 \\ 48 & 770 \end{pmatrix}$

(b) $\underline{n}' = \begin{pmatrix} 3 & 48 \\ 48 & 776 \end{pmatrix}$

(c) $\underline{n}' = \begin{pmatrix} 3 & 48 \\ 48 & 800 \end{pmatrix}$

(d) $\underline{n}' = \begin{pmatrix} 12 & 192 \\ 192 & 3080 \end{pmatrix}$

(e) $\underline{n}' = \begin{pmatrix} 48 & 768 \\ 768 & 12320 \end{pmatrix}$

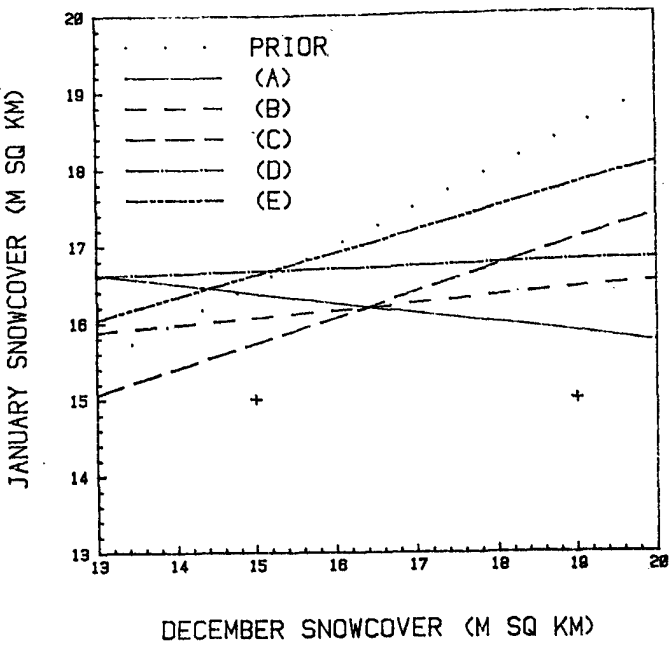


Figure 1. The prior regression line and five posterior regression lines (depending on the choice of the matrix $\underline{n}'$) after two hypothetical observations.

The two hypothetical observations were chosen to be well off the prior regression line so that they would indeed suggest a change of belief, but not so far off that they would appear incredible. It is difficult to evaluate evidence, real or imagined, that is not perceived to be credible.

With the two observations as hypothesized (and shown in Figure 1) I would, I believe, adopt a revised set of regression coefficients that would imply a smaller value of y corresponding to values of x in the range 16±3, and a reduced slope. Line (A), which corresponds to the prior matrix that has already been rejected, expectedly is an unacceptable posterior regression line. The negative slope results from trying to fit the hypothetical prior observations, which have a centroid at (16,17) and the hypothetical future data, whose center of mass is at (17,15), with the prior condition given that a negative slope is not unlikely.

At the other extreme, line (E) implies much more confidence in the prior regression line than I can accept as appropriate. The slope of line (D) is, to me, a credible compromise mixing prior views and subsequent observations, but places too much weight on the prior mean value assigned to y and too little on the two observed values of 15. On the other hand, prior (C) retains a slope that is too much like the prior slope. I choose matrix (B) as the best statement of my prior beliefs; line (B) represents an adjustment of views that I would indeed make if those observations did occur.

4. PREDICTIONS

Having determined the prior prameters that I would have adopted before the 1966-67 observations, I can proceed to make predictions of January snowcover, year by year, after observing the December snowcover. Also, each year, after observing the January snowcover, I can update the various parameters. No two predictions are made by the same equation; each one contains all the information--prior and empirical--that is currently available. Table 2 contains the results of the preeictions over the 16 years for which data are now available. Also shown in Table 2 are the predictions I would have made, following the same general methodology, if I had assumed that my prior information, in 1966, was vague, i.e. of so little value that its information content was negligible compared to even a single observation. In that case the regression equations are given by a least squares fit to the observations. (Of course initially there are no data to be fit, and thus no prediction can be made.)

Quite clearly, we can note from Table 2, the presence of the prior information led to less variation from year to year in the value of the regression coefficients. The slope is found to be negative over a substantial number of years in the absence of a prior assertion that a negative slope, while possible, is unlikely. It is also worth noticing the relationship of the prediction errors and the estimated standard errors of prediction. There are not substantially different errors of prediction between the two methods, but the ratios of error to standard error are much more nearly what one should expect when the prior information is of some substance; with the vague prior some unreasonable large values occur.

We can gain some insight into the evolution of the best judgments about the regression relationship from Figures 2 and 3. The points plotted in each of the figures are the observations accumulated through eight years (Fig. 2) and through 16 years (Fig. 3). In addition, there are three lines plotted in each figure: the prior regression line, the least squares regression line for the plotted points (which is also the regression line one should believe if the prior information were vague), and the regression line one is led to believe if the prior information is codified as described above.

The correlation between the first eight years' December and January snowcover amounts (Fig. 2) is clearly very weak; it suggests that the prior view that the coefficient of correlation was actually near 0.5 may have been an overestimate. One can also discern that the observations do not strongly suggest that the slope of the regression line must be negative. This allows the current best judgment, using the informative prior, to take the form of a regression line with positive slope, even though the prior information was, in effect, given a weight of only three observations, and the data consist of eight observations.

After 16 years (Fig. 3) the correlation between x and y is still weak (but less so) and the least squares regression line has a positive slope. There is less difference between the posterior judgments based on the informative and vague priors. As the amount of empirical data increases, the beliefs derived from a vague prior must converge with those derived from a reasonably stated informative prior. In other words, when there are many observations, the expert will accept the conclusion implied by the data; when the

			Informative prior (b)					Vague prior				
Year	x	y	y_c	σ_y	y_c-y	b_1	b_2	y_c	σ_y	y_c-y	b_1	b_2
66/67	15.1	17.0	16.55	1.34	-0.45	9.67	+0.465	-----	----	-----	-----	------
67/68	15.1	17.3	16.69	1.12	-0.61	10.37	+0.428	-----	----	-----	-----	------
68/69	16.4	17.4	17.39	1.03	-0.01	10.36	+0.429	-----	----	-----	14.25	+0.192
69/70	14.7	17.3	16.66	0.94	-0.64	11.33	+0.373	17.01	**	-0.22	15.62	+0.106
70/71	16.0	16.6	17.29	0.86	+0.69	11.58	+0.351	17.32	**	+0.72	18.47	-0.088
71/72	15.9	16.8	17.16	0.83	+0.36	11.65	+0.344	17.08	0.77	+0.28	19.17	-0.135
72/73	17.2	16.9	17.56	0.87	+0.66	12.72	+0.272	16.84	0.66	-0.06	18.89	-0.117
73/74	16.6	16.8	17.24	0.79	+0.44	13.04	+0.249	16.95	0.04	+0.15	19.20	-0.138
74/75	16.2	17.0	17.08	0.75	+0.08	13.06	+0.248	16.97	0.37	-0.03	19.17	-0.136
75/76	16.2	16.6	17.07	0.72	+0.47	13.16	+0.239	16.97	0.27	+0.37	19.42	-0.154
76/77	15.3	17.6	16.82	0.71	-0.78	13.77	+0.205	17.07	0.27	-0.53	20.33	-0.208
77/78	15.9	18.2	17.02	0.69	-1.18	13.85	+0.204	17.02	0.24	-1.18	20.38	-0.061
78/79	17.4	18.3	17.40	0.79	-0.90	12.64	+0.283	16.82	0.53	-1.48	16.24	+0.061
79/80	14.4	17.0	16.72	0.79	-0.28	13.02	+0.260	17.12	0.69	+0.12	15.96	+0.078
80/81	14.3	15.8	16.75	0.76	+0.95	11.87	+0.330	17.08	0.64	+1.28	13.49	+0.229
81/82	16.5	17.6	17.31	0.74	-0.89	11.49	+0.357	17.27	0.67	-0.93	12.79	+0.277

** Undefined

Table 2

Predictions, Prediction Errors and Revised Coefficients

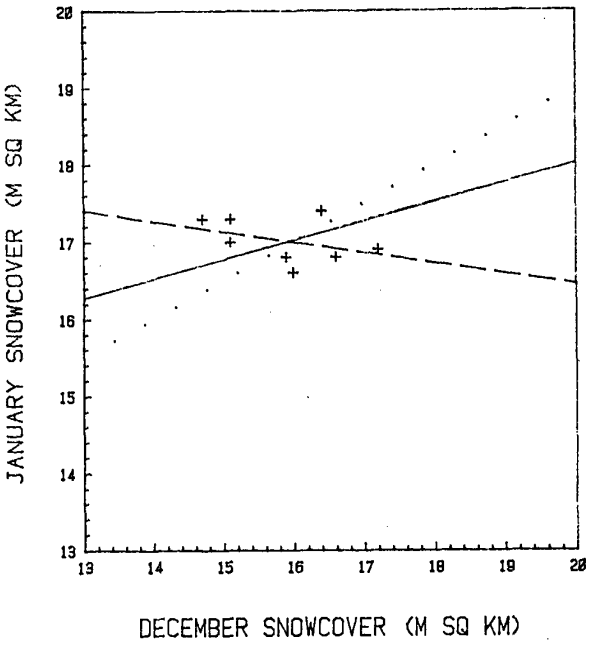


Fig. 2. The prior regression line (dotted) before any observations were available, the observations accumulated during 1966-1974, the least squares fit to the eight observations (dashed line), and the regression line (solid) posterior to those observations.

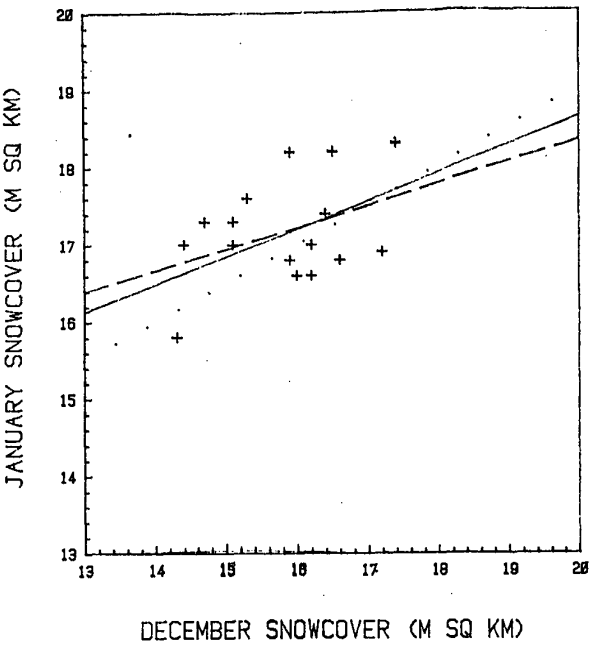


Fig. 3. Same as Figure 2 but including 16 observations over the period 1966-1982.

data are few the expert's prior judgment can make a very substantial difference.

5. CONCLUSIONS

If observations of a phenomenon are not plentiful, and a prediction is needed, then Bayes' Theorem provides a useful tool for incorporating the expert's best judgment with whatever data are available. The method we have described for simple linear regression requires that the expert, in order to fully codify his or her beliefs, specify seven scalar parameters. We have described a set of steps and thought processes whereby this can reasonably be done. Our approach is not unique. The only requirement is that the prior probabilities implied by the choice of parameters truly represent the experts' view. We found it particularly useful to consider the reaction one would have to somewhat anomalous hypothetical future data as a means for choosing from among different values of prior parameters.

The predictions of snowcover from December to January (one month predictions), made on the basis of the probabilities that we assert represented our prior beliefs, were reasonable. The observations lead us to conclude that December snowcover is not as useful a predictor of January snowcover as we had earlier believed. A comparison with predictions made using an uninformative (vague) prior reflects favorably on the use of an informative prior in that the prediction relationship is more stable from step to step, and the errors of the prediction bear a more credible relationship to the standard errors or prediction.

There is no way to arrive at a universal statement of prior beliefs, nor is there reason to expect experts' opinions to be the same. There will arise situations in which the expert, having available at most a few observations, must state his informed view, and then this approach is useful. We think that we were reasonable in our selection of prior parameters, and that others, while differing with us in some details, would not argue for grossly different expressions of prior belief. Because our prior parameters were reasonable from both probabilistic and physical points of view, it turned out that the predictions were reasonable. Unreasonable priors can lead to unreasonable predictions. The methods we describe can help to combine judgment and empirical evidence, but cannot overcome limitations of scant or faulty prior information.

6. REFERENCES

Dewey, K. F. and R. Heim, Jr., 1982: A digital archive of Northern Hemisphere snow cover. Bull. Amer. Meteor. Soc., 63(10), 1132-1141.

Raiffa, H. and R. Schlaifer, 1961: Applied Statistical Decision Theory. Harvard Business School, Boston.

PREDICTION OF SEASONAL TEMPERATURES BASED ON PERSISTENCES IN CYCLONE FREQUENCY EOFs

Bruce P. Hayden

Department of Environmental Sciences
University of Virginia
Charlottesville, Virginia 22903

1. INTRODUCTION

Season-to-season persistences in the EOFs of six-month cyclone frequencies for Eastern North America and the Western North Atlantic have recently been used to predict the spatial fields of cyclone numbers a season in advance (Hayden and Smith, 1982). In this paper I show that the persistences in cyclone frequency empirical orthogonal functions (EOFs) are maintained when station temperature data is added to the data matrix. The portion of the temperature variation that is associated with seasonal cyclone patterns that persist in time can in turn be forecast with good skill.

2. EARLIER WORK

Skillful forecasts of instantaneous weather patterns only seem possible out to about two weeks, but longer-period averaged properties of the atmosphere appear to be predictable for periods beyond two weeks (Lorenz, 1982). In spite of this encouraging prospect the recent comprehensive review of the current value, status and prospects of long-range prediction was less optimistic (Nicholls, 1980). The gap between promise and progress is, however, showing signs of narrowing. One of the promising lines of inquiry involves the use of multivariate statistical prediction models.

Early use of multivariate techniques focused on the prediction of pre-instrumental atmospheric conditions from modern data, (Fritts et al., 1971, Webb and Bryson, 1972, and CLIMAP, 1976). Gilman (1957) used the principal components of pressure, precipitation and temperature fields to predict one variable from the others. White et al. (1958) used lag correlations between two sets of EOFs to forecast surface pressure fields 12 h in advance. Bauer and Kutzbach (1974) used a similar scheme to predict heights of the 700 mb surface several days in advance. Davis (1976, 1978) also used persistence in EOF weightings but on a seasonal level in order to predict sea level pressure from SST fields.

In earlier work we (Hayden and Smith, 1982) used a multivariate statistical prediction scheme to forecast cyclone numbers a season in advance. Winter and summer half-year cyclone frequencies for eastern North America and the western North Atlantic were tabulated for 2.5° latitude by 5° longitude grid cells for the years 1885-1980. Correlation matrix eigenvectors were calculated for matrixes of both the winter and summer data for a dependent set of years (1885-1959). The first four eigenvectors for winter were highly similar in form to their summer counterparts and were correlated with them in the time domain. These correlations permitted a season-in-advance estimation of half-year cyclone frequencies. Forecast skill for independent years (1960-1980) in both winter and summer averaged 75% relative to the 1885-1959 means as forecasts and outperformed simple and damped persistence as well as climatology as forecasts.

Given the success at predicting seasonal cyclone frequencies and tracks, it seemed reasonable that some of the related variation in temperature fields might also be predictable. This paper reports on the forecast models developed and the success of forecast trials.

3. THE DATA

Cyclone frequencies for six-month seasons for the years 1900-1980 were tabulated for the 87 2.5° latitude by 5.0° longitude grid cells comprising the study areas shown in Fig. 1. From monthly charts of the "Tracks of the Centers of Cyclones at Sea Level" published by the Monthly Weather Review and in recent years by the Mariners Weather Log, seasonal totals of cyclones passing through each grid cell were recorded. Multiple entries of a given storm in a grid cell were ignored. Frequencies were not adjusted for latitude variations in grid-cell area because of distortions involved in such adjustments (Hayden, 1981b). Three-month temperature means for October-November-December, and January-February-March were collected for the period 1900 to 1980 at the eighteen U.S. stations shown in Figure 1. The matrix of the two three-month temperature means for six stations was added to the matrix of cyclone frequency counts. There were two consecutive three-month temperature records for each of the six stations. The resulting data matrixes had the dimensions of 87 cyclone frequency grid-cell variables plus 12 temperature variables and 81 yearly cases. Because 18 stations were selected but only six stations were used at one time, three separate data matrixes were assembled.

For the purpose of constructing and testing the prediction model, the data matrixes were divided into a dependent (1900-1959) part from which the eigenvectors were calculated and the forecast scheme constructed, and an independent (1960-1980) part which were used to evaluate forecast skill.

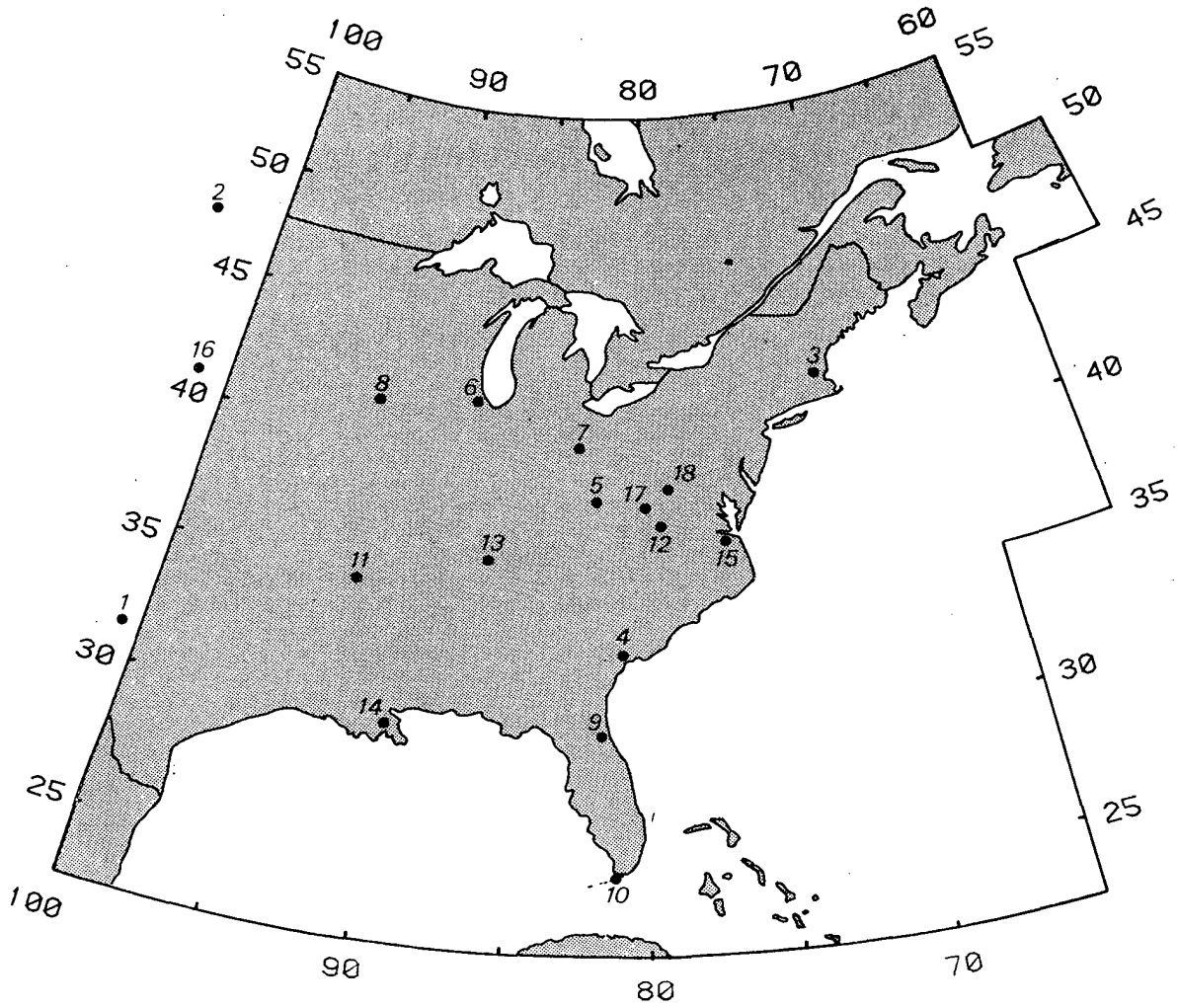


Figure 1. Study area indicating the location of the 18 stations used (see Table 1).

THE MULTIVARIATE STATISTICAL FORECAST SCHEME

4.1 The analysis

Principal components analysis has been used to resolve the variance structure in multivariate, geophysical data (Kutzbach, 1967; Fritts et al., 1971; Resio and Hayden, 1975; Hayden, 1981a), and in terms of least square errors, this type of analysis provides a method for determining patterns in large data fields (Lorenz, 1956; Gilman, 1957; Kutzbach, 1967). The objective of the analysis is to isolate characteristic, recurrent and independent modes of covariance among variables into a new set of independent variables. Basically, the analysis transforms a set of intercorrelated variables into a new coordinate system in which the axes are linear combinations of the original variates and are mutually orthogonal. The procedures for calculation of the eigenvectors of the correlation matrix follow those of Kutzbach (1967) and Vincent et al. (1976).

Principal components analysis provides a description of the uncorrelated major modes of variation in the data set. Typically, an attempt is made to identify the first several components with some property of the data field. The analysis also provides an index which measures the importance of each component within each case. The persistence identified in this index is the basis of the forecast skill reported in this paper. Finally, the analysis provides an estimate of the total percent of variance in the data set which can be explained on the basis of each component. Separate principal component analyses were run for the 1900-1959 winter and summer data matrixes.

Richman (1981) has shown that unrotated EOFs do not always adequately define actual patterns in the data field. He suggested using orthogonal rotations of the EOFs (VARIMAX) and oblique rotations of the EOFs (PROMAX) to better define organized variance in the data. In this paper EOFs, VARIMAX rotated EOFs, and Promax rotated EOFs are also used to construct forecast models. Each of the three model types is evaluated for forecast skill.

4.2 The predictive equations

The cyclone frequency values at each of the 87 grid cells and temperature data at each of the 18 stations (Fig. 1) for a given season (C_s) may be estimated from

$$C_s = X + a^*_s \sigma E_1 + a^*_2 \sigma E_2 + a^*_s \sigma E_3 + a^*_s \sigma E_4,$$

where X is the matrix of long-term (1885-1959) mean cyclone frequency and temperature, and σ is the standard deviation matrix of X; E_i are the four eigenvector components and a^*_i are the predicted eigenvector weightings. The a^*_s are calculated from regressions between the time series of eigenvector weightings for one six-month cyclone season (the predictor) and the time series of eigenvector weightings for the following prediction (the predictand) and season for the years (1900-1969). No cross lags between vectors are used in predicting a^*_s. In short, this forecast method states that the cyclone frequency and attendant temperature pattern for the coming season should be similar to the average for that season with the addition of four persistence components based upon the occurrences of cyclones during the preceding season. With this scheme a forecast of cyclone frequencies and temperatures is produced for each 2.5° latitude by 5.0° longitude grid cell of the study area and temperatures at the 18 stations (Fig.1).

As with all prediction models criteria for acceptance and use or rejection must be established. Usually the criteria are threshold values of skill based on forecast trials on independent data. Here the independent data period used in the forecast trials is the 21 years from 1960 to 1980. Two types of skill measures may be considered.

Since the cyclone frequency forecast for the predictor and predictand seasons selected (O-M and A-S) are published elsewhere (Hayden and Smith, 1982) only the forecast skill for the added (temperature) variables is treated here. The cyclone frequency and temperature forecasts for each grid cell can be compared to the observed value occurrence and forecast skill can be quantified.

5. FORECAST VERIFICATION

Local forecast skill is the skill at a station for a specific forecast period and lead time. Global skill is the aggregate or average skill over all stations in the network for a specific forecast period and lead time. Local skills are based on a smaller sample of forecast trials and independence from trial to trial is assumed. Spatial independence in global skill score can not be assumed (Livezey and Chen 1983) and statistical significance of the forecast method must be adjusted for the estimated degrees of freedom. If, however, the criterion of Madden and Julian (1971) is used where local skill score acceptance levels are raised to the point where occurrence by the accident of chance is unlikely ($p = 0.05$), then potential effects due to spatial dependence in global skill scores can be ignored (Livezey and Chen 1983). Livezey and Chen label this approach "over conservative," but where new methods for seasonal climate forecasts are offered for scrutiny, the over-conservative standard is justified. In this study forecast models for each vector for each station are screened and rejected if in the 21 years of trial forecasts at least 15 correct forecasts were made (71.4% correct; $p = 0.04$). Using this screening procedure the resulting forecast equation only includes those terms (eigenvectors) that individually pass the stringent local percent correct skill of 71.4%. In the unrotated models used here only the first four eigenvectors are included in the forecast trials because they each passed the statistical significance test of Overland and Preisendorfer (1982). The resulting forecast equation may thus have between one and four terms. In the varimax and promax rotation models the first five vectors are used.

6. FORECAST PROCEDURE

The EOFs of the predictor six-month season (April through September) include data for the season and year immediately preceding the season to be predicted. Thus on an operational basis forecasts could be made at the end of September for the October-November-December and January-February-March temperature seasons. The first has effectively no lead time and the latter a three-month lead time.

Each EOF is screened using the "over conservative" criterion of Livezey and Chen (1983) and the resulting forecast equations may have from one to four terms in the case of the unrotated EOF models and between one and five terms in the rotated EOF models. Generally the resulting equation has either two or three terms.

Each forecast equation then results in a prediction of either greater than the long-term mean temperature or less than the long-term mean temperature. These forecasts are compared to observed values for the independent data years and the percent of correct forecasts is determined.

7. RESULTS OF FORECAST TRIALS

Table I summarizes the results of forecast trials on independent data by station, season, and model type. In general, forecast skill for the winter three-month season exceeded that for the fall three-month season. The number of stations which were not forecastable using the "over conservative" skill criterion was larger in the fall (six stations) than in winter (one to five stations). Acceptable forecast skills are constrained to be no smaller than 71.4% and the average skill for stations forecast is about 80%. Stations in the western part of the study area are the most difficult to forecast with skill.

8. DISCUSSION AND CONCLUSION

Seasonal temperature forecast models were constructed and tested on independent data. Stringent statistical criteria were used for acceptance or rejection of model components. Forecast skills on a percent-correct basis in the

vicinity of 80% are demonstrated. Forecast skill varied by season, geography, and model type. Climate forecasts are subject to considerable scrutiny by the forecaster's peers, (Jolliffe, 1979; Nicholls, 1980; Harmack and Sammler, 1982). A unifying standard is called for because a diverse assortment of measures of forecast skill is available and each measure has a different interpretation. The "over conservative" station screening criteria suggested by Livezey and Chen (1983) is particularly attractive because a widely used statistical criterion (.05 significance level) is applied. With this criterion a measure of normalization for number of forecast trials is achieved. A reliance on a particular skill score is thus not required.

Seasonal temperature forecast trials have also been run on all other three-month seasons with commensurate forecast skill. Precipitation forecasts have likewise been run and verified with success.

SUMMARY OF RESULTS OF SEASONAL TEMPERATURE FORECAST TRIALS
ON INDEPENDENT DATA (1960-1980)
(Percent correct skill scores are given)

	AUTUMN			WINTER		
Station	EOF	Varimax	Oblique	EOF	Varimax	Oblique
1. Abilene, TX	*	*	*	71.4	76.2	76.2
2. Bismarck, ND	*	*	*	*	*	*
3. Blue Hill, MA	85.7	85.7	85.7	71.4	71.4	71.4
4. Charleston, SC	71.4	76.2	71.4	76.2	81.0	81.0
5. Charleston, WV	76.2	85.7	85.7	85.7	90.5	90.5
6. Chicago, IL	*	*	*	*	81.0	81.0
7. Columbus, OH	71.4	76.2	76.2	81.0	76.2	76.2
8. Des Moines, IA	*	*	*	*	71.4	71.4
9. Jacksonville, FL	*	*	*	71.4	76.2	81.0
10. Key West, FL	*	71.4	71.4	71.4	81.0	81.0
11. Little Rock, AR	71.4	*	*	*	81.0	81.0
12. Lynchburg, VA	76.2	71.4	76.2	90.5	85.7	81.0
13. Nashville, TN	71.4	76.2	76.2	81.0	81.0	85.7
14. New Orleans, LA	71.4	90.5	85.7	81.0	95.2	95.2
15. Norfolk, VA	71.4	76.2	76.2	81.0	85.7	85.7
16. North Platte, NE	76.2	81.0	81.0	*	*	71.4
17. Hot Springs, VA	76.2	71.4	71.4	80.0	81.0	76.2
18. Woodstock, VA	71.4	71.4	71.4	85.7	85.7	85.7
Average Percent Correct	74.2	77.8	77.8	78.3	81.4	80.4
No. Correct	12/18	12/18	12/18	13/18	16/18	17/18
Percent of Stations Forecast	66.7	66.7	66.7	72.2	88.8	94.4

*None of the vectors passed the 0.05 significance level - no model selected (71.4% correct = .040 level of significance).

REFERENCES

Bauer, K.G., and J.E. Kutzbach, 1974: Evaluation standards for dynamical prediction models. J. Appl. Meteor., **13**, 505-506.

CLIMAP Project members, 1976: The surface of the Ice-age earth, Science **191**, 1131-1137.

Davis, R.E., 1976: Predictability of sea surface temperature and sea level pressure anomalies over the north Pacific Ocean. J. Phys. Oceanogr.,**6**, 249-266.

_____, 1978: Predictability of sea level pressure anomalies over the north Pacific Ocean. J. Phys. Oceanogr., **8**, 233-246.

Fritts, H., T. Blasing, B. Hayden and J. Kutzbach, 1971: Multivariate techniques for specifying tree growth and climate relationships and for reconstructing anomalies in paleoclimate. J. Appl. Meteor., **10**, 845-868.

Gilman, D., 1957: Empirical orthogonal functions applied to thirty-day forecasting. Sci. Rep. No. 1, MIT, Dept. Meteor. CFSTI, 120 p. (NTIS AD117249).

Harnack, R.P. and W.R. Sammler: 1982: Performance of the 1976 University of Wisconsin model for United States long-range forecasts made for 1976-1980. Bull. Amer. Meteor. Soc., **63**(1), 23-28.

Hayden, B.P., 1981a: Secular variation in Atlantic coast extratropical cyclones. Mon. Wea. Rev., **109**, 159-167.

_____, 1981b: Cyclone occurrence mapping - equal area or raw frequencies? Mon. Wea. Rev., **109**, 168-172.

_____, and W. Smith, 1982: Season-to-season cyclone frequency prediction. Mon. Wea. Rev., **110**(4). 239-253.

Jolliffe, Ian, 1979: How accurate are long-range weather forecasts? New Scientist **84**(1177), 192-194.

Kutzbach, J., 1967: Empirical eigenvectors of sea level pressure, surface temperature and precipitation complex over North America. J. Appl. Meteor., **6**, 791-802.

Livezey, R.E. and W.Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. Mon. Wea. Rev., **111**(1), 46-59.

Lorenz, E.N., 1982: Atmospheric predictability experiments with a large numerical model. Tellus **34**, 505-513.

Lorenz, E.N., 1956: Empirical orthogonal functions and statistical weather prediction. Sci. Rep. 1, MIT, Dept. of Meteor. CFSTI, 19 p. (NTIS AD110268).

Madden, R.A. and P.R. Julian, 1971: Detection of a 40-50 day oscillation in the zonal wind in the tropical Pacific. J. Atmos. Sci., **28**, 702-708.

Nicholls, N., 1980: Long-range weather forecasting: values, status, and prospects. Rev. Geophys. Space Phys., **18**, 771-788.

Resio, D.T. and B.P. Hayden, 1975: Recent secular variations in mid-Atlantic winter extratropical storm climate. J. Appl. Meteor., **14**, 1223-1234.

Richman, M.B., 1981: Obliquely rotated principal components: an improved meteorological map typing technique? Jour. Appl. Meteor., **20**, 1145-1159.

Vincent, L., R. Dolan, B. Hayden and D. Resio, 1976: Systematic variations in barrier island topogrphy. J. Geol., **84**, 583-594.

White, R.M., D.S. Cooley, R.C. Derby and F.A. Seaver, 1958: The development of efficient linear statistical operators for the prediction of sea-level pressure. J. Meteor., **15**, 426-434.

SOME RESULTS OF APPLICATIONS OF THE STATISTICAL METHOD TO THE CLIMATE CHANGES AND SHORT-TERM CLIMATE PREDICTION IN CHINA

Mai-Tsun Li

Visiting Professor
Department of Meteorology and Physical Oceanography
Massachusetts Institute of Technology
Cambridge, MA
USA

Director
Laboratory of the General Circulation and Long-Range Forecasting
Institute of Atmospheric Physics
Academia Sinica
Beijing, People's Republic of China

Di-Rong Yao

Department of Geography
Hangzhou University
Hangzhou, People's Republic of China

1. INTRODUCTION

In recent years, much attention has been paid to climate changes and short-term climate prediction, because the climate anomaly of the large area has constantly occurred in many areas of China and various other countries since the 1960's. From 1974 to 1981, ten national meetings on climatology in China have been held. More than 100 papers were presented at each of these meetings. Two collections of papers presented at these meetings have been printed. At present, in addition to research at meteorological institutes and universities, the research of climate changes and short-term climate prediction is also being carried out in the weather bureaus of the 30 autonomous municipalities and provinces (including Taiwan). The gaps in the research in this field in China have been filled (Zhang, X., 1982).

In the study of climate changes and short-term climate prediction, statistical analysis is now an indispensable method. In recent years, widespread applications of statistical methods to the climatological data and short-term climatic prediction in China have progressed. There are many long-range climatic data. When these statistical methods are applied to these long-range data, many interesting results arise. This paper gives a summary of the results in this field using the following approaches, some of which have been refined by Chinese scientists: correlation analysis, periodicity analysis, time series predictions, regression models, stage analysis, variance analysis of series of pseudovariables, empirical orthogonal functions, Chebyshev polynomial analysis, and stochastic dynamic models.

2. CORRELATION ANALYSIS

The method of using correlation analysis to expose the laws of climatic data evolution has been widely developed.

In the analysis of droughts and floods for the last 500 years in China (Xu and Wang, 1981), the correlation coefficients among drought indices in the country and those in the latitude belts and among those in each pair of latitude belts were calculated. The results showed that the drought index changes of North China are representative of those of the country as a whole, while droughts in China as a whole and droughts in South China are not closely allied; the droughts in North China are often connected with

the droughts in the Huahe River basin or those in NE and NW China. Furthermore, the correlation analysis between the drought indices of each latitude belt and the sea level pressure field of the world for July of the same year showed that the subtropical high intensity in winter in the Southern Hemisphere is connected with the rainfall in the summer months (May to September) in southern areas in China. The results of the analysis are shown in Figs. 1-3.

Wang and Zhao (1981) calculated the correlation coefficients between the droughts-floods and atmospheric action centers and between the colds-warms and atmospheric action centers for initial series, long-range changes, and two- to three-year changes, respectively. The results showed that the relationship between the climate changes and general circulation varies with the length of the periodicities. Therefore, the mechanisms of the form of the variant periods may be different, but the periods of each kind have stationary areas. This fact shows that the climate changes and the circulation mechanisms may have the character of stationary waves.

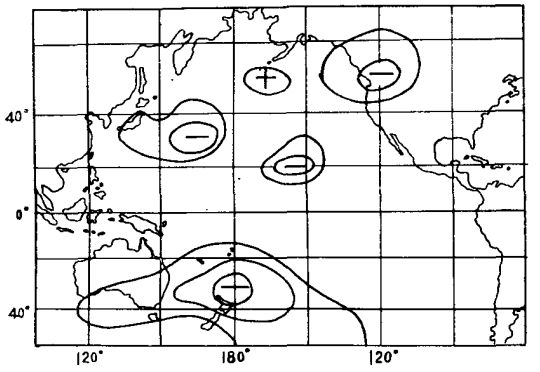


Fig. 1. The distribution diagram of the correlation region between the drought indices of the first latitude belt and the sea level pressure field in July. From the exterior to the interior, the levels of significance of the equiscalar lines are 0.05, 0.01, and 0.001, respectively.

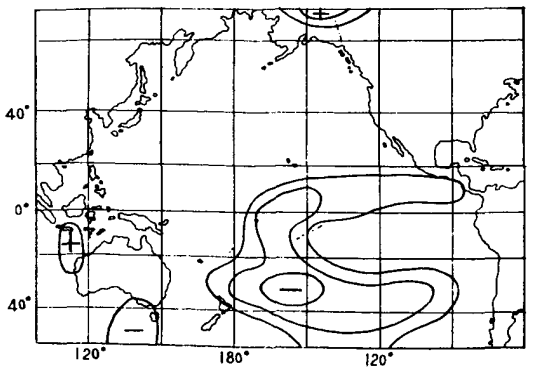


Fig. 2. The distribution diagram of the correlation region between the drought indices of the second latitude belt and the sea level pressure field in July.

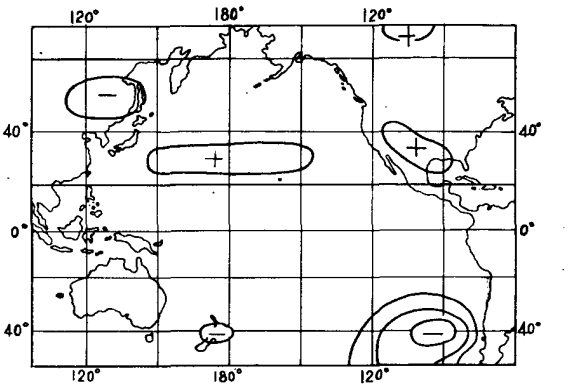


Fig. 3. The distribution diagram of the correlation region between the drought indices of the third latitude belt and the sea level pressure field in July.

The establishment of the long-range sequence of climate history is important for researching climate changes over historical periods. In the analysis of the temperature variation in winter in southern China for the last 100 years (Zhang and Zhu, 1981), the correlation coefficients among the temperatures of each station for each climatic region in China were calculated. The results showed that it is reasonable to use a temperature index series to express the temperature of a climatic region. Then we can get long-range series of temperature and apply them to the study of cold-warm laws over historical periods.

In the western areas of China, the research of dendroclimatology has been developed, because the references and data are lacking in this area. The correlation analysis between the breadth data of neighboring meteorological stations was made. The climate data in this area have been effectively extended (Lin and Wu, 1977; Wu et al., 1977; Meteorological Observatory of Hunan Province, 1977; Meteorological Observatory of Shanxi Province and Institute of Meteorology, 1977; Zhang et al., 1981; Zhuo et al., 1978; Gong et al., 1979).

Many correlation analyses showed that East Asia circulation indices, the South Oscillation indices, and the positions and intensities of the subtropical high and Indian low are all important circulation predictors for the droughts and floods in China (Yang, 1977; Plan Office of the Yangtze Region, 1975; Li, R. et al., 1977; Ci et al., 1974; Wang and Zhao, 1979). The correlations between the SST of the Pacific Kuroshio and the droughts-floods of eastern areas in China are very strong too (Institute of Atmospheric Physics, 1973, 1979; Institute of the Geographic, Academia Sinica, 1979; Zhao et al., 1981). Jang et al. (1981) and Q. Xu et al. (1975) pointed out that the beginning of plum rains, the amount of rainfall, and the duration of plum rains are related not only to the SST of Kuroshio but also to the coastal SST of Peru in winter. R. Xu and Wang (1981) pointed out that the sunspot number is connected with the yearly frequency of various circulation types, and the values of cross-correlation coefficients between the data, with a lag of several years, reached their maximum. This may provide a definite basis for the

short-term climate prediction of the circulation. In addition, the relationship between the number of sunspots and the droughts-floods of lower latitude areas in China is linear, while the linear correlation is less obvious in other areas. This latter result indicates that the relationship between the number of sunspots and the droughts-floods in China as a whole can be nonlinear. Lin (1981) pointed out that the oscillation of climatic periodicity determines the instability of the correlation coefficients.

3. PERIODICITY ANALYSIS

Periodicity analysis is widely used in the analysis of climate data. Variance analysis, energy spectrum analysis, and harmonic analysis are the most frequently used methods.

In research on climate changes over historical periods in the Qinghai-Xizang plateau, energy spectrum analysis and variance analysis for the sequence of the annual tree ring breadth were made (Lin and Wu, 1977). The results obtained were that the main periods of two to three years are consistent with the famous "two-year pulsation" in meteorology. This explains why the changes of the annual ring breadths and meteorological elements are affected by the same general circulation.

Energy analysis and variance analysis for the series of the mean flow of Hankou between June and September were performed (Plan Office of the Yangtze Region, 1977). The results showed that the flow changes of the Yangtze River form a mixed wave. It consists of waves of different periods, i.e., short waves of two to four years, medium waves of six to seven years and fourteen years, long waves of forty-one years, and ultralong waves of about one century.

In research on the periodicity of cold-warm and drought-flood, energy spectrum analysis and cross spectrum analysis of the temperature types in January and rainfall types in July were used (Department of Geophysics, 1977). The results showed that the temperature in January mainly has periods of more than twenty years and quasiperiods of two years and the rainfall in July mainly has periods of around four years and quasiperiods of two years and secondarily has periods of seven and thirty-five years. The data used in the energy spectrum analysis were ten-year moving averages. It can be seen that the period of around twenty years is very prominent for the temperature in January, while the period of around thirty-five years is very prominent for the rainfall types in July. The cross-spectrum analysis confirmed that the major relationship between the high in the South Pacific and rainfall in China is the period of around thirty-five years in the low-frequency part; the relationship of twenty-year periods between the high intensities in Siberia and the temperatures in China is very strong. This result shows that the general circulation plays a very important role in climate changes.

Energy analysis has been applied to the temperature and rainfall series in the Qinghai-Xizang after low-pass filtering (Wu and Lin, 1981). The results showed that many series have remarkable periods of general significance. These periods include quasipulsation of two years, the period of around eleven years, and the Bruckna period of more than thirty years.

The SP_{01} index expresses the major trend of the climate changes over the last 100 years in China. In Zhang, X. et al. (1981), energy spectrum analysis and harmonic analysis for the SP_{01} index series of ten-year moving averages were conducted. The results explained that the periods of the climate changes in China have been 600-700, 300-400, 180-220, 110-150, 80-90, 26, 13, and 2-3 years. These periods of the climate changes may be connected with the periods of sunspot movements.

In addition, the relationship between the increment of the breadth of the annual tree rings and intensity of the sunspot movement was analyzed. It was found that the increment of the breadth of the trees's annual rings is connected with the century periods of the sunspot movement, and the increment of the breadth of the trees's annual rings is somewhat related with a twenty-two year period of the sunspot movement.

The results of energy analysis and variance analysis for the drought index of the whole country showed that the periods above the 95% significance level are 2-3, 8-10, 22-26, 30-50, 140-170, and 320-350 years (Zhang, 1981). In the low-latitude areas, one must consider not only the changes of short periods, but also those of long periods. Therefore, it is more difficult to analyze the arid-moist changes south of the Yangtze River than north of the Yangtze River.

In Wang and Zhao (1981) the energy analysis of cold-warm and drought-flood for stations throughout the country was made. The results of the analysis are shown in Fig. 4.

It may be seen that the important periods are 36, 22, and 2-3 years. While for different areas predominant periods are not the same, the distribution areas of each period of the drought-flood are consistent with those of the cold-warm. The same analysis for the global general circulation has also been made. The results are shown in Fig. 5.

It can be seen that the periods of 36, 22, and 2-3 years are also significant in the global atmosphere. The periods of 36 and 22 years are most obvious in the Southern Pacific, Australia, and the zone of the Western Pacific to Asia, and these two periods are also reflected in East Asia general circulation. The periods of 2-3 years are concentrated in Eurasia, Eastern North America through the North Atlantic to Western Europe, and the Australia-New Zealand area.

In a word, a great deal of work in this field has been done, and many results of significance have been obtained (Zheng et al., 1977; Tang, 1977; Meteorological Observatory of Xinjiang, 1977; Bureau of Jiangxi Province, 1981).

Besides, for predictions of the climatic trends of typhoon movements, rainfall, temperature, SP_{01} index, and so forth, the method of analysis of variance was used widely (Zhang, X. et al., 1981; Ding, 1977; Xue, 1977; Institute of the Geographic, Hebei Province, 1977; Yan et al., 1981), and good results were obtained.

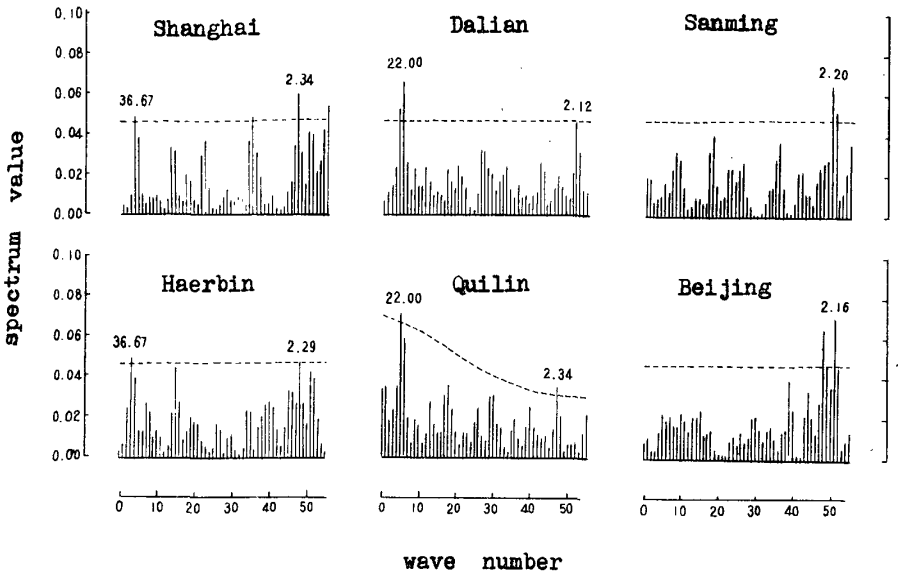


Fig. 4. The energy spectrum of drought-flood (top) and cold-warm (bottom). The dashed lines are the dividing lines at the 0.05 level of significance; values are the numbers of years of the major periods.

4. TIME SERIES PREDICTIONS

This method is mainly used for short-term climate predictions. In most of the cases the following additive model is assumed:

$$X(t) = f(t) + S(t) + \eta(t) + \xi(t) \qquad (1)$$

where $X(t)$ is a realization of a stochastic process, which is composed of a trend term, $f(t)$, a period term, $S(t)$, a stationary stochastic term, $\eta(t)$, and a white noise term, $\xi(t)$. This model was used in the rainfall predictions of twenty-three stations of the Haihe and Luanhe drainage (Institute of the Geographic, Hebei Province, 1977). Based on the data of yearly frequency number of the circulation types, W, C, and E, in the Northern Hemisphere since 1891, J. Zhang (1975 a, b) studied the law of secular changes of the general circulation and its relationship with the climatic oscillation in China. Using the additive model of nonstationary time series, Zhang also made short-term predictions of yearly frequency numbers of the circulation types and yearly rainfall in Eastern China.

According to the curves of historical evolution of rainfall and sunspots, the predictions of rainfall or rainfall and sunspots were made by using two-dimensional stochastic processes. In the predictions of the Haihe and Luanhe drainage this method was used most successfully (Institute of the Geographic, Hebei Province, 1977).

Yao (1981) first applied Markov chain theory by using pseudovariables in the study of historical climate changes. Using this method, we can estimate the cycle time of drought-flood or that of cold-warm, the amount of time that drought-flood or cold-warm is sustained, the yearly change ratio of the sustained state, and so on.

5. REGRESSION MODELS

Regression models were widely applied to predictions of temperature, rainfall, the breadth of trees's annual rings, and so on. Furthermore, in addition to linear regression models, nonlinear regression models were also used (Wu et al., 1977; Meteorological Observatory of Fujian Province, 1977). They have all obtained good results.

The methods for fitting the autoregressive model were also applied to climatic prediction. Yao (1982) pointed out that the last parameter of the autoregressive model used for climatic prediction is just the partial autocorrelation coefficient, and therefore the order of the model can be chosen by means of the t-test or F-test. The fitting of the autoregressive model can thus be performed by a simple procedure if the auto-regressive parameters are estimated by the recurrence formula. In the forecast of the monthly rainfall this method was used most successfully.

6. STAGE ANALYSIS

Stage analysis is a major method used in climatic prediction in China (Chang, 1977, 1981). It involves the use of the mean and variance of a meteorological element to identify the different stages of time evolution of a long-range time series and applications to climatic prediction (Shao, 1977). In this method, the number of stages can be adjusted by Mahalanobis distance D^2, so that the identification of the different stages may be suited to the scale of prediction. Then, with the F-test, suitable stage numbers can be obtained.

Using this method to analyze rainfall in Beijing we were able to obtain results which are in accord with the spectrum analysis. The method can be used to estimate the trend of future changes.

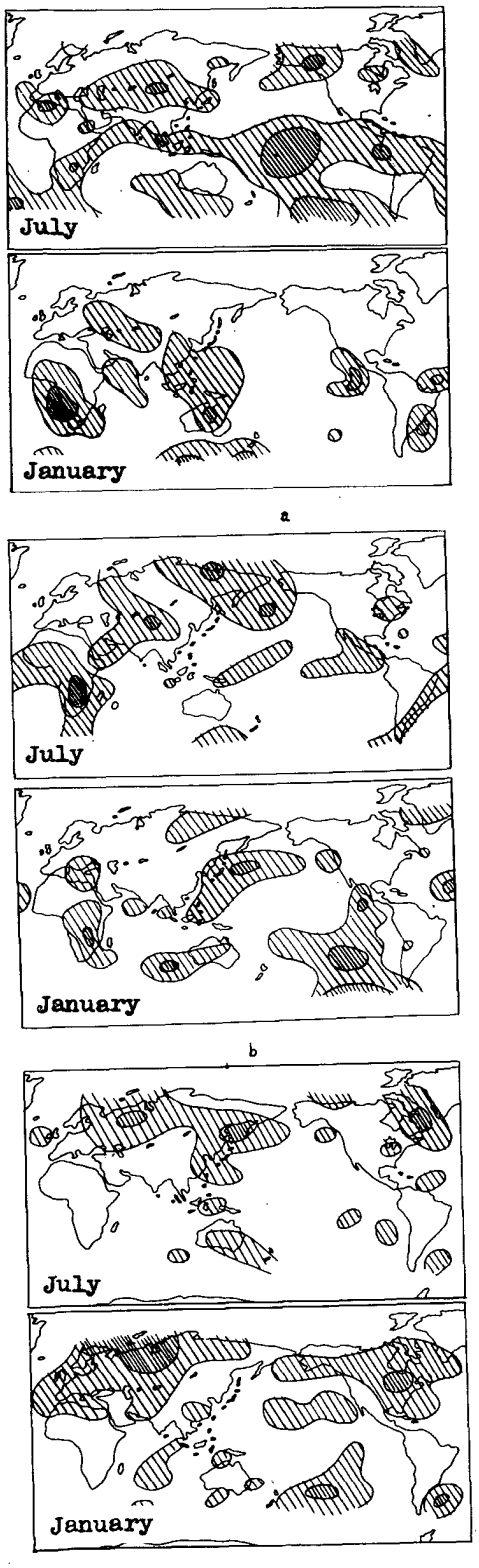


Fig. 5. The periods of (a) 36, (b) 22, and (c) 2-3 years for the general circulation in July (top) and in January (bottom). The regions with sparse and dense oblique lines represent two different levels of significance, 0.05 and 0.01.

In identifying the different stages of the drought index series, the method of optimal identification was used (Department of Meteorology, 1977). The results showed that drought also has stages and periods of different lengths. In the stage of the century length, a vibration of the second-order scale was obtained. In doing the above, one should make the variances between the drought indices at different times larger, but within a single period of time as small as possible.

In short, stage analysis has been widely used in the short-term climate prediction in China (Wang and Zhao, 1981; Ding, 1977).

7. VARIANCE ANALYSIS OF SERIES OF PSEUDOVARIABLES

In singular period analysis, Yao (1981 a) pointed out that it is better to use variance analysis of series of pseudovariables to test the significance of periods. Other methods such as spectrum analysis, periodgram analysis, and so on can obtain only those periods of frequent occurrence. Applying variance analysis to series of pseudovariables usually gives an undistorted analysis of ultralong periods.

In practical applications, variance analysis of series has proved superior to the method of the periodgram in testing the significance of periods.

8. EMPIRICAL ORTHOGONAL FUNCTIONS

In order to investigate the climatic variation of the last 140 years in China, researchers used the method of empirical orthogonal functions to analyze the types of temperature in January and the types of rainfall in July (Department of Geophysics, 1977). The January temperature at 90 stations throughout the country from 1951 to 1975 was analyzed. Five types were obtained:

Type I	warm (1952)
Type II	partially warm (1973)
Type III	normal (1953)
Type IV	partially cold (1955)
Type V	cold (1969).

The years in parentheses are typical years corresponding to the various types.

The same method was used to analyze rainfall in July at 100 stations throughout the country between 1951 and 1975. Again five types were obtained:

Type I	Abundance of rainfall throughout the country (1954)
Type II	rainy south of the Yangtze River, shortage of rain north of the Yangtze River (1974)
Type III	two rainbands (1969)
Type IV	rainy north of the Yangtze River, lack of rain south of the Yangtze River (1962)
Type V	not much rain throughout the country (1972).

Furthermore, the empirical orthogonal functions method was used to analyze the drought-flood orders at 100 stations and the cold-warm orders at 30 stations between 1871 and 1970, as well as the rainfall between June and August and the temperature in January during the years 1951 to 1974 (Wang and Zhao, 1981). The results showed

that the above-mentioned drought-flood and cold-warm series can be used to study climate changes in China.

The method of factor analysis was applied to SST fields of the North Pacific by M. Li and Huang (1981). The results showed that SST in the tropical areas (180°-90°W, 10°N-10°S) is more representative, and it is most representative in winter and spring. SST in tropical areas has the best persistence. The two-year period of SST changes is dominant in tropical low-latitude areas and in high-latitude areas, while in middle latitude sea areas the semiannual periods predominate. These characteristics of the climate changes of SST fields are very important to the study of the relationship between the pressure fields and SST, and between the rainfall fields and SST.

9 CHEBYSHEV POLYNOMIAL ANALYSIS

Previous investigations of applications of Chebyshev polynomials in meteorology dealt with only the expansion in equidistant grids. By introducing a nonlinear transformation, Chebyshev polynomials were generalized into irregular grids by Zhou (1981). This transformation is accomplished by replacing the independent variables with their ordinal numbers.

Two-dimensional ordinal numbers (ξ_1, ξ_2) are assigned to each station. The distributions of meteorological elements are expanded by using the expression,

$$\tilde{Z}_{ij} = \sum_{k=0}^{k_o} \sum_{s=0}^{s_o} A_{ks}\phi_k(i)\psi_s(j), \tag{2}$$

where $\tilde{Z}_{ij}$ denotes the meteorological elements at station (i, j), (i, j) is the two-dimensional ordinal number of the station, $\phi_k(i)$ and $\psi_s(j)$ are normalized Chebyshev polynomials, k and s are the orders of the polynomials, $A_{ks}(z)$ is the expansion coefficient, and K_o and s_o are the truncation orders of the expression.

Various distributions, including those of SST, precipitation over the middle and lower reaches of the Yangtze River, and temperatures of areas in NE China, were expanded in terms of Chebyshev polynomials to extract their main characteristics (Zhou, 1982 a, b). Moreover, by using expansion coefficients of SST as predictors of precipitation (or temperature), a series of prediction equations for coefficients $A_{ks}(R)$ or $A_{ks}(T)$ were established. The forecast of R_{ij} or T_{ij} for large areas can then be evaluated to form (2). In practical applications, this method was used most successfully.

It is worthwhile to point out that Chebyshev polynomials do not vary with time and they are suitable for predicting horizontal distributions of meteorological elements.

10. STOCHASTIC DYNAMIC MODELS

It is well known that oceans play an important role in the theory of long-range weather procedures and climatogenesis. M. Li (1981) presented a simple stochastic dynamic model to study the climatic anomaly, in which ocean and atmosphere form a coupled system.

Second-order linear differential equations with stochastic terms are

$$\frac{d^2T_s'}{dt^2} + (\lambda_2 + \lambda_3)\frac{dT_s'}{dt} + \lambda_2\lambda_4\left(\frac{\lambda_3}{\lambda_4} - \frac{\lambda_1}{\lambda_2}\right)T_s' = \lambda_2 W \tag{3}$$

and

$$\frac{d^2T'}{dt^2} + (\lambda_2 + \lambda_3)\frac{dT'}{dt} + \lambda_2\lambda_4\left(\frac{\lambda_3}{\lambda_4} - \frac{\lambda_1}{\lambda_2}\right)T' = \lambda_1\bar{W} \tag{4}$$

where

$$W = \frac{1}{\lambda_2}\frac{dW_2(t)}{dt} + \frac{\lambda_4}{\lambda_2}W_2(t) + W_1(t)$$

and

$$\bar{W} = W_2 + \frac{\lambda_3}{\lambda_1}W_1 + \frac{1}{\lambda_1}\frac{dW_1}{dt}. \tag{5}$$

If one is considering the forced effect of the stochastic process of the weather disturbance scale on the climate changes of SST, and one assumes that the spectral density of SST, T_s', is $Y(\omega)$, the solution of (3) is given by

$$Y(\omega) = \lambda_2^2\phi(\omega)/\left\{\left[\lambda_1\lambda_4\left(\frac{\lambda_3 - \lambda_1}{\lambda_4 \quad \lambda_2}\right) - \omega^2\right]^2 + (\lambda_2 + \lambda_3)^2\omega^2\right\} \tag{6}$$

where $\phi(\omega)$ is the statistical spectrum of the incoming stochastic weather disturbance; when $\lambda_2 >> \lambda_3$, (6) reduces to a first-order SST Markov model. Its feedback function is

$$\lambda_4\left(\frac{\lambda_3}{\lambda_4} - \frac{\lambda_1}{\lambda_2}\right).$$

If the stochastic input, W, is the white noise, i.e., $\phi(\omega) \sim \phi(0)$, and we let

$$\varepsilon(t) = \frac{1}{T}\int_{t-\frac{T}{2}}^{t+\frac{T}{2}} W(t)dt, \tag{7}$$

then

$$\phi(0) = T\overline{\varepsilon^2(t)}/\pi = \frac{D}{\pi}(°C)^2/yr, \tag{8}$$

where T is taken to be one year.

Taking the data studied by Frankignoul and Hasselmann (1977), from (6) we made computations, the results of which are shown in Fig. 6.

It may be seen that the results of the model of the second vibration conform more with the real situation than do those of the first model of Frankignoul and Hasselmann. When $D = 2.7(°C)^2/yr$, the statistical variance of SST obtained by Li is $1.0(°C)^2/yr$; this result agrees very well with the observations.

Stochastic dynamic models are very important for studying air-sea interaction, climate changes, and short-term climate predictions.

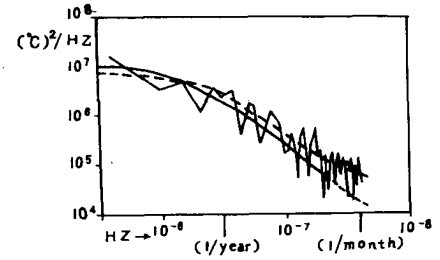


Fig. 6. Spectrum of SST departure. The unsmoothed line is the observed value; the solid line is the theoretical value according to Frankignoul and Hasselmann (1977), and the dashed line is the theoretical value according to M. Li (1981).

11. REFERENCES

Bureau of Meteorology of Jiangxi Province, 1981: Collection of Papers Presented at the China Symposium on Climatic Changes. Scientific Press, Beijing, People's Republic of China, 81-94.

Chang, Chiacheng, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 1-7.

____, 1981: Collection of Papers Presented at the China Symposium on Climatic Changes. Scientific Press, Beijing, People's Republic of China, 167-174.

Ci, Gongwang, Qinghua Zhou, and Di-Rong Yao, 1974: The Reference of Meteorological Science and Technology, No. 5.

Department of Geophysics, Beijing University, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 149-163.

Department of Meteorology, Nanking University, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 53-58.

Ding, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 123-125.

Frankignoul, C. and K. Hasselmann, 1977: Stochastic climate models, part II. Application to sea surface temperature anomalies and thermocline variability. Tellus, 29, 289-305.

Gong, Gaofa, Enjin Chen, and Wen Huanyan, 1979: Acta Geographica Sinica, 34, No. 2.

Institute of Atmospheric Physics, Academia Sinica, 1973: The Reference of Meteorological Science and Technology, No. 3.

____, 1979: Collection of Papers Presented at the China Symposium on Medium-Range and Long-Range Hydrologic Forecast, No. 1. Water-Power Press, Beijing, People's Republic of China.

Institute of the Geographic, Academia Sinica, 1979: The Reference of Meteorological Science and Technology, No. 3.

Institute of the Geographic, Hebei Province, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes. Scientific Press, Beijing, People's Republic of China, 146-148.

Jang, Yen, Yuehong Li, and Muying Bi, 1975: Collection of Papers Presented at the China Symposium on Hydrologic Forecast of the Yangtze Region, 164-177.

Li, Mai-Tsun, 1981: Acta Oceanologica Sinica, 3, No. 3.

____ and Hiayou Huang, 1981: Papers Presented at the Second China Symposium on Probabilistic and Statistical Methods in Weather Forecasting.

Li, Ruiyun et al., 1977: The Reference of Meteorological Science and Technology, May.

Lin, Shuechun, 1981: Collection of Papers Presented at the China Symposium on Climatic Changes. Scientific Press, Beijing, People's Republic of China, 131-142.

Lin, Zhenyao and Xiangding Wu, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes. Scientific Press, Beijing, People's Republic of China, 23-32.

Meteorological Observatory of Fujian Province, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 10-15.

Meteorological Observatory of Hunan Province, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 59-67.

Meteorological Observatory of Shanxi Province and Institute of Meteorology, Central Meteorological Bureau, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 72-76.

Meteorological Observatory of Xinjiang, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 117-118.

Plan Office of the Yangtze Region, 1975: Collection of Papers Presented at the China Symposium on Hydrologic Forecast of the Yangtze Region, 20-27.

____, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 139-145.

Tang, Zhongxing, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 45-49.

Wang, Shaowu and Zhongci Zhao, 1979: Acta Meteorologica Sinica, 37, No. 1.

____ and ____, 1981: Collection of Papers Presented at the China Symposium on Climatic Changes. Scientific Press, Beijing, People's Republic of China, 117-130.

Wu, Xiangding and Zhenyao Lin, 1981: Collection of Papers Presented at the China Symposium on Climatic Changes. Scientific Press, Beijing, People's Republic of China, 18-25.

____, Rongxin Tu, and Zijie Yan, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 68-71.

Xu, Qun et al., 1975: Collection of Papers Presented at the China Symposium on Hydrologic Forecast of the Yangtze Region, 142-153.

Xu, Ruizeng and Lei Wang, 1981: Collection of Papers Presented at the China Symposium on Climatic Changes. Scientific Press, Beijing, People's Republic of China, 52-63.

Xue, Heng, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 171-175.

Yan, Jiyuan, Weitong Xu et al., 1981: Collection of Papers Presented at the China Symposium on Climatic Changes. Scientific Press, Beijing, People's Republic of China, 71-77.

Yang, Jingxun, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 129-131.

Yao, Zhensheng, 1981a: Collection of Papers Presented at the China Symposium on Climatic Changes. Beijing, People's Republic of China, 160-166.

____, 1981b: Papers Presented at the Second China Symposium on Probabilistic and Statistical Methods in Weather Forecasting.

____, 1982: Scientia Atmospherica Sinica, 6, No. 4.

Zhang, Deer and Shulan Zhu, 1981: Collection of Papers Presented at the China Symposium on Climatic Changes. Scientific Press, Beijing, People's Republic of China, 64-77.

Zhang, Jijia, 1975a: Collection of Papers Presented at the China Symposium on Hydrologic Forecast of the Yangtze Region, 218-226.

____, 1975b: Collection of Papers Presented at the China Symposium on Hydrologic Forecast of the Yangtze Region, 227-236.

Zhang, Xiangong, 1981: Collection of Papers Presented at the China Symposium on Climatic Changes. Scientific Press, Beijing, People's Republic of China, 46-51.

____, 1982: The Reference of Meteorological Science and Technology, No. 1.

____, Zhen Zhao, and Ruizeng Xu, 1981: Collection of Papers Presented at the China Symposium on Climatic Changes. Scientific Press, Beijing, People's Republic of China, 26-35.

Zhao, Zhen, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 119-122.

Zhao, Zhongci, Shaowu Wang, and Zhenhua Chen, 1981: Collection of Papers Presented at the China Symposium on Long-Range Weather Forecast. Meteorological Press, Beijing, People's Republic of China, 66-73.

Zheng, Cizhong, Fuchan Zhang, and Gaofa Gong, 1977: Collection of Papers Presented at the China Symposium on Climatic Changes and Short-Term Climate Prediction. Scientific Press, Beijing, People's Republic of China, 29-32.

Zhou, Jiabin, 1981: Kexue Tongbao, No. 26.

____, 1982a: Papers Presented at the Second China Symposium on Probabilistic and Statistical Methods in Weather Forecasting.

____, 1982b: Scientia Atmospherica Sinica, 6, No. 4.

Zhuo, Zengdai, Xiangong Zhang, Yunsheng Wang et al., 1978: Journal of Lonzhou University, No. 2.

An application of the threshold autoregression procedure to climate analysis and forecasting

Jiuen Shi Qingang Zhou
National Meteorological Center
State Meteorological Administration
Beijing, China

Jingtian Xiang
Institute of Applied Mathematics, Academia Sinica
Beijing, China

1. INTRODUCTION

Meteorological data such as temperature, pressure and precipitation reflect the situation in certain field of the atmosphere state respectively. On one hand, they have their own evolution rules, and on the other hand, they are influenced by other factors. It is hard to get the determinate relation formula related to their own evolution. Hence, it is of importance to make studies on them by taking them as random series. Good results can be obtained by adopting time series analysis method(1) in analysing random series which have dependent relation between each other. However, in practise it is so complicated that linear model can not describe them.

In this paper, a non-linear model time series analysis--threshold autoregressive model in discrete time(2,3) is used. In recent years, this model is more perfect and it is very effective in application.

The basic idea of the threshold autoregressive model is to deal with non-linear systems by using the piecewise linearization. Owing to the threshold control role, the recursion stability is ensured, and limit cycle concept is introduced into the non-linear random systems. The threshold autoregressive model can effectively describe non-linear fluctuation and explain various kinds of the stable periodical cycles in nature, and reflect the catastrophe similar to jump resonance.(4)

The purpose of this paper is to make studies on monthly precipitation series by using threshold autoregressive model (TAR). The TAR procedure consists of four parts: model building, statistical identification, parameter estimation and forecasting, The stress is put on the studies on general definition of TAR, three special cases, open loop threshold autoregressive model and their application to forecasting.

2. THRESHOLD AUTOGRESSIVE MODELS IN DISCRETE TIME (TAR)

Let (Xn) be a k-dimensional time series and, for each n, let Jn be an observable random variable, taking integer values $(1,2,\ldots, l)$.

Definition: $(Xn; Jn)$ is said to be a general TAR if

$$Xn=B^{(Jn)}X_{n+A}{}^{(Jn)}X_{n-1+\varepsilon n}{}^{(Jn)}+C^{(Jn)}, \quad (1)$$

wher, for $Jn=j$, $A^{(j)}$ and $B^{(j)}$ are kxk (non-random) matrix coefficients, $C^{(j)}$ is a kxl vector of constants, and $(\varepsilon_n{}^{(j)})$ is a k-dimensional strict white noise sequence of independence random vectors with a diagonal covariance matrix. It is also assumed that $(\varepsilon_n{}^{(j)})$ and $(\varepsilon_n{}^{(j')})$ are independent for $j\neq j'$.

We now single out a few interesting special cases of the general TAR for further development.

First, let $(r_0, r_1, \ldots, r_l)$ denote a linearly ordered subset of the real numbers, such that $r_0\ r_1 \ldots r_l$, where r_0 and r_l are taken to be $-\infty$ and $+\infty$ respectively.

They define a partition of the real line R, i.e.

$$R=R_1 \cup R_2 \cup \ldots \cup R_l, \text{ say}$$

where $R_i=(r_{i-1}, r_i)$.

(1) Writing $Xn=(Xn, Xn-1, \ldots, Xn-k+1)^T$

$$A^{(j)}=\left[\begin{array}{c|c} a_1^{(j)}\ a_2^{(j)}\ \cdots\ a_{k-1}^{(j)} & a_k^{(j)} \\ \hline I_{k-1} & 0 \end{array}\right]$$

$$B^{(j)}=0,$$

$$\varepsilon_n(j)=(\varepsilon_n(j),\ 0,\ \ldots,0),$$

$$C^{(j)}=(a_0^{(j)},\ 0,\ \ldots,\ 0),$$

and $R_j^{(k)}=RxRx\ldots xRxR_jxRx\ldots xR$ is the cylinder set in the cartesian product of k real lines, on the interval R_j with dth coordinate space (d some fixed integer belonging to (1,2, ...,k)), and setting Jn=j if $Xn-1\in R_j^{(k)}$, we have

$$X_n=a_0^{(j)}+\sum_{i=1}^{k} a_i^{(j)} X_{n-i}+\varepsilon_n^{(j)}, \qquad (2)$$

conditional on $X_{n-d}\in R_j$; j=1,2,...,l. Since (J_n) is now a function of (X_n) itself, we call the univariate time series (X_n) given by equation (2) a self-exciting threshold autoregressive model of order (l;k,...,k) or SETAR (l; k,...,k) where k is repeated l times. If, for j=1,2,...,l,

$$a_i^{(j)}=0, \text{ for } i=k_j+1,\ k_j+2,\ldots,k,$$

then we call (X_n) a SETAR $(l;k_1,k_2,\ldots, k_l)$. We call $r_1,r_2,\ldots,r_{l-1}$ the thresholds. Note that an SETAR (1;k) is just a linear AR model of order k.

(2) (Xn, Yn) is called an open loop threshold autoregressive system with (Xn) as the observable output and (Yn) as the observable input, if

$$X_n=a_0^{(j)}+\sum_{i=1}^{mj} a_i^{(j)} X_{n-i}+\sum_{i=0}^{mj} b_i^{(j)} Y_{n-i}+\varepsilon_n^{(j)}, \qquad (3)$$

conditional on $Yn-d\in Rj$; (j=1,...,l), where $(\varepsilon_n^{(j)})$; j=1,...,l, are strict white noise sequences, with zero mean and finite variances and each being independent of (Yn). The l white noise sequences are assumed to be independent of one another. We denote this system by TARSO (l, $(m_1, m'_1),\ldots(m_l,m'_l)$).

(3) (Xn,Yn) is called a closed-loop threshold autoregressive system, or TARSC, if (Xn,Yn) and (Yn,Xn) are both TARSO. We assume that all the stationary white noise sequences involved are independent of one another.

3. TARSO MODELS FOR REAL DATA

(1) TARSO Model and forecasting of monthly total precipitation time series of Shanghai.

Let (Xt) be a monthly total precipitation time series of Shanghai and (Yt) be the mean kinetic energy at 500MB over Northern Hemisphere.

The following model is fitted, using only record of the first 300 months.

$$X_t=\begin{cases} {}^{(1)}\ -0.070-30.382X_{t-1}+24.488y_{t-1} \\ \quad +55.723y_{t-2}-33.994y_{t-3}+33.067y_{t-4} \\ \quad -21.055y_{t-5}+59.027y_{t-6}-9.243y_{t-7} \\ \quad \text{if } X_{t-5}\leq 78.7 \\ {}^{(2)}\ 0.058-0.110X_{t-1}-3.917X_{t-2} \\ \quad +13.139y_{t-1}-30.564y_{t-2}+81.479y_{t-3} \\ \quad -115.310y_{t-4}+183.548y_{t-5} \\ \quad -71.957y_{t-6}+32.849y_{t-7} \\ \quad \text{if } X_{t-5}>78.7 \end{cases}$$

Based on TARSO (2; (1,7),(2,7)) model, we have obtained one-step-ahead predictions of the next 3 months, and table 1 represents the results of independent test forecast (in mm) during June/1976 to Auguat/1976.

Table 1. Forecast verification on independent data (in mm)

	June	July	August
Prediction (P)	139.3	137.3	176.4
Observation (O)	139.1	142.7	232.4
|P-O|	0.2	5.4	56.0
Climate mean	178.0	144.1	134.6

(2) TARSO model and forecasting of monthly total precipitation time series of Beijing.

Let (Xt) be the monthly rainfall data of Beijing, and (Yt) be the mean kinetic energy at 500MB.

We have fitted the following non-linear time series model, using 300 months.

$$X_t = \begin{cases} {}^{(1)} \\ 0.126-0.132X_{t-1}-0.223X_{t-2} \\ +0.929X_{t-3}+10.272X_{t-4}-31.271y_{t-1} \\ -25.709y_{t-2}-2.822y_{t-3}+62.845y_{t-4} \\ +60.163y_{t-5}+72.005y_{t-6}-243.160y_{t-7} \\ \qquad \text{if } X_{t-4} \leq 38.1 \\ {}^{(2)} \\ -0.109-20.655X_{t-1}+24.280y_{t-1} \\ -68.050y_{t-2}+12.204y_{t-3} \\ +137.792y_{t-4}+103.062y_{t-5} \\ -107.265y_{t-6}+10.842y_{t-7} \\ \qquad \text{if } X_{t-4} > 38.1 \end{cases}$$

We obtained TARSO(2; (4,7), (1,7)) model, and table 2 shows the results of independent test forecast during July/1976 and August/1976.

Table 2. Forecast verification of monthly rainfall of Beijing (in mm)

	July	August
P	170.7	188.1
O	173.9	284.8
|P-O|	3.2	96.7
Climate mean	210.5	172.2

The results are good.

4. DISCUSSION

(1) So far as a stable linear system is concerned, ARMA (p,q) model and statistical identification can be built according to a certain method when limited data (Xn) (namely, the length of samples is limited) is available. As for non-linear system, it is not possible to find a general method for model building. A special non-linear method is built only in special limitation. The target for TAR to describe is non-linear system, but it is based on the piece-wise linearization of non-linear models. Therefore, TAR model can be built in accordance with the means for linear autoregressive model building. The essential difficulties in model building can be avoided. This is one of the advantages of TAR.

(2) The authors made climate analysis and forecasting of monthly precipitation and monthly mean temperature by using threshold autoregressive model. The practices show that the forecasts made with TARSO are better than those with SETAR. The good results in forecasting the tendency of climate parameter series are obtained. However, as regard quantitative forecasts, it is not so desirable to forecast extremely abnormal situation.

(3) A minor change in selecting the thresholds for TARSO has been made, namely, $Y_{n-d} \in R_j$ in formula (3) is replaced by $X_{n-d} \in R_j$. The result is good.

In the future, it is proposed that on the basis of TAR, experiments be made by adopting threshold autoregressive moving average model (TARMA) [5]. It is anticipated that results will be even better.

REFERENCE

(1) Box, G.E.P. and G.M.Jenkins, 1970: Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco.

(2) Tong, H., 1978:

On a threshold model. In Pattern Recognition and Signal Processing (C.H.Chen, ed.). Sijthoff and Noordhoff, The Netherlands.

(3) Tong, H., and K.S.Lim, 1980:

Threshold Autoregression, Limit Cycles and Cyclical Data. The Journal of the Royal Statistical Society, series B (Methodological), 42, 245-292.

(4) Tong, H., 1980:

Catastrophe Theory and Threshold Autoregressive Modelling. Tech. Rep. No. 125, Dept. of Mathematies, UMIST.

(5) Xiang, Jingtian, et al, 1983:

Non-linear threshold model of Time Series. (in preparation).

A NEW FORECAST METHOD OF TIME SERIES

Zhou Jiabin

Institute of Atmospheric Physics
Academia Sinica
Beijing, China

1. INTRODUCTION

The importance is always attached to the investigation of time series at single station. The evolution curve for meteoeological elements is a fundamental tool in county observational station in particular.

Yang(1953) proposed a method using the rule of evolution of meteorological elements in 1950s. Owing to that a new way for long-range forecasting in county meteorological station is found. Up to now this method is still used in meteorological station in China. However this method greatly depends on experience because of the lack of requisite mathematical treatment in it.

Since 1960s the stationary time series has been introduced into meteorological forecast. The ARIMA scheme has been developed in recent years and some improvements of forecast were obtained. Nevertheless the common operation of time series is too complex to be widely used in meteorological station.

Using Chebyshev polynomial in the study of time series a new forecasting method is proposed in this paper.

2. FORECAST FORMULAE

Giving T_0 grids, t=1,2,...T_0, the values of meteorological element are asigned as $Z(1)$, $Z(2)$, ..., $Z(T_0)$. The value of $Z(T_0+1)$ at grid T_0+1 should be predicted.

The function $Z(t)$ at grids t=1,2, ..., T_0, T_0+1 can be expanded in Chebyshev polynomials (Bagrov, 1958), that is

$$\tilde{Z}_{K_0}(t)=\sum_{K=0}^{K_0} A_K \varphi_K(t) \qquad (1)$$

(t=1,2,..., T_0+1.)

In particular, there is

$$\tilde{Z}_{K_0}(T_0+1)=\sum_{K=0}^{K_0} A_K \varphi_K(T_0+1) \qquad (2)$$

Here $\tilde{Z}_{K_0}(t)$ is the fitting of $Z(t)$ at the grid t, $\varphi_K(t)$ is the value of normalized Chebyshev polynomial with k order at the grid t, K_0, truncation order in the Epxs. (1) and (2), A_K, Chebyshev coefficients, which can be evaluated from the following (Bagrov, 1958),

$$A_K=\sum_{t=1}^{T_0+1} Z(t)\,\varphi_K(t) \qquad (3)$$

Eq. (3) indicates that A_K is unknown becasue it depends on the value of function $Z(t)$ at grid T_0+1. If the approximation (A_K') of A_K could be calculated by using $Z(1)$, $Z(2)$,...$Z(T_0)$, the approximation of $\tilde{Z}_{K_0}(T_0+1)$, $\tilde{Z}'_{K_0}(T_0+1)$, can be obtained from the following,

$$\tilde{Z}'_{K_0}(T_0+1)=\sum_{K=0}^{K_0} A'_K \varphi_K(T_0+1) \qquad (4)$$

Now we discuss the way to get the A'_K.

Assuming the $Z^{(0)}_{K_0}(T_0+1)$ as the initial guess of $Z(T_0+1)$, and replacing

$Z(T_0+1)$ in (3) with $Z^{(0)}_{K_0}(T_0+1)$, the first approximation of A_K is obtained, that is,

$$A_K^{(1)} = \sum_{t=1}^{T_0} Z(t)\varphi_K(t) + Z^{(0)}_{K_0}(T_0+1)\varphi_K(T_0+1) \quad (5)$$

Replacing A_K in (2) with $A_K^{(1)}$, the first approximation of $Z(T_0+1)$ is got,

$$Z^{(1)}_{K_0}(T_0+1) = \sum_{K=0}^{K_0} A_K^{(1)}\varphi_K(T_0+1). \quad (6)$$

Similarly, replacing $Z(T_0+1)$ in (3) by $Z^{(1)}_{K_0}(T_0+1)$, the second approximation of A_K can be calculated,

$$A_K^{(2)} = \sum_{t=1}^{T_0} Z(t)\varphi_K(t) + Z^{(1)}_{K_0}(T_0+1)\varphi_K(T_0+1) \quad (7)$$

Then the second approximation of $Z(T_0+1)$ can be obtained from the Exp.(2),

$$Z^{(2)}_{K_0} = \sum_{K=0}^{K_0} A_K^{(2)}\varphi_K(T_0+1). \quad (8)$$

In general, we have,

$$A_K^{(\nu)} = \sum_{t=1}^{T_0} Z(t)\varphi_K(t) + Z^{(\nu-1)}_{K_0}(T_0+1)\varphi_K(T_0+1) \quad (9)$$

$$Z^{(\nu)}_{K_0}(T_0+1) = \sum_{K=0}^{K_0} A_K^{(\nu)}\varphi_K(T_0+1) \quad (10)$$

$Z^{(\nu)}_{K_0}(T_0+1)$ is the forecast value of Z at the grid T_0+1 which we are looking for.

3. ERROR ESTIMATION

For convenience the following symbols or definitions are introduced.

$$Z = Z(T_0+1), \quad (11)$$

$$\tilde{Z} = \tilde{Z}_{K_0}(T_0+1), \quad (12)$$

$$Z^{(\nu)} = \tilde{Z}^{(\nu)}_{K_0}(T_0+1), \quad (13)$$

$$\varepsilon^{(\nu)} = \varepsilon^{(\nu)}_{K_0}(T_0+1) = Z^{(\nu)} - Z \quad (\nu=0,1,2,\cdots) \quad (14)$$

$$\tilde{\varepsilon} = \tilde{\varepsilon}_{K_0}(T_0+1) = \tilde{Z} - Z, \quad (15)$$

$$\sigma = \sigma_{K_0}(T_0+1) = \sum_{K=0}^{K_0} \varphi_K^2(T_0+1). \quad (16)$$

From the Eqs. (14), (6), (2), (5), (3), (15) and (16), it is obtained,

$$\varepsilon^{(1)} = Z^{(1)} - Z = (Z^{(1)} - \tilde{Z}) + (\tilde{Z} - Z) = \sigma\varepsilon^{(0)} + \tilde{\varepsilon}. \quad (17)$$

Similarly,

$$\varepsilon^{(2)} = \sigma\varepsilon^{(1)} + \tilde{\varepsilon}. \quad (18)$$

Substituting the Exp. (17) in (18), we obtain

$$\varepsilon^{(2)} = \sigma^2\varepsilon^{(0)} + (1+\sigma)\tilde{\varepsilon}. \quad (19)$$

Using the mathematical induction, it can be proved,

$$\varepsilon^{(\nu)} = \sigma^\nu\varepsilon^{(0)} + \sum_{r=0}^{\nu-1}\sigma^r\tilde{\varepsilon}. \quad (20)$$

Introducing symbols,

$$\varepsilon_1^{(\nu)} = \sigma^\nu\varepsilon^{(0)}, \quad (21)$$

$$\varepsilon_2^{(\nu)} = \sum_{r=0}^{\nu-1}\sigma^r\tilde{\varepsilon}, \quad (22)$$

we obtain

$$\varepsilon^{(\nu)} = \varepsilon_1^{(\nu)} + \varepsilon_2^{(\nu)}. \quad (23)$$

It is thus obvious that the prediction error is composed of two parts. One ($\varepsilon_1^{(\nu)}$) depends on the error of initial guess ($\varepsilon^{(0)}$), and the other ($\varepsilon_2^{(\nu)}$) relates to the fitting error ($\tilde{\varepsilon}$). So-called fitting here refers to the $\tilde{Z}_{K_0}(T_0+1)$ obtained from the Exp. (2). When $K_0=T_0$, $\tilde{Z}_{K_0}(T_0+1)=Z(T_0+1)$, $\tilde{\varepsilon}=0$; $\tilde{\varepsilon}\neq 0$ when $K_0<T_0$. Therefore $\tilde{\varepsilon}$ is caused by that truncation order K_0 is less than T_0.

It can be proved

$$\sigma_{K_0}(T_0+1) < 1 \quad \text{for} \quad K_0 < T_0. \quad (24)$$

Using the Exp. (24) it could be obtained from the Eqs. (21), (22) and (23),

$$\lim_{\nu\to\infty} \varepsilon_1^{(\nu)} = 0, \quad (25)$$

$$\lim_{\nu\to\infty} \varepsilon_2^{(\nu)} = \frac{\tilde{\varepsilon}}{1-\sigma}, \quad (26)$$

$$\lim_{\nu\to\infty} \varepsilon^{(\nu)} = \frac{\tilde{\varepsilon}}{1-\sigma}. \quad (27)$$

It means that the iteration is convergent. The error $\varepsilon_1^{(\nu)}$ relating to initial guess convergents to zero, the error $\varepsilon_2^{(\nu)}$ associated with the fitting convergents to $\frac{\tilde{\varepsilon}}{1-\sigma}$ and the total error also convergents to $\frac{\tilde{\varepsilon}}{1-\sigma}$. Owing to that $0<\sigma<1$, the $\varepsilon_1^{(\nu)}$ and $\varepsilon^{(0)}$ are of the same sign and the $\varepsilon_2^{(\nu)}$ and $\tilde{\varepsilon}$ are of the same sign. These two error vary

monotonously with the increasing of ν. The two errors will overlap each other when the sign of $\varepsilon^{(0)}$ is the same as that of $\tilde{\varepsilon}$. They will cancel each other if they have opposite signs, and then the error will be smaller.

4. FACTORS EFFECTING FORECASTING ERROR

4.1 The Error of Initial Guess $\varepsilon^{(0)}$

As shown in preceding section $\varepsilon^{(1)}$ and $\varepsilon^{(2)}$ will cancel each other in a special condition and the total error $\varepsilon^{(\nu)}$ will be smaller in this case. According to this we indicated that such a initial guess is ideal which satisfy the following conditions:

1) $\tilde{\varepsilon}$ and $\varepsilon^{(0)}$ with opposite signs.

2) $|\varepsilon^{(0)}| > \frac{|\tilde{\varepsilon}|}{\sigma}$.

If we use the initial guess of this kind, the forecast error $\varepsilon^{(\nu)}$ will be smaller at a special value ν_c.

The variation of $\varepsilon^{(\nu)}$ with ν is shown in Fig.1

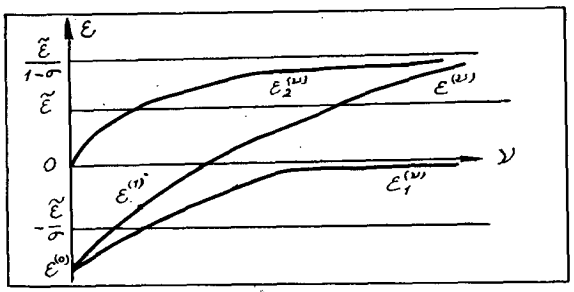


Fig.1 The variation of $\varepsilon^{(\nu)}$ with ν for ideal initial guess

4.2 The Tuncation Order K_0

When K_0 increases, it has two influences upon $\varepsilon^{(\nu)}$. (1) K_0 increases, then $\sigma_{K_0}(T_0+1)$ increases, hence $\varepsilon_1^{(\nu)}$ increases. (2) K_0 increases, $\tilde{\varepsilon}_{K_0}$ decreases, $\sum_{r=0}^{\nu-1}\sigma^r$ increases. Thus $\varepsilon_2^{(\nu)}$ approaches its minimum at certain K_0. To sum up, $\varepsilon^{(\nu)}$ achieves the minimum at a moderate value of K_0.

4.3 The Fitting Error $\tilde{\varepsilon}$

$\tilde{\varepsilon}$ influences $\varepsilon_2^{(\nu)}$ and $\varepsilon^{(\nu)}$. So it is important to get a smaller $\tilde{\varepsilon}$ for improving forecasting. Increase of K_0 can reduce the value of $\tilde{\varepsilon}$. Larger K_0, however, has other shortcomings as indicated previously. Therefore it is necessary to pretreat the time series in order to smooth it and decrease the $\tilde{\varepsilon}$.

4.4 The Eteration Time ν

The relationship between ν and $\varepsilon^{(\nu)}$ is indicated in the Eq. (20). That is, the larger the value of ν, the less the influence of initial guess and the more the influence of fitting.

4.5 The Number of Grid T_0+1

$\sigma_{K_0}(T_0+1)$ and $\tilde{\varepsilon}_{K_0}(T_0+1)$ vary with different T_0. Thus the accuracy of forecasting is also influenced by the selection of T_0.

5. FORECAST FOR MORE POINTS

The forecast for single point T_0+1 has been discused in previous sections. Here we discuss the forecast for more points briefly.

The forecast for more points can be realized by two ways.

(1) First the forecast for the grids T_0+1, T_0+2,..., T_0+n-1 are made and regarded as known. Then forecasting for the grid T_0+n is made by using the formula,

$$Z_{K_0}^{(\nu)}(T_0+n)=\sum_{K=0}^{T_0} A_K^{(\nu)}\varphi_K(T_0+n), \qquad (28)$$

where

$$A_K^{(\nu)}=\sum_{t=1}^{T_0} Z(t)\varphi_k(t)+\sum_{t=T_0+1}^{T_0+n-1} Z_{K_0}^{(\nu)}(t)\varphi_k(t) + Z_{K_0}^{(\nu-1)}(T_0+1)\varphi_k(T_0+1). \qquad (29)$$

(2) Using the data at grids $t=1,2,\ldots,T_0$, the forecasts for T_0+1, T_0+2,..., T_0+n are made simultaneously. The computation formula is

$$Z_{K_0}^{(\nu)}(T_0+\gamma)=\sum_{K=0}^{K_0} A_K^{(\nu)}\varphi_k(T_0+\gamma) \qquad (30)$$

$(\gamma=1,2,\ldots n.)$

where

$$A_k^{(\omega)} = \sum_{t=1}^{T_0} Z(t)\varphi_{kt}(t) + \sum_{t=T_0+1}^{T_0+n} Z_{k\cdot}^{(\omega-1)}(t)\varphi_{k\cdot}(t) \quad (31)$$

The forecast error for the Exps.(28) and (30) may be discussed in similar way.

Owing to that the sample interval is not defined in this method, the forecast for more points may also be made with the method for one point above recounted. For example, the forecast for next year is made with one year interval and the forecast for the year after next is made by using two year interval.

6. OPERATION PERCEDURES

6.1 Treatment of Dependent Data

(1) Pretreating data

Different pretreatment could be used with respect to different predictands. For example we should calculate the percentage of precipitation depature for rainfall to reduce the influence of season. Then the normalization can be realized by using the following Exp.

$$Z' = \frac{Z - Z_{min}}{Z_{Max} - Z_{min}}, \quad (32)$$

There is

$$0 \le Z' \le 1. \quad (33)$$

(2) Determining the number of grid, T_0+1.

(3) Fitting and evaluating $\tilde{\varepsilon}_{K_0}$ for different K_0.

(4) Iterating using different initial guesses for every K_0.

(5) Selecting the ideal initial guess.

(6) Calculating forecasting for dependent data by using the critical iteration times which appeared frequently as the iteration time in fitting.

(7) The final value of iteration may also be used as the forecast, if the error involved is small enough. This can be done without the selection of initial guess and the determination of ν_c.

(8) The computation results for different T_0 and K_0 are aggregated (Zhou, 1975) and final forecasting is obtained.

6.2 Use in the Independent Data

(1) Pretreating data (It is the same with that in the treatment of deppendent data).

(2) Determining the number of grid, T_0+1.

(3) Fitting and evaluating $\tilde{\varepsilon}_{K_0}$ for different K_0.

(4) Determining the requisite type of initial guess by analysing the variation of iteration results. Using the ν_c which is determined based on dependent data as the ν_c in independent data, the prediction is evaluated.

(5) If the final value of iteration in dependent data is used, the final value of iteration in independent data could also be taken as the forecast result.

(6) Aggregating the predicted results.

7. PRACTICAL APPLICATION

The method in this paper has been used in the forecast of annual precipitation trendency over the Xinanjiang valley. Weighted average of the total January-September rainfall is taken as the predictand. The ten stations selected are Tunxi, Yuliang, Sangangkeng, Jiekou, Shangxikou, Yixian, Huangshan, Bailuoban, Jiangjia and Pailing. The data from 1932-1981 are used. All data of percentage of departure are normalized.

Number of grid is taken as 7, 9, 11. Iteration was made for different K_0 in every T_0. The following initial guesses were used: -4, -3.5, ..., -0.5, 0, 0.1, 0.2, ..., 1, 1.5, 2.5, ..., 5. The analyses of result shows that either $Z' \le -1$ or $Z' \ge 2$ must belong to the ideal

initial guess. Therefore, it is enough to judge which of above two kinds of initial guess belongs to the initial guess of this kind. For example the final value of iteration for 1938 is $Z'^{(\infty)}=Z'+\frac{\tilde{\varepsilon}}{1-\sigma}=0.861$. Thus the true value must be between 0.861 and 2 if $Z'=2$ belong to the ideal initial guess. Here $K_0=2$, $\sigma=0.76$. So the error of final value is $\frac{\tilde{\varepsilon}}{0.24}$. This value is in general larger. Thus, true value will be very large and even exceed 1. But it is impossible. Thus it is judged that $Z'\leq -1$ belongs to the ideal initial guess. It can be seen that the transformation of data into Z' not only makes the time series more stable, but the determination of the type of initial guess easier.

Fig.2 illustrates the results for $T_0+1=11$, $\nu=2$, $K_0=1$, The dashed line represents the calculation and the full line denotes the observation. It is clear that the result is better and the main droughts and floods are obtained. By the way this result has not been aggregated.

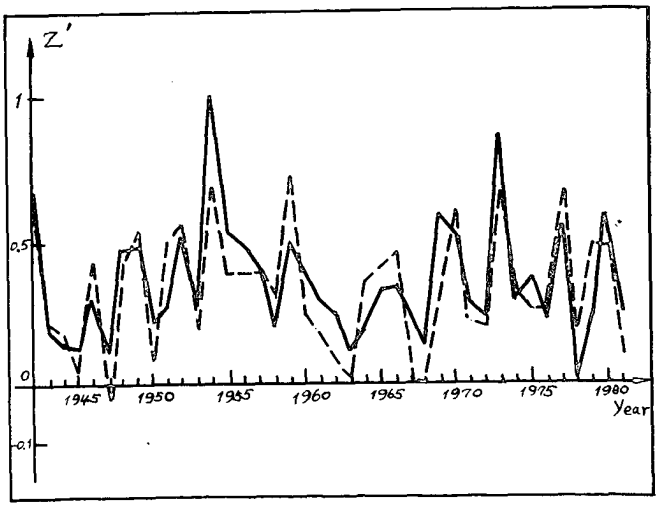


Fig.2 Calculated result of rainfall from Jan. to Sept. over the Xinanjiang valley ($T_0+1=11$, $K_0=1$, $\nu=2$)

———— Observation

-------- Calculation

It is obtained from the Exp. (32)

$$Z=Z_{min}+Z'(Z_{Max}-Z_{min}). \quad (34)$$

For this example we have

$$Z=1.01Z'-0.369\approx Z'-0.369. \quad (35)$$

Therefore the difference between Z and Z' is just a constant and the result of forecast for percentage of precipitation departure could be known from Fig.2

Substituting Z=0 into the Exp.(35) leads to Z'=0.365. Therefore 0.365 corresponds to zero departure.

This method has been used in operational forecast of ten days and monthly rainfall by meteorological group of the Xinanjiang Electric Power Plant in 1982. The forecast of precipitation trerdency for 1983 was provided for the plant by us. Some hydrological and meteorological observatories in China are testing this method. Using this method the precipitation forecasting from April to Jine in Beijing and in rainy season in middle and lower reaches of the Changjing River have been made in spring in 1982 in Lab. 5 of the Institute of Atmospheric Physics, Academia Sinica and provided for the hydrological and meteorological depantments concerned.

8. SUMMARY

The forecast method of time series proposed in this paper is simple and convenient for use. Its operation could be conducted with computer or by hand. Thus this method can be widely tested with computer and may also be used in hydrological and meteorological stations where no any computer is available. There is no limit to the sample inteval of this method, so it may be used in short, medium and lang range forecasting.

9.REFERENCES

Bagrov, N.A., 1958: Analiticheskoe pregstavlenie polei, Trugy TsIP, 64:3-25.

Yang Jianchu, 1953: Long-rang forecasting over one year using the rule of evolution of meteorological element, Acta Meteorologica Sinica, 24, 100-117.

Zhou Jiabin, 1975: On the aggregation of forecast results, Investigations on Statistical Methods of Weather Forecast in Single Station, Collected Papers of the Institute of Atmospheric Physics, Academia Sinica, 3:62-71. Science Press, Beijing.

IMPROVEMENT OF SEASONAL CLIMATIC FORECAST IN HUNGARY

J. Bartholy - M. Kaba

Central Meteorological Institute
H-1525 Budapest, 114, P.O.Box No. 38

Now, in the decade of the World Climate Programme, an increased interest has been shown in all branches of climatology, thus also in climatic forecasting. The users' increasingly complex demands are motivated by complex social and economic problems yet unsolved.

The Hungarian Meteorological Service has been engaged in climatic forecasting activity relating to both its research and operational aspects for several years. In a publication regularly issued each month since 1976, the temperature and precipitation conditions to be expected in Hungary during the following three months are given for eight stations. According to the aims of the climatic forecasts as well as to the possibilities of their approach the prognostic model was elaborated on the basis of a simple but meteorologically and mathematically well established statistical analogue method, the Nearest Neighbour Decision Rule. The results of the forecasts are given in intervals corresponding to 60, 70 and 80 per cent probability levels. In the decision making process of statistical character the selection of analogies is carried out with the help of the surface pressure and the 500 mb contour fields of the Atlantic-European region as well as with that of the long data series /from 1901 to 1980/ of the monthly temperature fields of the entire Northern Hemisphere.

Originally, analogue periods were selected from the archive to match the monthly code series of the Hess-Bresowsky macro-scale circulation classification. However, the limited validity of the different synoptic types in space as well as the subjectivity of coding unfavourably affected our forecasts. A comprehensive study was undertaken, therefore, in order to establish a new and more objective system of synoptic types. In the course of this study the cluster analysis of the 500-mb fields of the Atlantic-European region characterized by 80 grid-points was carried out. The region in question in shown in Fig. 1. In the clustering method well known from mathematical statistics /McQueen's algorithm/ the daily 500-mb fields for the period from 1949 to 1972 were used. For each of the four seasons a separate system of types was established.

The classification broken down into seasons, contains 19 types for spring, 8 for summer, 15 for autumn and 17 for winter.

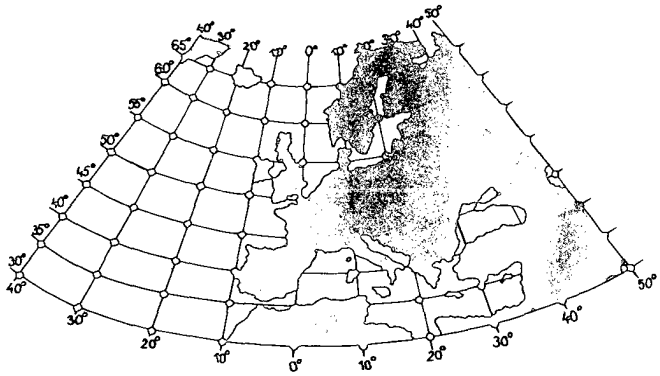


Fig. 1. Grid system covering the investigated area

The similarity indices often used for characterizing a cluster system are the so called external and internal distances. Being suitable for an intercomparison, these indices were used by us, too. Taking the arithmetical mean of all the elements contained in a class for its centre and the distance between the centres for the external distance, the following conclusions can be drawn. The external distances of the new seasonal classification system are on the average 60 to 70 per cent greater as compared with those of the Hess-Brezowsky classification. Consequently, this classification stretches an 80-dimensional space of the 500 mb fields better than the Hess-Brezowsky types to. This is also proven by the fact that the maximum external distances of the system increased by 23 per cent on the average. The increase of the minimum external distances by 69 per cent on the average shows a better separation of the synoptic types.

The internal distance of a class within the system is defined as the mean distance of its elements to its centre. In Fig. 2. the internal distances of the clusters are shown for the four seasons /the numbering of the seasonal clusters corresponding to the succession of growing internal distances/. This figure proves that it has been right to cluster the

seasons separately. The curves of the summer and winter seasons, respectively, are completely separated from each other while those of the in-between-seasons have a nearby course. The significant separation of the summer and winter curves is due not to the differences between the absolute values of the 500-mb fields but rather to the fact that in summer the pressure gradients are much smaller, the standard deviations around the centres of the clusters being, therefore, also smaller.

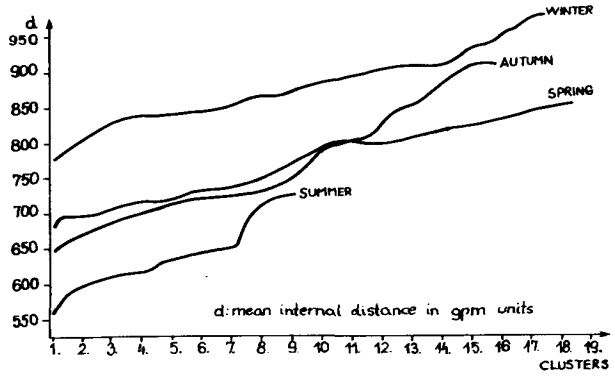
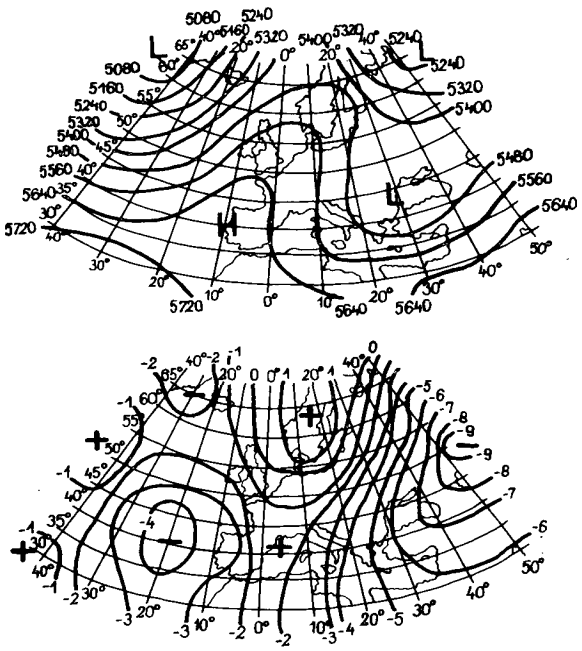


Fig. 2. Inner distances /radii/ of the new classes in different seasons /The clusters have been ordered according to their increasing inner distances; for determining distance Euclid metrics has been used/

Each synoptic type was characterized by the fields of temperature anomalies calculated from the 500/1000-mb fields, the surface pressure and the 500-mb fields. In Fig. 3. all the three fields of weather type No. 9 are shown.



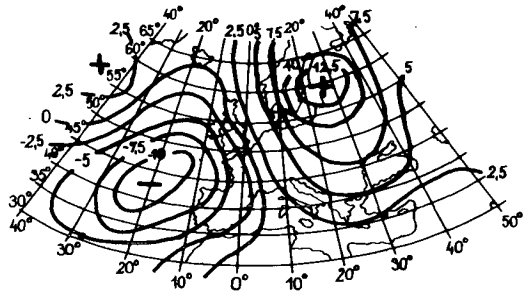


Fig. 3. 9th class of the type system for spring

All the data available for the whole period were catalogued, i.e. on the basis of the 500-mb field of the Atlantic-European region, each day was clustered into the appropriate class of the weather type system. This data series represents the data bank of our analogue method. The new classes as well as the catalogue were tested by a programme package written for an IBM 370/155 computer. By entering them into the above model, operative forecasts for three and six months, respectively, were prepared.

As a firs step in verifying the forecasting method, the change in time of the similarity indices of the periods classified according to the degree of similarity was studied /Fig.4./ In the figure presented, along the axis of abscissa the degree of temporal extrapolation in indicated and along the axis of ordinates: the series of the analogue periods with growing similarity indices. Several interesting conclusions can be drawn, concerning, on the one hand the rather reliable extrapolability of the global analogies /no fast decay was observed after the third month/. Already as a first approximation, this fact permits us to conclude that the temporal validity range of the forecasts can be extended. On the other hand, as far as the optimum number of analogue periods to be used in forecasting is concerned, it can be seen that it is not worth-while to take into account the similarity of the fields received over the 5 to 6 periods showing the highest analogy. A third conclusion relates to the possibility of observing the change of the so called weight factors expressing the degree of importance of the analogue periods used for establishing the forecasted fields or data as a function of time.

Thus, further on, the forecasts of climatic character were prepared using the 5 optimum periods with the highest degree of similarity.

The verification was carried out on the test material of the 10-year period from 1961 to 1970 with the Bayes-type

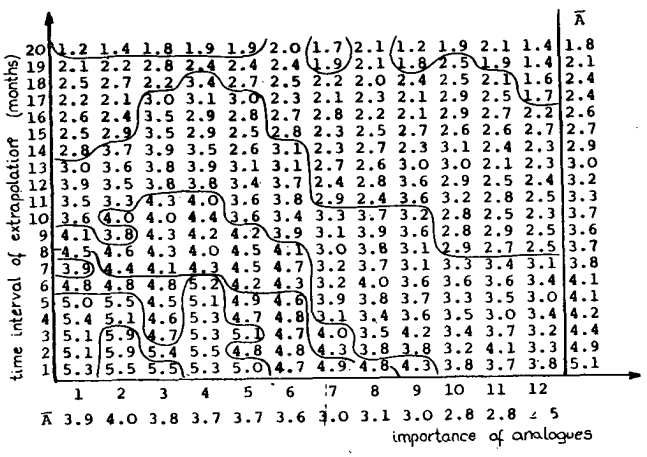


Fig. 4. Changes of weighting factors of the different analogous periods

loss estimation /for absolute error, quadratic loss function and for different intervals/. Evaluating the estimates of the intervals corresponding to 70 per cent probability level it was concluded that the results of the procedure were promising /70 to 78 per cent/. The forecast relating to the second month proved to be the most reliable.

The forecasts of the mean precipitation amounts for the territory of the country was analyzed in a similar way. Our results are modest so far; the estimation by intervals approaches to the 70 per cent verification level. Since September 1982, tentative half-year forecasts have been prepared regularly.

In the second part of the present paper the activity is described which defines the direction of further researches in the field of climatic forecasting.

An attempt has been made to establish a model which could meet the requirement to increase the number of predictors used for climatic forecasts.

After having investigated the principal methodological problems of climatic forecasting it became obvious that the extension of the period of validity would require the elaboration of a model in which the effects of the spheres surrounding the atmosphere - so far neglected - shoul be looked upon as climatic forcing effects /i.e. as predictors/. Among these spheres, primarily the effect of the oceans, i.e.the relation between the atmosphere and the ocean have to be discussed.

As it is known, the duration of the internal processes of the ocean varies from 10^{-2} to 10^{3} years, i.e. it commprises the time intervals being of greatest interest for studying the climate, thus also for climatic forecasting. Since the atmosphere-ocean interactions take place at the boundary of these media and deep-sea processes can be neglected even if the forecasts cover a period of several months, it is sufficient to take into consideration the so-called mixed isotherm upper layer of the ocean. In approaches of that kind it will suffice to take into account the variability of the sea surface temperature /SST/.

The studies in question had a double aim:

1/ The establishment of a complex climatic prognostical model using several atmospheric elements of terrestrial effects for predictors thus permitting to forecast the monthly mean temperature and monthly precipitation amounts for the following six months for the territory of Hungary;

2/ In addition to the analogy selecting model considered as a simple statistical approach, the elaboration of a multiregression model satisfying higher meteorological as well as mathematical-statistical requirements.

Using the factor analysis, the data set of the following predictors was analyzed in order to reveal structural characteristics:

i/ The field of monthly temperature anomalies of the Northern Hemisphere was studied for the period from 1901 to 1960 /22 stations/. The groups representing different "properties" of the field for the winter season is presented in Fig. 5.

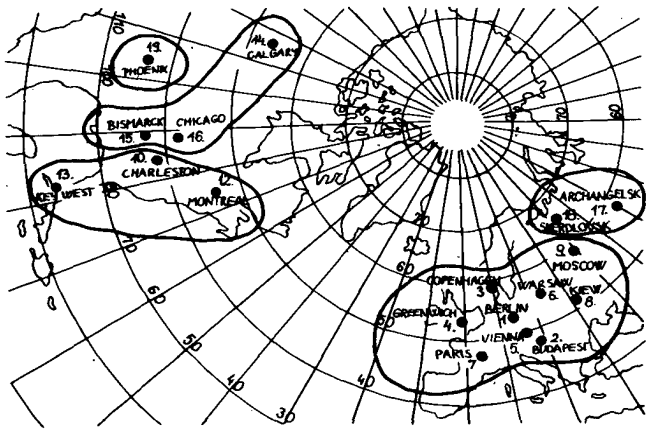


Fig. 5. Sectors found with the factor analysis of the monthly temperature field /winter season/

ii/ The time series of the SST values of the Atlantic Ocean and the Mediterranean from 1949 to 1969 were analyzed on the basis of the data of the section between 90°W and 30°E of the Northern Hemisphere interpolated with 5°-grid spacing. Revealing its internal correlation conditions, the field was devided into 10 sectors /Fig. 6/.

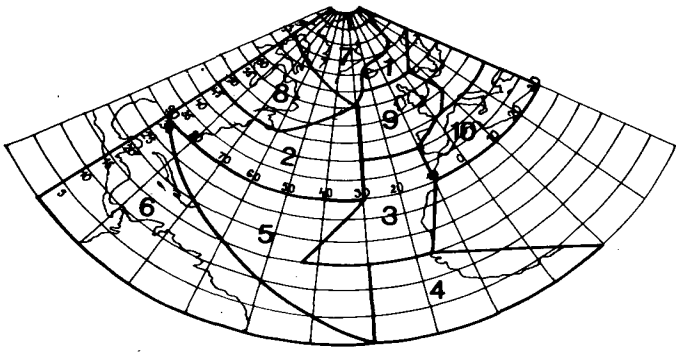


Fig. 6. Sectors of the SST data 1/ North Atlantic Drift /north of 60°N/; 2/ Golf Stream; 3/ Canaries; 4/ Guinea Current; 5/ North Equatorial Current; 6/ South Equatorial Drift; 7/ East Greenland Current; 8/ Labrador Current; 9/ North Atlantic Drift /south of 60°N/; 10/ Effect of the Mediterranean.

iii/ The fields of the monthly precipitation amounts were studied using the data sets of 33 stations from 1901 to 1960. The separation of the field for the winter season can be seen in Fig.7

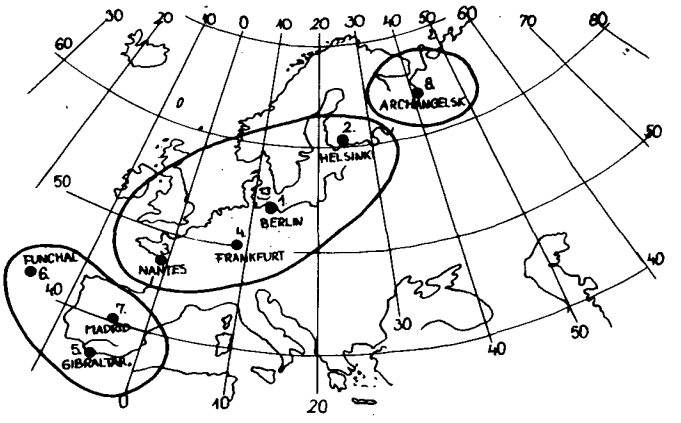


Fig. 7. Sectors of the precipitation field /winter season/

Thereafter we intended to establish the multiregression model using the above predictors so as to be able to predict the monthly mean precipitation amounts and monthly mean temperatures for a half year for each of the 18 catchment sectors for the Carpathian Basin geographically comprizing the territory of our country, i.e. for the representative stations /Fig. 8/./Hungary was represented by station

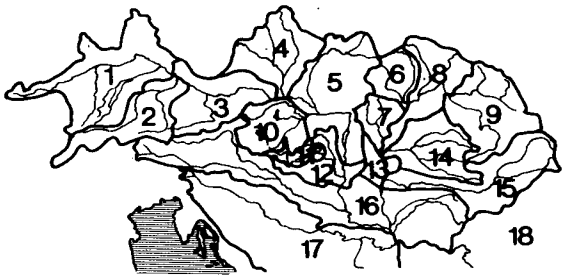


Fig. 8. Catchment areas of the Carpathian Basin

Budapest, but the experiment can be carried out for any other representative station/. On the basis of the main factor method of factor analysis it could be stated which predictors should be used for the forecasting by seasons of the different factors. /As examples the spring-aut mn systems are presented in Figs. 9 and 10./ The multiregression approach

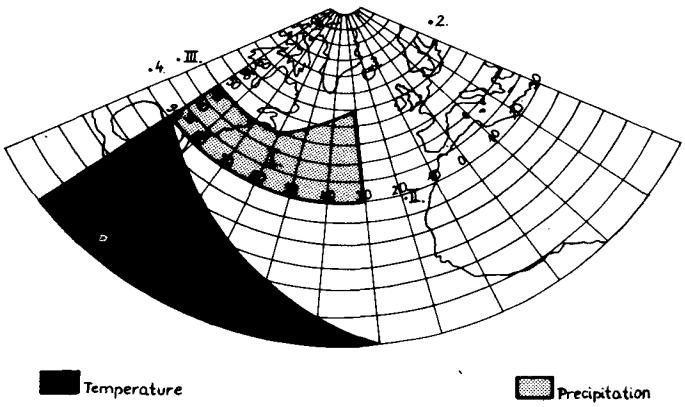


Fig. 9. T { 1/ South Equatorial Current /SST/; 2/ Archangelsk /precip./; 3/ Thorshöfn /precip./; 4/ Phoenix /temp./
P { I. Golf Stream /SST/; II. La Laguna /precip./ III. Phoenix /temp./

separately applied for both parameters by seasons resulted in establishing the regression coefficients. Using them, the verification of the forecasts for the period from 1949 to 1960, studying the relation between the deviations relating

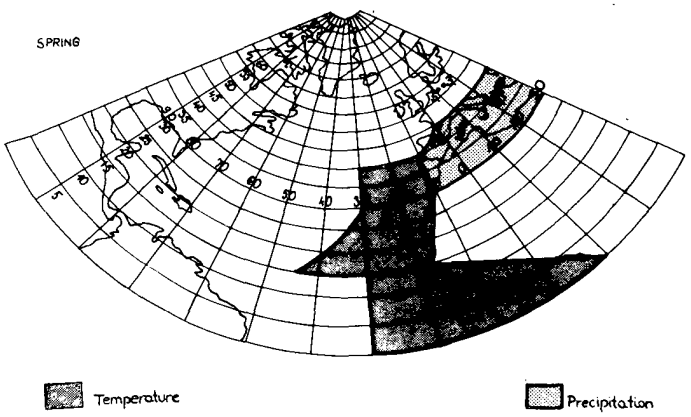


Fig. 10. T {1/ Guinea Current /SST/; 2/ Canarian /SST/; 3/ Warsaw /temp./ 4/ Madrid /precip./; 5/ Berlin /precip./
P {I. Effect of the Mediterranean /SST/; II. Athens /precip./

to the different months and the standard errors of the estimates the results were better than those of a forecast giving the climatological average, in 68 per cent of all the cases of seasonal forecasts and in 63 per cent of all the cases of half-year forecasts. If the researches related to the application of satellite images will give positive results, i.e. if real-time data will be available, the use of the model in operational work seems to be a solved problem.

REFERENCES

Bartholy, J., Gulyás, O., 1980: A method of analysing macrosynoptic types using analogy indices. Acta Climatologica, Szeged, 1980. XVI-XVII. 1-4. 11-17.

Cradock, J.M. 1980: Work in synoptic climatology with a digitized data bank. Met. Mag. Vol. 99. No. 1177. 221-232.

Götz, G., 1982: Some thoughts about the dynamic meteorological interpretation of climate. Időjárás, Vol. 86 1982. No. 5.

Gulyás, O., Bartholy, J., Légrády, G. 1977: The notion of analogy and its use for type formation II. Időjárás, 81. 1977. 6.

Kaba, M., Faragó, T., 1975: Mathematical model of forecasting methods based on the Principle of Analogy. Időjárás, Vol. 79. 1975. No. 3.

Kaba, M. Faragó., T., 1976: Statistical prognosis of Monthly mean temperatures and estimation of Bayes' risk. Időjárás, Vol. 80, 1976. No. 6.

Kaba, M., Bartholy, J., Bán, M., 1980: Long-range precipitation forecast for the territory of Hungary. A study, prepared for the Hungarian Water Authority. Budapest, 1980.

Kaba, M., Bartholy, J., Bán, M., Légrády, G., 1982: Forecasting of the precipitation conditions in the Carpathian Basin. Vizügyi Közlemények, Vol. LXIV., No. 3.

Koppány, Gy., 1975: Long-range extrapolation of analogies /in Hungarian/. Időjárás, Vol. 79. No. 1. 17-27.

Kruizinga, S. 1979: Objective classification of daily 500 mb paat-rn. Bull. of sixth conference on Probability and Statistics in Atmospheric Science of the Merican Meteor. Society. 1979. Oct. Pariff, Alta; Canada.

Mac Queen, J., 1967: Some methods for classification and analysis of multivariate observations. Proc. of the Fifth Berkeley Symp. on Prob. and Stat. F. Univ. of. Calif. Press. Calif. 1967. 281-297.

AN ADAPTIVE METHOD OF PERIODIC ANALYSIS AND ITS APPLICATION TO CLIMATE PREDICTION

Hongxing Cao *

Academy of Meteorological Science
Central Meteorological Bureau
46 Baishiqiaolu, Beijing, China

Zhijun Lü

Department of Mathematics
Shanxi University
Taiyuan, China

1. INTRUDUCTION

Several methods intending to demonstrate periodic phenomena which indeed exist in atmospheric motion have been proposed. The periodogram based on harmonic analysis was used earlier and the power spectral analysis in view of the autocorrelation functions has been widely employed. In recent years, the maximum entropy method (MEM) has been developed to improve the periodic estimation (Burg, 1967; Cao and Lu, 1979). In the meantime a simple but efficient method of periodic estimation was suggested and improved in China by both meteorologists and mathematicians (Tian, 1981; Tan, 1981). It is often called the periodic extrapolation with variance analysis, but in this paper, the phase mean superposition (PMS). We describe a version of this method and present a case to predict the precipitation amount in the Meiyu season over the Middle and Down Yantse River.

2. MATHEMATICAL PRINCIPLE

Consider a time series

$$X(t) : x(1), x(2), \ldots, x(n)$$

and suppose that periods exist in X(t). Dividing X(t) into K groups on the order of the time length $p\ (2 \leq p \leq K)$, a array of them is made up as follows:

phase	1	i	p
	$x(1) \ldots., x((i) \ldots\ldots\ldots\ldots., x(p)$		
	$x(1+p) \ldots, x(i+p) \ldots\ldots\ldots\ldots., x(2p)$		
	$\ldots\ldots\ldots

$$\bar{S}^{(p_1)} \triangleq S^{(p_1)}/p_1 = \max_p s^{(p)}(x)/p \qquad (6)$$

as a indicator and ignore the F-test. Selecting the prevailing period of X(t) corresponding to $S^{(p_1)}$, the new series is made up as

$$\bar{X}_1(t): \bar{x}^{(p_1)}(1), \bar{x}^{(p_1)}(2), \ldots, \bar{x}^{(p_1)}(i), \ldots,$$

$$\bar{x}^{(p_1)}(p_1), \bar{x}^{(p_1)}(1), \ldots, \bar{x}^{(p_1)}(i) \ldots, \bar{x}^{(p_1)}(1) \qquad (7)$$

$$\ldots, \bar{x}^{(p_1)}(m) \qquad (1 \leq m \leq p_1)$$

Further we set up the new series

$$X_1(t) = X(t) - \bar{X}_1(t) \qquad (t = 1, 2, \ldots, n).$$

Repeating the same procedure with (2), (4), (6) and (7) for $X_1(t)$, the new mean series $\bar{X}_2(t)$ is obtained again. We may meanwhile compute a variance reduction

$$\Delta S = [\, S_r(x) - S_{r-1}(x) \,] \,/\, S_{r-1}(x)$$

where r is the order of calculation step. If $\Delta s < \Delta S^*$ (pre-specified), we stop the calculation. Finally the estimation of the series X(t)

$$\hat{X}(t) = \bar{X}_1(t) + \bar{X}_2(t) + \ldots + \bar{X}_R(t)$$

is obtained, where R is the number of the mean series which are to fit the original series X(t). This method is able to adapt the initimate features of the data series under analysis, since a set of mean series are led from the data self. Since the key of this method is to add several mean series, which are obtained from averaging the elements at the same phase, it is called the phase mean superposition.

The periodogram and the power spectral analysis are confined by a specific function, which is pre-prescribed. As in the MEM and the component analysis, the 'base' function of PMS is taken in accordance with the data feature. Using the MEM one can accurately demonstrate the period length, but cannot estimate the values of time series. However, using PMS both period length and series values can be obtained. The latter is useful for practice.

Denoting the residual error square

$$Q_R = \sum_{t=1}^{n} [\, X(t) - \sum_{i=1}^{R} \bar{X}_i(t) \,]^2,$$

if $n > 15$, the convergence of Q_R has been proved (Tian, 1981), i.e. there exists

$$\lim_{R \to \infty} Q_R = 0.$$

According to numerous practical computations the fitting capacity of PMS is generally more efficient than the other periodic analysis, especially for fitting the extreme value.

3. PREDICTING MEIYU

Meiyu is a climate phenomenum appearing over the Middle and Down Yantse River in summer and is caused by the moonson over Southern Asia. During the Meiyu season there is a heavy rain and is persistently overcast. A longe-range forecast of Meiyu is very important for agriculture, irrigation works and water transport.

The data under sutdy are the total precipitation at Shanghai, Nanjing, Wuhu, Jiujiang and Wuhan within the Meiyu season from 1951 to 1980. Three cases with different size of records 1951-1980, 1951-1979 and 1951-1978 are calculated. As an example, the result of case 1951-1980 is given in Tab. 1 and Tab. 2. Four periods: 13, 7, 10 and 8 years are selected in this case. The detailed calculation procedure is shown by the phase mean series, the residual variance and variance contribution percentage in Tab. 1. Comparing the observations and the estimations of case 1951-1980 in Tab. 2, it is found that the extreme values in several years are perfectly fitted. For example, 1954, 1969 and 1980 were a serious flood and 1958, 1965 and 1978 were a terrible drought. All these years are successfully estimated. In general, to fit the extreme value is not done well by a autoregrssive model - even by a ARIMA model (Huang et al., 1980).

In May 1981 we predict Meiyu with the estimation series in Tab. 1 for three years 1981, 1982 and 1983 in advance. The forecast for 1981 agrees well with the observations. In 1982, the rainy tendency in the Meiyu season is also predicted.

It is worth remarking that the extreme record has an impact on the prevailing period. A list of the length of the selected periods for three cases is given in Tab. 3. Both case 1951-1979 and case 1951-1978 demonstrate the same periods, i.e., 5, 7, 3, 13 and 10 years, but case 1951-1980 demonstrate the other set of period length, i.e., 13, 7, 10 and 8 years; especially the first period changes from 5 years to 13 years and the order of the periods is also changed. Naturally it is reduced to the 1980's extreme precipitation 267% in the Meiyu season, because this extreme value leads to the variation of the variance between the groups, $S^{(p)}(x)$, so as to changing the periods selected. From that we may think this is also a robust problem for the PMS.

4. ACKNOWLEDGEMENTS

I wish to thank Prof. Dr. G. Fischer of the Meteorologisches Institut der Universität Hamburg for his help. Thanks are also due Mrs. Zinecker for her typing the manuscript.

Table 1:

order	length (year)	phase mean series						residual variance*	variance contribution (%)
1st period	13	11.6 7.3 -43.2	-62.7 -70.2	24.6 -48.7	109.3 - 42.7	-11.2 -33.7	75.3 43.3	53639	0.625
2nd period	7	5.2 - 9.7	20.7	-47.8	13.9	26.8	-15.5	37239	0.740
3rd period	10	14.3 -36.1	8.1 -29.8	-0.1 4.3	39.1 37.9	-16.1	-21.6	18734	0.869
4th period	8	-26.2 - 1.6	6.5 3.1	12.4	5.1	15.0	-13.9	13401	0.906

* Total variance of 1951-1980 series is 143196

Table 2

Year	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967
Observations*	139	91	93	302	122	126	91	9	36	52	68	149	38	92	5	120	52
Estimation	102.6	70.3	86.8	265.0	112.1	122.0	57.6	6.0	47.9	51.5	104.6	180.9	53.9	124.8	22.4	124.6	96.9

Year	1968	1969	1980	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980		1981	1982	1983
Observations	51	220	119**	46	62	58	60	133	71	97	9	154	267		56	105	
Estimations	77.1	216.5	132.5	47.1	48.4	74.0	58.4	112.6	66.1	70.2	0.6	146.8	251.7	Fore-cast	51.3	198.1	105.5

* Climatical mean is 100, observations is relative percentage
** Mistaken for 129 in calculation

Table 3

order of period case	1	2	3	4	5
1951-1980	13	7	10	8	
1951-1979	5	7	3	13	10
1951-1978	5	7	3	13	10

REFERENCES

Burg, J.P., 1967: Maximum entropy spectral analysis, presented at the 37th Annual Meeting Society Explor. Geophys., Oklahoma City, Okla.

Cao, H.X. and Q.L. Lu, 1979: Spectral analysis of meteorological historical series using maximum entropy method, Kexue Tongbao, Vol. 24, 351-355 (in chinese).

Huang, W.J. et al., 1981: Application of ARIMA seasonal model of time series to long-range weather forecasting, Kexue Tongbao, Vol. 26, 433-438.

Tan, G.R., 1981: Stepwise periodic analysis. Sci. Atm. Sin., Vol. 5, 434-443 (in Chinese).

Tian, C.J., 1981: Convergence of periodic extrapolation with variance analysis, Kexue Tongbao, Vol. 26, 140-142.

STATISTICAL ANALYSIS OF GENERAL CIRCULATION MODEL CLIMATE SIMULATION, SENSITIVITY, AND PREDICTION EXPERIMENTS

Robert E. Livezey

Climate Analysis Center
NMC, NWS, NOAA
Washington, D. C. 20233 USA

1. INTRODUCTION

Models of all kinds are an essential ingredient in the study of climate. Not only is it impossible to set up controlled experiments in the real climate system, the system itself is and will remain inadequately observed. With the advance of computer technology and the recognition of the fact that short-term climate change is the result of a global scale interplay of many low frequency internal and external forces, it was inevitable that these models would become more and more complicated. The most complex of them, general circulation models (GCM's), cover a global domain, are calculated using high horizontal and vertical resolution, incorporate sophisticated physical parameterizations, and require powerful computers to integrate for any meaningful length of time. During the last decade reliance on GCMs to study short-term (up to inter-decadal) climate fluctuations has steadily increased.

In their most common application, the "sensitivity" test, these models are typically integrated twice, with and without a specified anomalous external influence - for example an area of abnormal sea surface temperatures. The climatic statistics of the two runs are then compared and a decision made about whether the null hypothesis that the statistics are the same can be rejected at some prescribed level of confidence; i.e. whether it can be confidently inferred that the model's climate "changed".

More recently, studies have also addressed the question of whether the climate of two GCM's or a GCM and the real world are the same or significantly different, the "simulation" experiment. A third type of investigation, the "predictability" experiment, has sought to discover at what point the time-mean state (30-day for example) of long runs from initial conditions perturbed with simulated random observational errors start to differ significantly from unperturbed control runs. An offspring of this genre of study is the "prediction" experiment in which the evolution of time averages in runs started from analyzed real-data initial states are compared to what actually occurred in nature.

Collectively, evaluation of the GCM experiments presents one of the most challenging statistical inference problems in the study of climate. The problem is similar to (but even more intractable than) that encountered in empirical studies of global climate. This is because the data sets consist of small samples of large numbers of cross-correlated climate statistics. Samples are small because individual model runs require considerable computer resources. Each run, however, generates an enormous amount of information about the climate state all the way from grid-size to planetary scales. The typical experimental situation places a large burden on even the most sophisticated multivariate parametric test techniques.

Nevertheless, it is essential that effective test methodologies be worked out. Computer power will continue to be a severely limiting factor in experimental design at many research centers around the world for some time. Meanwhile the need to develop detailed scenarios of climate change continues to grow in importance, both in terms of what changes society can expect, and of how to detect quickly those changes to which society is sensitive. This is especially crucial for CO_2-related changes.

Bell (1983) has shown how to average data intelligently to optimize a test's sensitivity to a hypothesized climate scenario and Epstein (1982) has demonstrated how likelihood ratios can be used to define odds of not detecting a hypothesized climate change that actually is in progress or of incorrectly identifying one that actually is not. These approaches will remain ineffective as long as the hypotheses are ill-formed. The best tool to remedy this weakness is the GCM, as long as it is honed by careful experimental design and decisive experimental evaluation.

In the next section, the test problem will be described in more detail. Then in Section 3 the development of ways to deal with the problem will be traced from the first univariate/parametric techniques to the most recent multivariate and permutation/Monte Carlo approaches. Recent examples of the different kinds of experiments

described above will be presented in Section 4. The paper will conclude with summary remarks in Section 5.

2. THE PROBLEMS

It is a simple consequence of sampling theory that even with an unchanging "climate" a finite time mean (for example the January mean temperature at a location) will vary from one year to the next. This is the result of the nearly stochastic nature of daily weather fluctuations and the finite length of sampling periods. This "natural variability" or "climate noise" is present in both the real climate system and in sophisticated GCM's and implies a degree of uncertainty in estimates of climate statistics from finite time series (c.f. Madden, 1981). It is this uncertainty that makes the evaluation of GCM climate experiments so difficult.

If model dynamics are transitive or ergodic it is theoretically possible, in the January mean temperature case, to average a very large number of model Januaries to reduce the estimation error of the mean to inconsequential levels. This averaging would permit direct conclusions to be drawn in experiments like those described in Section 1 from simple examination of the differences in model runs. Unfortunately, the number of runs necessary to use this approach coupled with the expense of producing just one run preclude its further consideration.

In practice generally only a few runs are available for an experiment, and estimated errors will be quite large (Leith, 1973). The implications of this limitation will now be explored and in the next section the manner in which researchers have dealt with these implications will be examined. It should be remembered throughout that all of the problems in experimental evaluation addressed below can be partially ameliorated by increasing sample sizes through additional model runs.

First, consider the simple univariate problem of testing whether one mean is significantly different from another. Chervin and Schneider (1976b) suggest as a test statistic the Student's t, approximated by $r=|\Delta|/\sigma$, where Δ is the difference in means and σ an estimate of the standard deviation of the mean - the climate noise. Through simple examples they show that, given a preset level of significance, the ability of this test to reject correctly a null hypothesis decreases rapidly for any r as the sample used to estimate either the numerator or denominator decreases.

While Chervin and Schneider's study is suggestive, the evaluation of a GCM experiment is truly a multivariate problem. In that problem, it also turns out that sample size places severe limits on the tests, by constraining their dimensionality and controlling in what detail the tremendous amount of information in a GCM climate experiment outcome can be effectively evaluated.

Before these effects are described, the difficulties of interpreting the results of performing many separate univariate tests (for example, at individual grid points on a grid) need to be thoroughly understood. Madden and Julian (1971) and Reynolds (1978), among others, have pointed out the simple fact that if you perform a number of significance tests the odds of passing one increase rapidly over those for a single test. For instance, 5% of a number of 95% significance tests will be expected to pass even with the null hypothesis true. The situation is actually more complicated than this in two respects.

First, there can be a rather large probability that substantially more then 5% of the tests will be passed in the face of the null hypothesis. For example, binomial odds are 1 in 5 that, given 30 tests, at least 3 or (10%) will pass at the 95% level. This tilt of the odds can be evaluated exactly if individual tests are independent, which is rarely the case. One either can search for a test outcome on the grid that passes at a significance level (say 99.99%) that would be unlikely to occur anywhere on the grid (Madden and Julian, 1971), or one can count up the number of tests passed at some fixed level and use the binomial distribution to define the odds of achieving or exceeding that particular outcome (Livezey and Chen, 1983).

Because there usually is dependence between the tests the former procedure becomes overconservative and the latter overliberal. This can be understood in the framework of binomial arguments using an effective number of independent tests. In terms of percent of passed tests the distribution of Thus, if the probability of passing 3 or more independent tests (out of 30) by accident at the 95% level is .20, the probability of passing 3 or more dependent tests by chance will be larger - in some cases substantially so.

This phenomenon has been shown formally by Hasselmann (1979) to reduce the sensitivity of tests to a given climate signal at a given spatial scale as the spacing of grid points is decreased. To see this, imagine a stationary climate signal with some characteristic spatial scale. When the signal is sampled at just one point near its apex, it can be detected confidently (at the 95% level) if its amplitude exceeds the 95% noise level. If, however, it is sampled instead at a finite network of points near its apex an additional play of odds will be present. Covariance between points will be less than perfect because there will be uncorrelated noise present at all scales (and consequently at all grid points). Thus the probability of detecting a given signal amplitude by accident will be higher than before, so only a stronger signal can be confidently observed. In terms of Livezey and Chen's (1983) arguments, the number of points individually significant at the 95% level needed to detect a signal confidently on a fixed grid will descend toward a minimum defined by the binomial distribution as the signal's scale decreases to the grid spacing or

smaller. In the limit, tests at all grid points will be effectively independent. As Hasselmann states:

> ...the inclusion of additional information which adds to the noise without signficantly improving the definition of the signal deteriorates the signal-to-noise ratio and makes the signal more difficult to detect.

A final problem related to evaluation of collections of univariate tests following the methods of Livezey and Chen (1983) is that only global decisions about experiments are possible. There is no way to discriminate between parts of the experimental outcome that pass the significance test because of climate change and those that pass by chance. Only local tests that passed at the very high levels of Madden and Julian (1971) would escape this dilemma. Local decisions, then, are generally not straightforward.

Multivariate techniques are designed both to deal with this problem and also to account directly for signal structure and scale. Unfortunately, these methods are severely limited in small sample situations with the practical effect that little signal can be resolved anyway without considerable a priori insight into what one is searching for. These limitations are explictly or implicitly inherent in all multivariate approaches. For example, with the T^2 test that Hayashi (1982) advocates the dimensionality of the test is limited, in the case where the true covariances are known and unequal, to less than the sample size of the experiments being tested against each other. Rarely, however, are the covariance matrices "known", and their estimation errors reduce permissable dimensionality even further. Chervin (1981) cites Kendall (1975), who recommends as a practical limit that dimensionality of the T^2 test be no greater than one tenth of the sample size used to estimate the covariances. Tests like Hasselmann's that use the ρ^2 statistic do not have an explicit constraint on dimensionality, but nevertheless require estimation of covariances.

Even if covariances can be independently well estimated and a ρ^2 test employed, the reduction in sensitivity with higher resolution that was described earlier, coupled with the high resolution of GCM climate experiments, requires drastic a priori reduction in test dimensionality. This can be achieved by judicious combination of theoretical argument, examination of real world data, independent experiments, and filtering of model output. Whatever the approach, it needs to be done a priori. After-the-fact reduction in dimensionality, such as that by Hayashi (1982) for the experiments of Manabe et al. (1981), seriously undermines the validity of the tests employed and the credibility of the conclusions reached.

The manner in which the problems described above came to the attention of the climate research community and the solutions that members have proposed will now be briefly traced.

3. DEVELOPMENT OF SOLUTION METHODOLOGIES

In this section the development of univariate and multivariate techniques will be separately discussed, as will parametric and permutation approaches.

3.1 Parametric Methods

The traditional approach to statistical inference in GCM climate experiments has been to formulate test statistics, with given degrees of freedom, that approximately or asymptotically fit classical distributions of known properties. The obvious shortcomings of these approaches occur either when test conditions are not approximately or asymptotically close enough to those required, or when the degrees of freedom are not accurately determined. Despite their shortcomings, these methods have been of considerable value in focusing attention on the relevant issues and in providing a base for less limited techniques.

3.1.1 Univariate Methods

Following Leith's (1973) discussion of the error in estimating climate means from finite times series, the first serious attempt of which this author is aware to evaluate objectively the outcome of a GCM climate sensitivity experiment through the use of model noise estimates was made by Shukla (1975)[1]. His interest was the influence of a cold sea surface temperature anomaly (SSTA) near the Somali coast and over the western Arabian Sea on monsoon rainfall over India in particular, but also over the rest of the globe. To study this he compared the differences in 48-day rainfall with and without the SSTA to estimates of model noise. The latter were obtained using the autocorrelation function estimated from a single model run. Signal to noise ratios greater than one were taken to be significant changes, but no attempt was made to define confidence levels or place the results in a probabilistic framework.

That step was taken by Chervin and his collaborators in a series of papers that culminated in the development of a method for determining the statistical significance of differences in mean responses in sensitivity experiments and its application to an actual experiment (Chervin et al., 1974; Chervin and Schneider, 1976a, 1976b; Chervin et al., 1976). Their formulation, the so-called Chervin-Schneider algorithm, depends on formation of a signal to noise ratio quite similar to that used by Shukla (1975). Climatic noise - that is, the variability of the time mean - was estimated directly from the combined variance of time means from both the "prescribed change" experiment and a number of control experiments, rather than from the auto-

---

1 Before Leith, Warshaw and Rapp (1973) used analysis of variance to evaluate the statistical significance of changes in the 30-day mean zonal GCM climate resulting from removal of polar ice.

correlation function of a single experiment. Consequently, the resulting ratio is proportional to chi-squared, permitting the use of a t-test. A large portion of subsequent univariate work has consisted of either applications, refinement, or extensions of this technique.

For example, Chervin and Schneider (1976b) required that the population variances of control and prescribed change experiments be the same. Along with a recapitulation of a number of other testing considerations, Laurmann and Gates (1977) pointed out that this assumption could be relaxed with the use of the Behrens-Fisher statistic along with multiple prescribed change runs (Chervin and Schneider used only one) in addition to the multiple control runs. Chervin (1980) himself was able to relax this constraint by relating the Chervin-Schneider t-statistic to the Cochran and Cox approximation. Additionally, he established the feasibility of using the F-statistic to test whether differences in the population variances actually were present. This paper was innovative in one other regard: The model comparison did not involve a sensitivity test, but instead juxtaposed the model's January and July climates.

More recently, Katz (1982) reopened the question of whether the climate noise can be estimated from the detailed time series, rather than from a sample of time averages, and then used to form a statistic with a known distribution. He correctly pointed out that the Chervin-Schneider approach does not use all of the information available. He showed that if the model runs are long enough (greater than 50 days or so) and well-enough behaved (statistics stationary and Gaussian), time-series modeling techniques can be employed to estimate a σ that can be used to form a statistic similar to Chervin and Schneider's t, but Gaussian in its distribution. Katz's statistic is more sensitive to a climate signal than one of equal magnitude with a t-distribution. The spectral analysis approach to estimation of noise levels suggested by Madden and Ramanathan (1980) is equivalent to Katz's method but is slightly more subjective. Katz (1983) has also addressed the problem of test methodology for discontinuous, non-Gaussian series, such as daily precipitation.

The principal contribution to the univariate problem of the paper by Hayashi (1982) is the notion that the confidence interval defined by the test distribution can sometimes be used to decide if the difference between two experiments is significantly small if not significantly large.

3.1.2 Multivariate methods

Until Hasselmann's (1979) elegant and thorough treatment, the multivariate problem was largely ignored. Since then Chervin (1981) and Katz (1982) have recognized the problem's existence, Storch (1982b) and Livezey and Chen (1983) have illustrated in practical terms what it can mean in an experimental situation, and - in addition to Hasselmann - Storch (1982a) and Hayashi (1982) have attempted to deal with the problem.

Hasselmann employs the statistic

$$\rho^2(\delta\phi) = \sum_{ij} C^{-1}_{ij}\, \delta\phi_i\, \delta\phi_j$$

where $\delta\phi = \phi - \overline{\phi}$, the difference between the perturbed run mean and the control run mean, the subscripts refer to n different gridpoints, wavenumbers, etc., and $C_{ij} = \langle \delta\phi_i \delta\phi_j \rangle$ is the perturbation covariance matrix. This statistic has a χ^2 distribution with n degrees of freedom. To reduce the dimensionality of the problem to acceptable levels in line with the discussion in Section 2, Hasselmann suggests the *a priori* choice of representation of the experimental outcomes $\delta\phi_i$ by their set of empirical orthogonal functions (EOF's). This is the set of basis vectors that most, in sequence, efficiently reduces the remaining differences between the truncated representation and the complete outcome. This set acts as a filter that concentrates most of what most likely is signal into a relatively few ordered vectors whose incremental significance can be tested by the use of ρ^2 and by a hierarchy of ordered "guesses" or truncated representations. If the coefficients of these basis vectors are denoted by a_1, a_2, etc., and the vector guesses by $\underset{\sim}{g}_1$, $\underset{\sim}{g}_2$, etc., then the response $\underset{\sim}{g}_1 = (a_1)$ is tested first, $\underset{\sim}{g} = (a_1, a_2)$ next, and so forth until the multivariate test fails. Other approaches for constructing guesses suggested by Hasselmann include the solutions of other climate models, particularly simpler ones, or a spherical harmonic decomposition and ordering by harmonics of the experimental solution. Obviously insight obtained from observation of the climate system could provide other guesses. Hasselmann states that, based on the experience of constructing prediction models from data, the dimensionality of multivariate tests will be limited to between 5 and 20. This author believes that without much larger samples than have been available for experiments to date this range is much too optimistic.

Hayashi (1982) offers an approximation to a T^2 test as an alternative to the approach of Hasselmann. He provides little insight into the problem of reduction in test dimensionality, because his reduction is performed *a posteriori*. His subsequent use of a liberal *a priori* 90% acceptance level and the lack of realism inherent in the tested climate simulation (the comparison being between climates with current and quadrupled CO_2, cf. Manabe *et al.*, 1981) provide little basis for using the results to refine scenarios or hypotheses of the results of an increase in atmospheric CO_2 concentrations. These comments are offered to underline the need for proper *a priori* experiment and test design.

3.2 Permutation techniques

The principal difficulty in the use of collections of univariate tests is the uncertainty in the number of independent tests. Multivariate

schemes do not share this difficulty but, for small samples, are severely limited in their dimensionality by their insensitivity, uncertainties in estimating covariance matrices and, in some cases, their explicit form. Both approaches are handicapped by the necessity to formulate test statistics that have known distributions. Often it is difficult to justify the various assumptions of normality, equal population statistics, asymptotic behavior, etc. required for a rigorous formulation.

These problems, with the possible exception of the limits on multivariate dimensionality, frequently can be bypassed by the application of permutation or Monte Carlo techniques. Even the severe limits on dimensionality can sometimes be ameliorated. The main difficulty in using permutation methods is their cost in computer time. Nevertheless, this cost will normally be orders of magnitude less than that required for additional GCM runs. Certain test situations, moreover, will remain intractable without them.

Basically, the terms permutation or Monte Carlo techniques refer here to the general strategy of formulating a test so that the probability distribution of possible joint outcomes (between for example control and prescribed change runs) under a null hypothesis can be directly estimated with sufficient accuracy by permutations or combinations of relatively small samples of separate sets of outcomes.

For instance, using an example from Livezey and Chen (1983), suppose the 30-day mean circulation in a single prescribed-change experiment is to be compared to that from a six-case mean control. Suppose further that nine more non-prescribed cases are available. This makes fifteen cases from which 30-day means sampled from the same population (or climate) can be computed. These can be generated from a very long single run sampled at intervals widely enough separated to ensure independence, or from multiple runs started with different initial conditions but integrated with the same fixed external forcing.

From this sample of fifteen, seven at a time can be randomly selected and six of these can be compared to the seventh. Because there exist more than 40,000 different comparisons, the probability distributions of a number of *a priori* univariate or multivariate hypotheses can be accurately reconstructed, and tests performed directly against these distributions. There is no need to estimate distribution parameters or degrees of freedom. These are merely descriptors of a distribution that is already reasonably well-known, often in more detail than the parameters could reveal. Thus, this and other permutation methods are *non-parametric* in nature.

There are many different ways these tests can be set up depending on the experimental configuration and test goals. Often their application is limited only by the imagination of the researcher. The recent papers by Mielke *et al.* (1981), Livezey and Chen (1983), and Preisendorfer and Barnett (1983) contain a number of ideas suitable for a wide variety of situations as well as some general design considerations.

4. EXAMPLES OF RECENT APPLICATIONS

The techniques summarized in the previous section have been applied principally to sensitivity or prescribed change GCM climate experiments. Of these, most have been studies of the response of models to SSTAs and have used ancestors or descendants of the Chervin-Schneider univariate algorithm. Reference has already been made to the papers by Shukla (1975) and Chervin *et al.* (1976). Others that share these characteristics include Kutzbach *et al.* (1977), Washington *et al.* (1977), Julian and Chervin (1978), Schneider *et al.* (1978), Chervin *et al.* (1980) and Keshavamurty (1982). Prescribed changes other than SSTAs have included the replacement of "standard" boundary conditions with 18000 B.P. conditions to simulate an ice age climate (Manabe and Hahn, 1977), the addition of 10 ppb of chlorofluoromethanes to a January model atmosphere to study its impact on the global radiative balance (Dickinson and Chervin, 1979), and the addition of 90 Wm^{-2} of simulated thermal pollution in a large-scale control area in both January and July model atmospheres to study local and "teleconnected" climate changes (Washington and Chervin, 1979).

More recently, researchers have tried to apply statistical methods to classes of experiments other than the prescribed change type. Chervin's (1980) study of the differences in modeled Januaries and Julies mentioned earlier is an example of a climate "simulation" experiment. The idea in these studies is to test the ability of the GCM to reproduce the actual climate. Chervin (1981) has also developed univariate methodologies for evaluating comparisons between observed and modeled first- and second-moment climate statistics, and has applied them to an analysis of the January vertically averaged transient heat flux.

Multivariate techniques have also been applied to the question of model climate verisimilitude. In fact, the study by Storch (1982a) may have been the first to use the multivariate test framework for any type of GCM experiment. In it Storch partially followed the suggestions of Hasselmann (1979). A null hypothesis that the model January climate was the same as the observed January climate was first formulated. The representation used to describe both model and real climates was a set of 30 EOF's of 500 mb heights produced by others, while the statistic chosen for the test was ρ^2 for the departures of daily mean EOF amplitudes from their climate means. The covariance matrix and climate estimates for this statistic were formed from eleven *observed* Januarys. Rather than test the significance of a hierarchy of guesses built by successive addition of the next higher-order EOF to the previous guess (as suggested by Hasselmann) Storch simply took

advantage of a defective preliminary experiment to identify seven unusually sensitive EOF amplitudes to examine in a second run. This latter, a perpetual January run, was integrated for 75 days.

On 15 (20%) of the days ρ^2 exceeded its 95% confidence level. The 75-day model time series probably contained approximately ten independent events (cf. Paegle and Haslam's, 1982, estimates of the characteristic time between independent 500 mb winter height samples). Therefore, from binomial arguments his result probably had about a one in ten chance of happening by accident (cf. Livezey and Chen, 1983), so it appears likely that the model climate was not the same as that observed.

As a last step Storch performed univariate tests on each of the EOF amplitudes separately to help identify the source of the large deviations from observed climate. One or the other of two EOFs that were associated with zonal wavenumbers five failed the univariate tests in four out of five of the intervals in which the multivariate test failed, suggesting strongly that this scale was the principal cause of the unrealistic model behavior. Tests like these can not only usefully validate a climate model, but can also sometimes provide important diagnostic feedback to the modeler.

The last GCM experiment that will be cited here was the first of its kind also, both in terms of the type of experiment and the application of statistical tests to such an experiment. This is the study of the dynamical predictability of monthly means by Shukla (1981). It consisted of a set of 60-day integrations of a GCM from three different January 1 initial conditions. For each January/February case a control and at least one additional perturbed run (started from the control initial conditions with random error added) were produced.

From this data Shukla formed a statistic at different stages in the runs similar to F, proportional to the ratio of two measures of variability. The numerator measured the variability among monthly means following substantially different initial conditions and the denominator measured the variability among monthly means for randomly, slightly perturbed initial conditions. Shukla argued that the monthly mean remains predictable at least as long as the numerator remains significantly larger than the denominator. He used the F-distribution to make this test for 30-day mean 500 mb heights at each point on a Northern Hemisphere grid. This set of univariate tests were performed separately for days 1-30, 8-37, 15-44, 22-51, and 29-58, respectively, of his multiple GCM integrations. His results suggest that there may be theoretical predictability for the monthly mean out to days 15-44, but there are a number of problems associated with his test design.

First, a global test was not performed and, because the covariance matrix is unknown, it is not possible to assert at what point the experimental significance drops below the 95% or any other level. Second, Hayashi (personal communication) has suggested that Shukla's statistic does not have an F-distribution. If this is correct, it would invalidate the univariate tests. Finally, Hayashi has also correctly pointed out that failure to pass Shukla's test does not reliably preclude the possibility of predictability unless confidence limits are also narrow.

Shukla's experiment does, however, contain the raw material for tests without most of these difficulties and joint preparations to try such tests are currently being made by this author and Shukla. Global tests that use a set of 700 mb height statistics closely related to anomaly correlations[2] can be conducted for a number of different time means (5-, 10-, 15-, and 30-day). Predictability for each time-average can be examined by noting at what point in the GCM integrations the appropriate test statistic can no longer be confidently considered to be larger than that computed from random mismatches of cases drawn from the population.

Even though this proposed experiment deals with a theoretical question, it has a decidedly pragmatic cast. The choices of statistics place emphasis on the characteristic of a height anomaly field most important to extended and long-range forecasters - the positions or phases of anomaly centers. The test itself addresses the following question: When does it become difficult to distinguish the mean forecast outcome of the GCM system from that of a simple random analog selection approach? The latter refers to the random choice of a "forecast" January/February from the population - or, in practice, the sample of observed past January/Februaries. If at some point in a run the chance of making better forecasts by the unsophisticated, no-skill approach exceeds some predetermined amount, say one out of four, then a basis exists for a management decision to terminate GCM forecasts at that point.

For this experiment a fourth independent pair of January 1 control and perturbation runs will be available, so test statistics can consist of four-case means. Their probability distributions can be estimated by Monte Carlo methods from over 30 years of Northern Hemisphere observed daily 700 mb heights available at the Climate Analysis Center. This procedure will tend to overestimate predictability if the model's climate noise is smaller than that of the atmosphere. However, the error of overestimation is believed to be small. Whatever the case, upper bounds for predictability can be defined in the context of forecasting practice, something to which little attention has been paid in previous dynamical predictability studies.

---

2 The test statistics can be similar to those designed by Preisendorfer and Mobley (1982).

In fact, the experimental design and evaluation procedures recommended here for a reexamination of Shukla's predictability experiments can be even more cleanly applied to assess prediction experiments. Then the question is not what the skill might be but what the skill is! The use of an observed data base to estimate probability distributions no longer constitutes an approximation but is the correct approach. Recent qualified success with a GCM in reproducing the 30-day mean atmosphere in certain Januaries from their respective January 1 initial conditions (Miyakoda et al., 1983) mandate that consideration of this or another suitable evaluation technique become a first step in the design of future prediction experiments.

5. CONCLUDING REMARKS

In the preceding sections an attempt has been made to trace the rapid development (over just one decade) of the specialized, but increasingly important, application of statistical testing to GCM climate studies. The characteristic that sets this testing problem apart from others is its particular difficulty; the test sample is small, but multivariate of large dimensionality with a considerable amount of superimposed noise. Thus a multivariate test framework will be severely limited in its sensitivity to climate signals without considerable a priori reduction in the number of tested variables, that is in the degree of specificity or amount of detail in the tested hypothesis.

A variety of strategies have been recommended to accomplish this reduction. These include spatial averaging of experimental results, decomposition into orthogonal bases and subsequent reordering into a hierarchy of guesses, sequential but independent integrations of models with refinement of guesses in each successive integration, and exploitation of empirical experience to suggest possible outcomes. Clearly a successful but parsimonious reduction of dimensionality, and subsequent effective application of the parametric and non-parametric (permutation) techniques outlined in Section 3, may well require the combined talents of modeler, empiricist, and statistician in the process of experimental design. Without careful thought to the considerations outlined here, substantive climate signals in GCM experiments often will not be credibly detected, or - worse - sampling fluctuations will be misrepresented as signals.

In regard to the former - failure to credibly detect a signal - the GCM study by Rind (1982) is instructive. In this potentially very important work, Rind examined the role that late-spring/early-summer ground dryness over the United States plays in the initiation or maintenance of heat waves and drought. Many of the results of this experiment are suggestive, but nonetheless inconclusive. This is because their statistical treatment is incomplete, lacking either global univariate or multivariate analyses. For the more serious problem - misrepresentation of noise as signal - a scenario associated with the doubling of CO_2 concentrations in the next century provides an example. This is the idea that there will be major shifts in growing season rain belts that will lead to disruption of agricultural patterns over North America. While the idea has been the object of a good deal of concern, it is largely based on empirical and model-related evidence that has not held up well under critical post-examinations by others than those who presented the evidence.[3]

Both examples highlight the continued neglect in many studies of the effects of climate noise on observing or predicting climate signals. These are effects that have been thoroughly described by Leith (1973), Hasselmann (1979), and Madden (1981), while numerous examples have been presented by others. Therefore, it seems reasonable that such effects should be routinely accounted for in future experiments in an objective, quantitative fashion.

ACKNOWLEDGMENTS

Discussions with Bob Chervin kept the number of overlooked papers in this version to a (hopefully) small number. Don Gilman and Tony Barnston invested an afternoon each in correcting, clarifying, and polishing the manuscript.

REFERENCES

Bell, T. L., 1982: Optimal weighting of data to detect climate change: Application to the carbon dioxide problem. J. Geophys. Res., 87, 11161-11170.

Chervin, R.M., 1980: On the simulation of climate and climate change with general circulation models. J. Atmos. Sci., 37, 1903-1913.

_____, 1981: On the comparison of observed and GCM simulated climate ensembles. J. Atmos. Sci., 38, 885-901.

_____, and S. H. Schneider, 1976a: A study of the response of NCAR GCM climatological statistics to random perturbations: Estimated noise levels. J. Atmos. Sci., 33, 391-404.

_____, and S. H. Schneider, 1976b: On determining the statistical significance of climate experiments with general circulation models. J. Atmos. Sci., 33, 405-412.

_____, W. L. Gates, and S. H. Schneider, 1974: The effect of time averaging on the noise level of climatological statistics generated by atmospheric general circulation models. J. Atmos. Sci., 31, 2216-2219.

_____, W. M. Washington, and S. H. Schneider, 1976: Testing the significance of the response of the NCAR general circulation model to North Pacific ocean surface temperature anomalies. J. Atmos. Sci., 33, 413-423.

---

3 See the remarks in Section 3 on Hayashi's (1982) treatment of the experiments of Manabe et al. (1981). Also refer to Livezey and Chen's (1983) critique of the paper by Williams (1980). Their comments apply equally well to the study of Jäger and Kellogg (1983).

_____, J. E. Kutzbach, D. D. Houghton, and R. G. Gallimore, 1980: Response of the NCAR general circulation model to prescribed changes in ocean surface temperature. Part II: Midlatitude and subtropical changes. J. Atmos. Sci., 37, 308-332.

Dickinson, R. E., and R. M. Chervin, 1979: Sensitivity of a general circulation model to changes in infrared cooling due to chlorofluoromethanes with and without prescribed zonal ocean surface temperature change. J. Atmos. Sci., 36, 2304-2319.

Epstein, E. S., 1982: Detecting climate change. J. Appl. Meteor., 21, 1172-1182.

Hasselmann, K., 1974: On the signal-to-noise problem in atmospheric response studies. Meteorology of the Tropical Oceans, Roy. Meteor. Soc., 251-259.

Hayashi, Y., 1982: Confidence intervals of a climatic signal. J. Atmos. Sci., 39, 1895-1905.

Jäger, J., and W. W. Kellogg, 1983: Anomalies in temperature and rainfall during warm Arctic seasons. Climatic Change, 5, 39-60.

Julian, P. R., and R. M. Chervin, 1978: A study of the Southern Oscillation and Walker Circulation phenomenon. Mon. Wea. Rev., 106, 1433-1451.

Katz, R. W., 1982: Statistical evaluation of climate experiments with general circulation models: A parametric time series modeling approach. J. Atmos. Sci., 39, 1446-1455.

_____, 1983: Statistical procedures for making inferences about precipitation changes simulated by an atmospheric general circulation model. J. Atmos. Sci. (in press).

Kendall, M. G., 1975: Multivariate Analysis, Griffin, 210 pp.

Keshavamurty, R. N., 1982: Response of the atmosphere to sea surface temperature anomalies over the equatorial Pacific and the teleconnections of the Southern Oscillation. J. Atmos. Sci., 39, 1241-1259.

Kutzbach, J. E., R. M. Chervin, and D. D. Houghton, 1977: Response of the NCAR general circulation model to prescribed changes in ocean surface temperature. Part I: Midlatitude changes. J. Atmos. Sci., 34, 1200-1213.

Laurmann, J. A., and W. L. Gates, 1977: Statistical considerations in the evaluation of climatic experiments with atmospheric general circulation models. J. Atmos. Sci., 34, 1187-1199.

Leith, C. E., 1973: The standard error of time-average estimates of climatic means. J. Appl. Meteor., 12, 1066-1069.

Livezey, R. E., and W. Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. Mon. Wea. Rev., 111, 46-59.

Madden, R. A., 1981: A quantitative approach to long-range prediction. J. Geophys. Res., 86, 9817-9825.

_____, and P. R. Julian, 1971: Detection of a 40-50 day oscillation in the zonal wind in the tropical Pacific. J. Atmos. Sci., 28, 702-708.

_____, and V. Ramanathan, 1980: Detecting climate change due to increasing carbon dioxide. Science, 209, 763-768.

Manabe, S., and D. G. Hahn, 1977: Simulation of the tropical climate of an ice age. J. Geophys. Res., 82, 3889-3912.

_____, R. T. Wetherald, and R. J. Stouffer, 1981: Summer dryness due to an increase of atmospheric CO_2 concentration. Climatic Change, 3, 347-386.

Mielke, P. W., Jr., K. J. Berry, and G. W. Brier, 1981: Application of multi-response permutation procedures for examining seasonal changes in monthly mean sea-level pressure patterns. Mon. Wea. Rev., 109, 120-126.

Miyakoda, K., T. Gordon, R. Caverly, W. Stern, J. Sirutis, and W. Bourke, 1983: Simulation of a blocking event in January 1977. Mon. Wea. Rev., 111, 846-869.

Paegle, J. N., and R. B. Haslam, 1982: Empirical orthogonal function estimates of local predictability. J. Appl. Meteor., 117-126.

Preisendorfer, R. W., and T. P. Barnett, 1983: Numerical model/reality intercomparison tests using small-sample statistics. J. Atmos. Sci. (in press).

_____, and C. D. Mobley, 1983: Data Intercomparison Theory. III. S-Phase and T-Phase Tests for Spatial Pattern and Temporal Evolution, NOAA Technical Memorandum ERL PMEL-40, Pacific Marine Environmental Laboratory, Seattle, Washington.

Reynolds, G., 1978: Two statistical heresies. Weather, 33, 74-76.

Rind, D., 1982: The influence of ground moisture conditions in North America on summer climate as modeled in the GISS GCM. Mon. Wea. Rev., 110, 1487-1494.

Schneider, S. H., W. M. Washington, and R. M. Chervin, 1978: Cloudiness as a climatic feedback mechanism: Effects on cloud amounts of prescribed global and regional surface temperature changes in the NCAR GCM. J. Atmos. Sci., 35, 2207-2221.

Shukla, J., 1975: Effect of Arabian sea-surface temperature anomaly on Indian summer monsoon: A numerical experiment with the GFDL model. J. Atmos. Sci., 32, 503-511.

_____, 1981: Dynamical predictablity of monthly means. J. Atmos. Sci., 38, 2547-2572.

Storch, H. v., 1982a: Comparison of a sequence of model generated 500 mb-topographies with climate. Tellus, 34, 89-91.

_____, 1982b: A remark on Chervin-Schneider's algorithm to test significance of climate experiments with GCM's. J. Atmos. Sci., 39, 187-189.

Warshaw, M., and R. R. Rapp, 1973: An experiment on the sensitivity of a global circulation model. J. Appl. Meteor., 12, 43-49.

Washington, W. M. and R. M. Chervin, 1979: Regional climatic effects of large-scale thermal pollution: Simulation studies with the NCAR general circulation model. J. Appl. Meteor., 18, 3-16.

_____, R. M. Chervin, and G. V. Rao, 1977: Effects of a variety of Indian Ocean surface temperature anomaly patterns on the summer monsoon circulation: Experiments with the NCAR general circulation model. Pure Appl. Geophys., 115, 1335-1356.

Williams, J., 1980: Anomalies in temperature and rainfall during warm Arctic seasons as a guide to the formulation of climate scenarios. Climatic Change, 2, 249-266.

EVALUATING CLIMATE CHANGE EXPERIMENTS

F.W. Zwiers

Department of Mathematics
University of Saskatchewan
Saskatoon, Saskatchewan
S7N OWO Canada

1. Introduction

The purpose of this paper is two fold. The first is to examine some aspects of the problem of evaluating climate change experiments. The second is to partially examine the outcome of an eleven month perpetual July run of the Canadian Climate Centre (CCC) General Circulation Model (GCM) and to compare it with the outcome of a five year annual cycle run of the same model.

Climate change experiments are frequently evaluated using a procedure described by Chervin and Schneider (1976b). These authors will be referred to as CS from this point onwards. In this procedure a simple test of hypothesis is conducted at each grid point of a field to determine whether the means of the "control" and "experimental" climates at that point are significantly different. An advantage of the procedure is that the statistical test used is an exact test. That is, with suitable assumptions the exact probability distribution of the test statistic is known. This is because independent realizations of control and experimental climates are generated. One draw back of this procedure is that there is considerable computing overhead involved in generating independent climate realizations. Another is that the result of such a procedure, generally a map showing locations at which the hypothesis of equality of means is rejected, is very difficult to interpret. We describe a modification of the CS procedure in §3 which will allow us to ascribe a level of significance to the outcome of an experiment.

The output of an eleven month simulation with the CCC-GCM run in the "perpetual July" mode is examined in §4. In particular, the assumption that GCM's of this kind fairly rapidly reach statistical equilibrium (especially when run with constant forcing) is tested. The assumption that statistically independent realizations can be readily obtained is basic to the usual CS approach. In §5 a comparison is made between the outcome of the "perpetual July" simulation and the July climate of a five year "annual cycle" simulation run with the same model. These two climates are found to be significantly different from each other.

2. Experiment Design Considerations

In the usual CS setup a number of control and experimental climate simulations are generated by starting the model off from a set a initial conditions which are perturbed only slightly from run to run. To insure that realizations are independent the model is integrated for 30-days before any observations are taken. Then the model is integrated a further 30-days which are averaged to form a single realization of a monthly climate mean. When the experiment is completed a standard difference of means test is conducted at each grid point to determine whether the means of the control and experimental climates are significantly different.

Two things can be immediately noted about the procedure. One is that there is a 30-day overhead associated with each and every realization of a thirty day mean. The second is that 30-day means are computed. If a model is running in perpetual mode and if it can be reasonably assumed that the output of the model is stationary then there is no reason to chose to compute 30-day means. One is free to optimize the procedure by chosing the length of the averaging interval to maximize the power of the statistical test for equality of means subject to a cost contraint or to minimize computing costs for a given power or level of sensitivity.

For example, suppose that at a particular point both the control and experimental climates have stochastic structure similar to that of a Markov process, say

$$(x(t)-\mu) - a(x(t-1)-\mu) + \epsilon(t) \qquad (2.1)$$

where μ is the mean of the process, with the same innovation variance σ^2_ϵ and the same lag one autocorrelation coefficient a. Then the variance of an ℓ-day time average may be expressed as

$$\sigma^2_{\bar{x}_\ell} = \frac{\sigma^2_x}{\ell}\left[1 + \frac{2a}{(1-a)^2}\left(\frac{\ell-1}{\ell} - a + \frac{a^\ell}{\ell}\right)\right] \qquad (2.2)$$

so that the variance of the mean of m realizations of an ℓ-day time average is given by

$$\sigma^2_{\bar{x}} = \sigma^2_{\bar{x}_\ell}/m \qquad (2.3)$$

With the assumptions above the time averages for both climates will have the same variance so that the test for equality of means which assumes equality of variances can be used. If the same number of realizations are generated from each climate then the test statistic is given by

$$t = (\bar{x} - \bar{y})/S_p\left(\frac{1}{m_x} + \frac{1}{m_y}\right)^{\frac{1}{2}} \qquad (2.4)$$

where

$$S_p^2 = \left[\sum_{i=1}^{m} (\bar{x}_{\ell i} - \bar{x})^2 + \sum_{j=1}^{m} (\bar{y}_{\ell j} - \bar{y})^2\right]/(2m-2)$$

Here X and Y denote control and experimental climates respectively.

When normality is assumed for the ℓ-day climate means this test statistic has a Student's t distribution with 2m-2 degrees of freedom (df). The power curves for this test statistic are well known and may be found in Pearson and Hartley (1951). These curves express the power of the t-test in terms of a standardized departure ϕ from the null hypothesis of equality of means. This quantity is given by

$$\phi = \Delta/\sqrt{2}\,\sigma_{(\bar{x}-\bar{y})} \qquad (2.5)$$

where Δ is the departure from the null hypothesis.

For example, suppose that an experiment is to be designed so that a departure of one standard deviation of the daily observations from the mean can be detected with probability 0.9. Then the relationship between the sample length ℓ, the number of realizations m, the correlation coefficient a and the standardized departure ϕ is given by

$$\ell = \frac{4\phi^2}{m}\left[1 + \frac{2a}{(1-a)^2}\left(\frac{\ell-1}{\ell} - a + \frac{a^\ell}{\ell}\right)\right] \qquad (2.6)$$

Standardized departures from the null hypothesis which can be detected with probability 0.9 are given in the third row of Table 2.1. Assuming a lag one correlation coefficient of 0.75 (giving a mean time between "independent" realizations of about 7 days) the corresponding solutions of equation (2.6) for ℓ are given in the fourth row of Table 2.1. Taking the overhead of 30 days per realization into account, the cost of an experiment with the required power is given in the fifth row of Table 2.1. The most economical experiment with power 0.9 against a departure of one standard deviation of the daily obeservations requires four realizations of 50-day time averages from each climate and requires that the GCM be integrated for a total of 640 days. If the size of the departure is reduced to the standard deviation of a 30-day mean it is quite easily shown that with the assumptions made above the most economical experiment with power 0.9 consists of six realizations of a 130-day time average of each climate. Such an experiment requires that the GCM be integrated for a total of 19,200 days.

The procedure can also be reversed so that the use of available computer time is optimized.

TABLE 2.1

m	3	4	5	6	7
12	4	6	8	10	12
ϕ	3.1	2.75	2.62	2.55	2.5
ℓ	87	50	35	27	21
cost	702	640	650	684	714

Better use of the available computer time can be made if time series techniques can be used to derive tests for the difference of means. The difficulty here is that only the asymptotic distributions of the available tests are known. This places an additional constaint on the design of an experiment in that samples must be long enough to insure that the actual significance level of the test is close to the nominal level at which the test is being conducted. This question has been looked at in some detail by Zwiers (1981) and Thiebaux and Zwiers (1981).

One approach, described by Jones (1975, 1976), is based on the observation that the standard deviation of an -day time average is related to the spectral density function $f_{xx}(\lambda)$ of the observed process by

$$\sigma_{\bar{x}}^2 = 2\pi f_{xx}(0)/\ell \qquad (2.7)$$

This suggests that an apropriate test statistic for the difference of means of two time series is given by

$$t = \frac{\sqrt{\ell}\,(\bar{x} - \bar{y})}{\sqrt{2\pi}(\hat{f}_{xx}(0) + \hat{f}_{yy}(0))^{\frac{1}{2}}} \qquad (2.8)$$

where $f_{xx}(0)$ and $f_{yy}(0)$ are estimates of the spectral densities at the origin of the control and experimental climates. One estimator of the spectral density function is the Daniel estimator given by

$$\hat{f}_{xx}(0) = \frac{1}{m}\sum_{j=1}^{m} I_{xx}(\tfrac{2\pi j}{\ell}) \qquad (2.9)$$

where $I_{xx}(\lambda)$ is the squared modulus of the finite fourier transform of the observations (the Periodogram). In this case it can be shown that the test statistic proposed above is asymptotically a Student's t statistic with 4m df (Brillinger (1981)). Here m is the number of periodogram ordinates which are averaged together to form estimates of the spectral densities at the origin.

Zwiers (1981) discussed experimental design considerations in this case. For example, suppose we make the same assumptions as above and suppose we want to construct a test which can detect departures of one standard deviation of daily observations from the null hypothesis with probability 0.9. Then the most economical experiment would consist of one realization of each climate of length 175 days and would employ the Daniel estimator which averages three periodogram ordinates. Taking into account the 30-day start up time required for each realization it would be necessary to integrate the GCM for 410 days which is considerably less than the figure arrived at using the CS procedure.

Another approach which is similar to the one described above is based on the observation that the variance of an ℓ-day mean is given by

$$\sigma_{\bar{x}}^2 = \frac{1}{\ell}\left(C_{xx}(0) + 2\sum_{\tau=1}^{\ell-1}(1 - \tfrac{\tau}{\ell})C_{xx}(\tau)\right) \qquad (2.10)$$

where $C_{xx}(\tau)$ is the autocovariance function (acf) of the observed process. Thus an estimate of the variance of the ℓ-day mean may be made by replacing the acf in (2.10) with an estimate. One must be a little careful here because the asymptotic distribution of the resulting test statistic for the difference of means will depend upon how the acf is estimated (Thiebaux and Zwiers (1983)). If it is estmated directly from the data and we assume equality of variances for control and experimental climate means then it can be shown that the test statistic will be asymptotically distributed as a Student's t variable with 4 df (Brillinger (1981)). On the other hand, if the estimate of the acf is derived from an autoregressive time series model fitted to the data the resultant test statistic for the difference of means will asymptotically have a standard Normal distribution and be more powerful than the test employing the direct estimate of the acf. The latter approach has been used by Katz (1982) and will be used below in an examination of the perpetual July simulation mentioned above. Thiebaux and Zwiers (1981) have examined in part the question of the length of sample necessary so that the distribution of the test statistic for the difference of means is close to its asymptotic distribution. For the example examined above it appears that realizations of length approximately 100 to 150 days will yield tests with true significance level close to their nominal levels.

3. Experiment Evaluation Considerations

A further difficulty with the CS approach lies with the interpretation of results. Even if the null hypothesis of equality of means were true at every gridpoint on which the GCM is defined we would still expect it to be rejected at about 5% of points if the tests were conducted at the 5% significance level. However, since the field of test statistics is spatially correlated, these points will not be scattered about the field at random. Rather, they will be gathered in pools giving the appearance that climate changes occured in these regions when in fact no change has occured. This problem has been illustrated with a numerical example by Storch (1982).

Clearly what is required for the correct interpretation of the results of climate change experiments is

(i) to ascribe a level of significance to the differences between the control and experimental mean fields.

(ii) to develop some sort of multiple comparison procedure to determine the regions in which significant departures from the null hypothesis occured.

We will address only the former.

In order to ascribe a level of significance to the differences between two mean fields it is necessary to conduct a multivariate test of hypothesis. However, while the statistical theory of Multivariate Analysis is well developed it cannot be

employed directly. The reason is that the problem dictates the use of unobtainable amounts of data with standard multivariate tests. This has been noted in the literature by Hasselmann (1979) amongst others.

One possible solution is to reduce the dimensionality of the problem, perhaps by using empirical orthogonal functions (EOF's) of the observed atmosphere directly or in a two stage procedure as has been suggested by Storch (1982). Another would be to represent the spatial covariance structure of climate means with a model depending upon only a few free parameters. This approach is taken in objective analysis and Thiebaux (1976) has developed accurate anisotropic spatial corelation functions for daily 500mb height over North America.

A third approach which we would like to suggest uses a "randomization" test to ascribe a level of significance to the outcome of a CS type experiment. Points in the field are randomly labeled as being either "control" or "experimental" points. Three simulations are needed: two "control" and one "experimental". At control points tests are conducted comparing the two control climates at that point. At experimental points one of the control climates is compared with the experimental climate. Then the difference is found between the rate of rejections at experimental and control points. A significance level is ascribed to this value as follows:

(i) the points in the field of test statistics generated above are randomly relabeled as "control" or "experimental". The values at the points are not recomputed.

(ii) the difference between rejections rates at experimental and control points is computed.

(iii) steps (i) and (ii) are repeated a large number of times and the results are used to estimate the probability distribution of the difference between rejection rates.

(iv) having estimated the distribution of the difference between rejection rates, we can ascribe a level of significance to the observed difference.

The procedure works because when the null hypothesis of equality of means is true everywhere it makes no difference how points are labeled. Each t-value, whether "control" or "experimental", will be a realization of the same test statistic.

A nice property of the procedure is that the exact significance level of the statistical test used at individual points does not have to be known. Therefore, as long as we can argue that the significance level of individual tests is more or less constant across the field, the procedure can be used. This is not to say that the form of the individual tests is not important. These will no doubt have a considerable impact upon the sensitivity of this procedure. An application of this procedure will be described below.

4. Examination of a CCC-GCM perpetual July simulation.

The CCC-GCM is a 10 level, global, primitive equations, spectral model which contains radiative processes including a diurnal cycle, a surface heat budget and a hydrological cycle including surface hydrology, soil moisture, and snow amounts. The sea-surface temperatures in the model are fixed as are the zonally averaged amount and location of clouds. The model is integrated numerically using a semi-implicit time stepping method with time step of about 30 minutes.

The model was integrated for 11 months with "perpetual July" solar declination. The last 240 days of the simulation (8 months) were used in the testing process. This allowed some three months for the model to attain equilibrium.

The 240 day run was split into two sections of 120 days each and the model's 500mb height field for the two periods was compared as follows. Autoregressive time series models were fitted to the sequence of observations at each of 2080 gridpoints using program UNIMAR from the TIMSAC-78 library (Akaike, et al (1979)). The order of model fitted was determined objectively using the information theory based AIC criterion (Akaike (1974)). Acf's were derived from the fitted models and these were used in expression (2.10) to provide estimates $\hat{\sigma}^2_{\bar{x}}$ and $\hat{\sigma}^2_{\bar{y}}$ of the variance of the time average at each grid point for the first and second 120-day period respectively.

The hypothesis that the mean 500mb height during the first 120-day period was equal to that during the second 120-day period was tested by comparing the value of the test statistic

$$t = (\bar{x} - \bar{y})/(\hat{\sigma}^2_{\bar{x}} + \hat{\sigma}^2_{\bar{y}})^{\frac{1}{2}} \qquad (4.1)$$

with critical values from its asymptotic distribution. The computed t-values are displayed in Figure 4.1. The shaded regions of this map indicate areas in which the null hypothesis of equality of means is rejected at the NOMINAL 5% significance level. They cover roughly

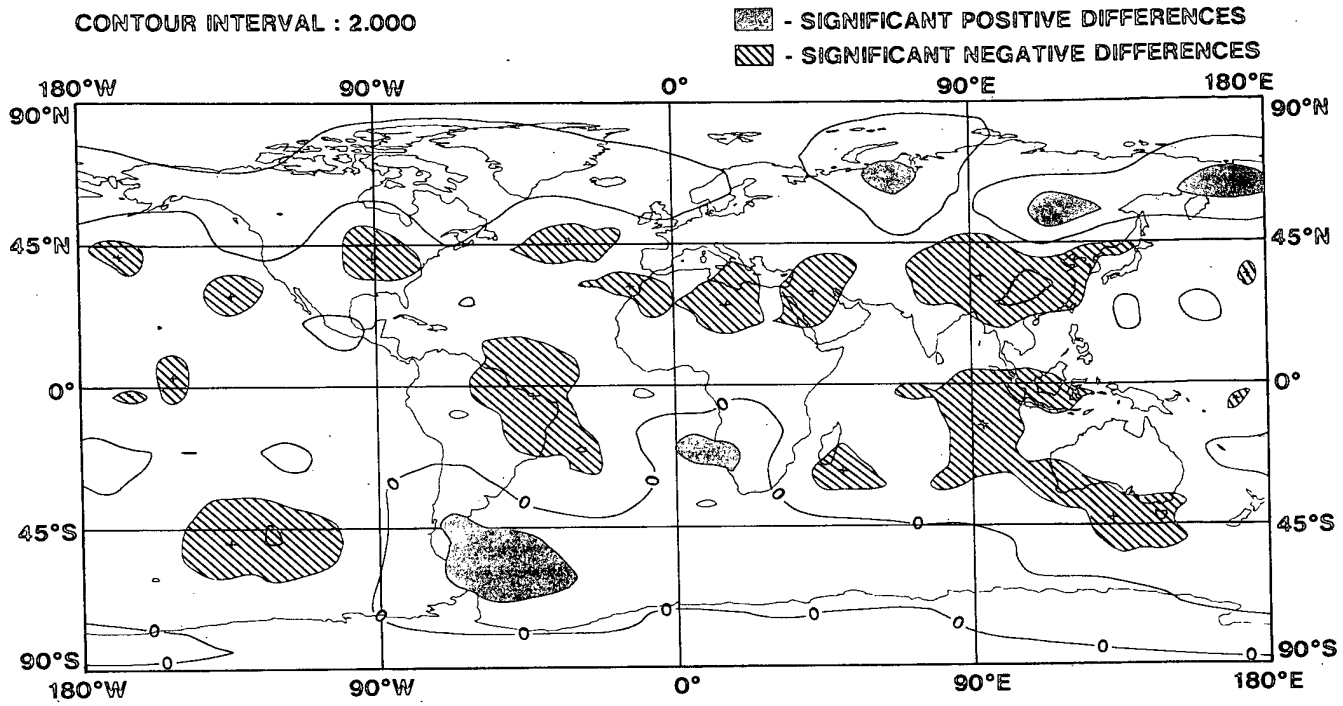


Figure 4.1

T-values comparing mean 500mb height fields for two halves of a 240-day perpetual July simulation.

18.8% of the globe. The true significance level of the tests is unknown but is not felt to be higher than 10%.

As noted previously, because the t-values are spatially correlated it is not possible to ascribe a significance level to the observed differences. However the map does suggest that the 500mb height field gradually subsides with time when the model is operated in perpetual July mode. The map seems to indicate that there is some land-sea interaction involved in whatever mechanism is causing the subsidence of the 500mb height field. In particular there is quite a large band of positive significant differences at northern mid latitudes.

The differences between the two periods are too large to be attributable to a loss of mass from the model's atmosphere from numerical sources as can be evidenced from Figure 4.2. In this figure the mean global 500mb height is displayed for each of eight 30-day periods contained within the 240-day segment of the simulation under consideration. A linear least squares fit to these values indicated by the dashed line shows that the mean global 500mb height subsides at the rate of about 1.8 meters per month during the eight month period.

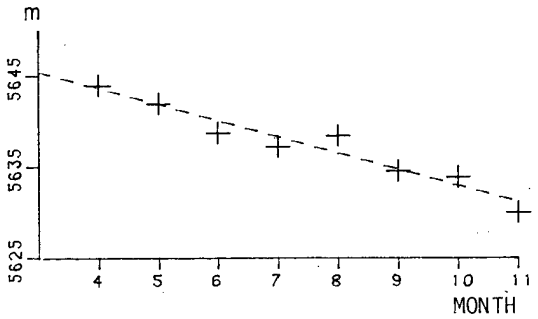


Figure 4.2
Mean global 500mb height for the last eight months of the perpetual July run.

5. Comparison of two July climates.

In addition to the perpetual mode July simulation discussed above, an annual cycle simulation extending over slightly more than five years has also been made with the CCC-GCM. Six July monthly mean 500mb height fields were available from this run. In order to make a comparison between the annual cycle July climate of the model and its perpetual July climate the 240 days of the perpetual July run were divided into 8 "months" from which monthly mean 500mb height fields were computed. Every other mean field was discarded so that there would be 30-day intervals between the sequences of observations used to compute the mean fields. In this way the fields which were used could be regarded as

being approximately statistically independent.

We had then six realizations of the model's annual cycle July monthly mean 500mb height field and four realizations of the same field from the model's perpetual mode July climate. To compare these two climates we followed the usual CS procedure by computing the t-statistic for the difference of means which assumes equality of the variance of the monthly means in both climates. The variablity of the monthly means had previously been compared using a standard F test. At the 10% significance level the null hypothesis of equality of variances was rejected over about 14% of the globe. As is discussed further below, we felt that this did not provide very strong evidence for a difference in variability of the two climates. The t-values which were computed are displayed in Figure 5.1. The shaded regions indicate where the null hypothesis of equality of means is rejected at the 5% significance level. These regions cover approximately 53% of the globe. The mean zonal profiles of the two sets of mean fields are displayed in Figure 5.2 and give an indication of the size of the differences. There seems to be no question that the annual cycle July climate of the model is generally cooler than the perpetual mode July climate.

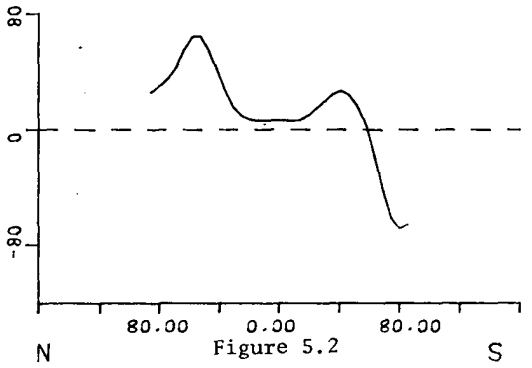


Figure 5.2

Difference of mean July 500mb height zonal profiles.

As an illustration of its use, the randomization test described above was used to ascribe a level of significance to this result. The first step was to divide the six annual cycle July mean fields into two groups of size three which could act as the control simulations. The mean 500mb height field of the first half of the annual cycle run was compared with the mean 500mb height field of the second half of the annual cycle run using test statistic (2.4). The computed t-values are shown in Figure 5.3. The shaded regions indicate grid points at which the hypothesis of equality of means was rejected at the 5% significance level. These regions cover

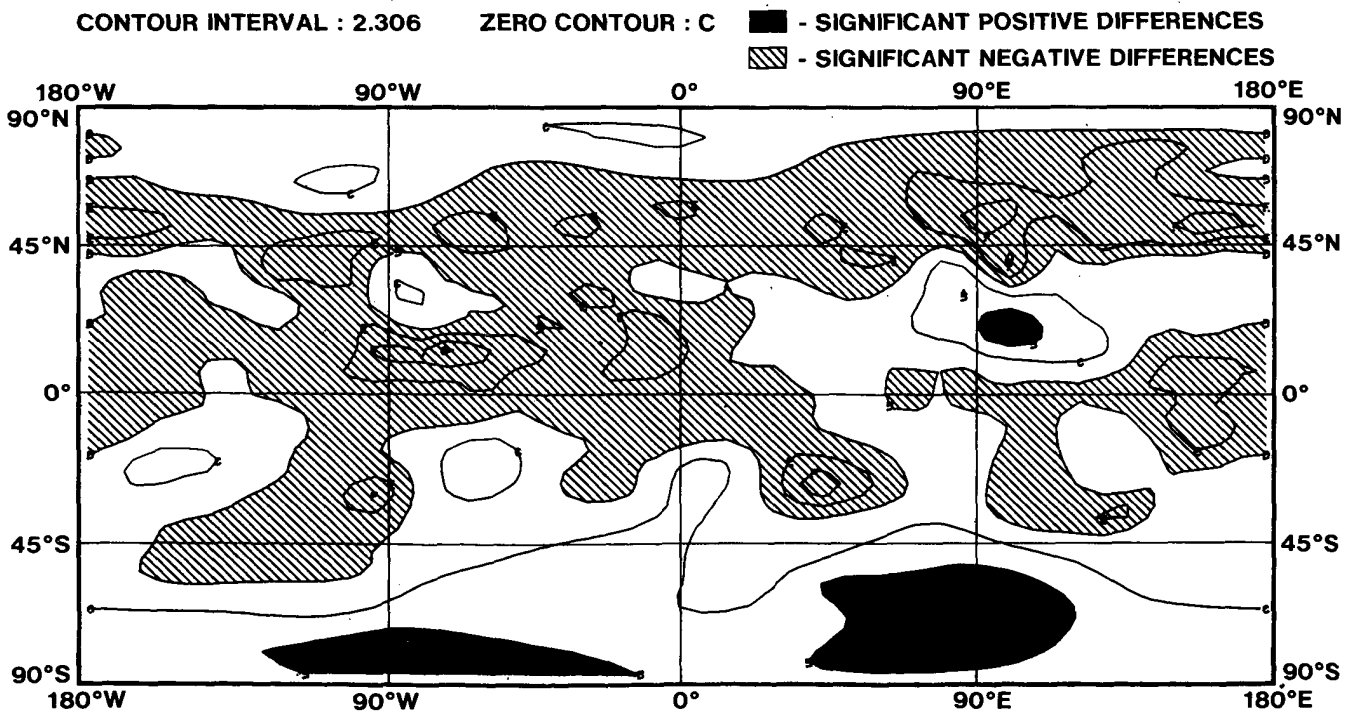


Figure 5.1

T-values comparing annual cycle and perpetual mode July 500mb height fields.

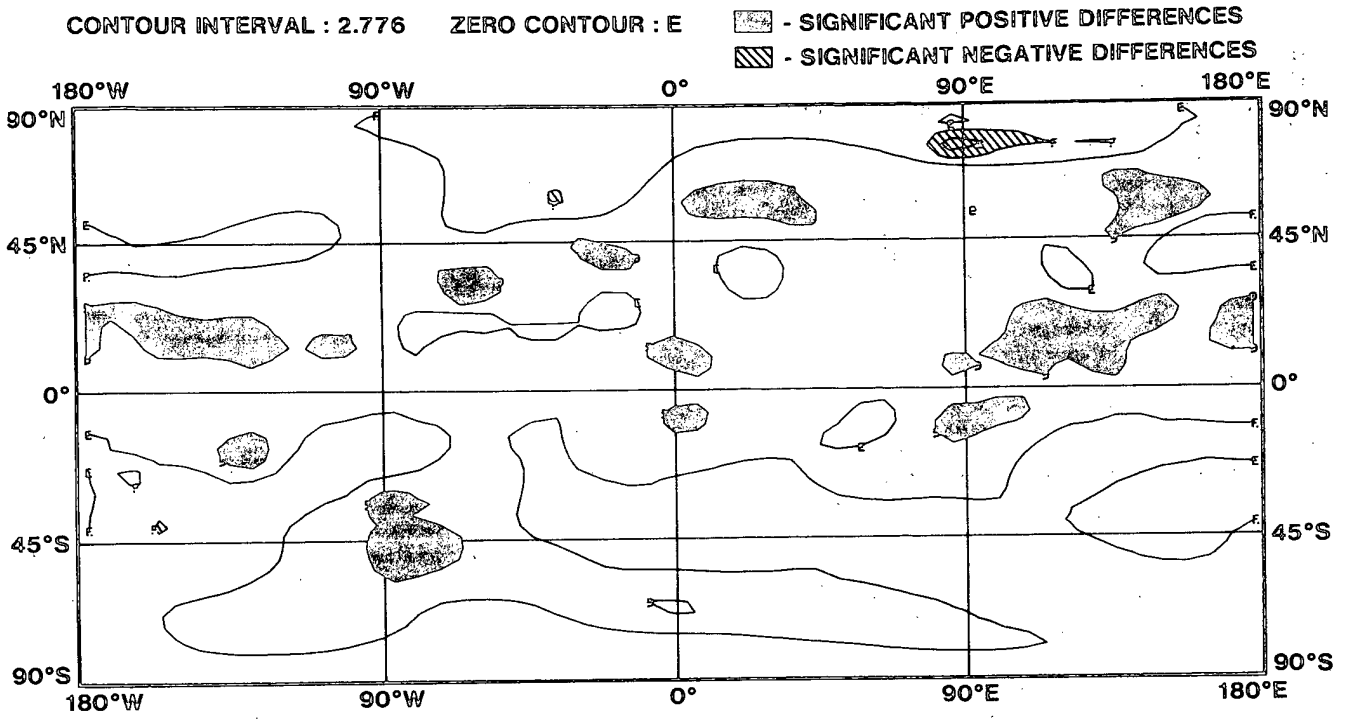


Figure 5.3

T-values comparing the first three and last three July monthly mean 500mb height fields of the annual cycle simulation.

about 8.8% of the globe. The slightly higher than expected value and the fact that most significant differences were positive is probably due to a gradual decrease in the mean July 500mb height field as is seen from Figure 5.4. The average rate of decrease of about 1.6 meters per year corresponds reasonably well to the annual rate of mass lost from the model's atmosphere due to numerical reasons.

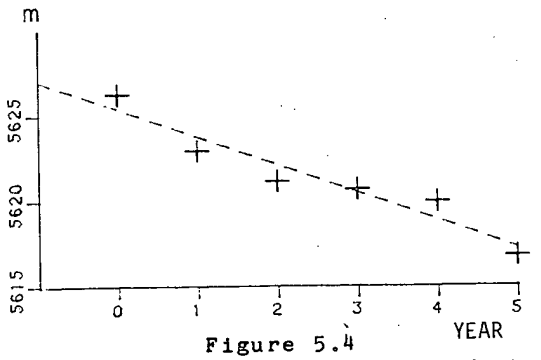


Figure 5.4
Mean July global 500mb height of the "annual cycle" climate.

The second step was to compare the mean 500mb height field of the first half of the annual cycle run with the mean of the first three realizations of the perpetual mode July mean 500mb height fields. These three mean fields from the perpetual July run serve as the experimental climate simulation. Only three realizations were used so that control vs control and control vs experiment tests would have the same power. T-values were computed and plotted. It was found that regions in which the null hypothesis of equality of means was rejected cover approximately 33.4% of the globe. This decrease in rejection rate from the case in which all available data is used may be attributed to a decrease in the power of the test statistic. The structure of the graph was very similar to that of Figure 5.1. Regions in which rejections occurred were smaller than in Figure 5.1 but were found in the same locations.

Finally the permution procedure was applied as described in §3. In total 1000 realizations of the difference in rejection rates between control-control comparisons and control-experiment comparisons were made. The observed difference of rejection rates ranked at the 100th percentile of such differences and thus the hypothesis of equality of means can be rejected at the 1% significance level (Ripley (1981)).

The variablity of the monthly means of the annual cycle and perpetual mode climates was also compared with the randomization test. When F tests were conducted at the 10% significance level

comparing the variability of monthly means during the first and second half of the annual cycle run rejections of the null hypothesis occured over approximately 7.9% of the globe. The corresponding figure for the comparison of the first half of the annual cycle run with the three realizations from the perpetual mode run was 10.6%. The difference between these rates was ranked at the 69th percentile of such differences generated by the permutation procedure.

6. Discussion

There is apparently fairly strong evidence that GCM's such as the CCC-GCM reach equilibrium quite slowly. The reason for the generally decreasing trend in mean 500mb height has not been identified but it is speculated that the soil moisture component of the hydrology cycle may somehow be implicated. This problem will have to be examined more closely because it has serious implications for some of the options available for the design of efficient climate change experiments as discussed in §2.

Also we discussed the problem of the interpretation of the results of climate change experiments using the usual CS approach and have illustrated a method which seems to be reasonably competant at ascribing global significance levels to the outcomes of climate change experiments. It should be pointed out that the results of our comparisons between the annual cycle and perpetual mode July climates of the CCC-GCM should not be taken to literally. Since our primary purpose was to illustrate the use of the randomization test, we did not attempt to remove the small trends that have been identified in the 500mb height fields of the climates. In particular, the failure to remove these trends make the results of the tests for equality of variances difficult to interpret.

7. References

Akaike, H., Kitagawa, G., Arahata, E. and Tada, F., 1979: TIMSAC-78, Computer Science Monographs, No. 11, The Institute of Statistical Mathematics, Tokyo.

Akaike, H., 1974: A new look at the statistical model indentification. IEEE Trans. Auto. Control, AC-19, 716-723.

Brillinger, D.R., 1981: Time series data analysis and theory. Holden-Day.

Chervin, R.M., and S.H. Schneider, 1976: On determining the statistical significance of climate experiments with general circulation models, J. Atmos. Sci., 33, 405-412.

Hasselmann, K., 1979: On the signal-to-noise problem in atmospheric response studies, in Meteorology over the Tropical Oceans, ed. D.R. Shaw, Roy. Met. Soc., Bracknell, England, 251-259.

Jones, R.H., 1975: Estimating the variance of time averages. J. Appl. Met., 14, 159-163.

Jones, R.H., 1976: On estimating the variance of time averages. J. Appl. Met., 15, 514-515.

Katz, R.W., 1982: Statistical evaluation of climate experiments with general circulation models: A parametric time series modelling approach, J. Atmos. Sci.,39,1446-1455.

Pearson, E.S. and Hartley, H.O., 1951: Charts of the power function for analysis of variance tests derived from the non-central F-distribution. Biometrika, 38, 112-130.

Ripley, B.D., 1981: Spatial statistics. Wiley.

Storch, H.V., 1982: A remark on Chervin-Schneider's algorithm to test significance on climate experiments with GCM's, J. Atmos. Sci., 39, 187-189.

Thiebaux, H.J., 1976: Anisotropic correlation functions for objective analysis. Mon. Wea. Rev., 104, 994-1002.

Thiebaux, H.J., and Zwiers, F.W., 1981: Some alternative approaches to testing for the significance of differences in climate model outcomes, Preprint volume, 7th conf. on Prob. and Stat. in Atmos. Sci., Nov. 2-6, Monterey, CA, 34-39.

Thiebaux, H.J., and Zwiers, F.W., 1983: On the interpretation and estimation of effective sample size, Preprint volume, 2nd Int. Meeting on Stat. Climatology, Sept. 26-30, Lisbon, Portugal.

Zwiers, F.W., 1981: Simulation experiments with a test for the mean of a time series for use in climate change experiments, Preprint volume, 7th Conf. on Prob. and Stat. in Atmos. Sci., Nov. 2-6, Monterey, CA, 28-33.

Acknowledgments

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (Grant A5448) and in part by the Canadian Climate Centre. I would like to than Dr. G.J. Boer of the Canadian Climate Centre for his many helpful discussions and for making available the data from the CCC-GCM. Also, I would like to thank A. Kelm for his assistance with plotting the figures.

A STUDY OF THE STATISTICAL SIGNIFICANCE OF THE RESPONSE OF A SPECTRAL GENERAL CIRCULATION MODEL TO A MODIFICATION OF SOLAR RADIATION.

Jean-Michel Veysseire

Centre National de la Recherche Météorologique
Unité de Dynamique du Climat
31057 TOULOUSE CEDEX
FRANCE

1. INTRODUCTION

One of the most important use of a General Circulation Model is to simulate a change in the conditions that rule the atmospheric motions (external forcing, modifications in the air composition or in boundary conditions), and to compute the modifications in time averages of climatological variables. However, every GCM (like the real atmosphere) has an intrinsic variability in the way that in several similar experiments, without changing the model, the studied time averages vary. The problem arising now is to determine in what degree a change computed in an experiment where the design of the model has been modified is really significant and not only due to the intrinsic variability of the model. Many techniques have been proposed (Laurmann and Gates (1977), Katz (1982), Chervin (1980)), the one proposed by Chervin and Schneider (1976), and using a Student test, being the most widely employed.

In this paper we look, by means of a GCM, into the response of the atmosphere to the modifications of incident solar radiation intensity and repartition, due to the variation of astronomic parameters of the Earth's orbit. We dispose of two simulations of a complete annual cycle, with the conditions of insolation as they were 125,000 and 115,000 years ago, these dates approximately corresponding to the beginning and the end of the optimum of the last interglacial (perturbed experiments). We have also five years of climatic simulation in the present conditions (control experiments).

The model used is a global GCM based on the spectral method and developped by the FERM (Etablissement d'Etudes et de Recherches Météorologiques - Paris). It is a model in primitive equations with ten vertical layers in σ-coordinate. The truncation is a trapezoidal S 10-13 one, and the associated "Gaussian" grid, where all the fields used here are given, has 20 latitude circles and 32 equidistants points on each latitude circle. The model takes into account the main physical processes: parametrization of radiation, convection, cloudiness, synoptic and convective precipitations, and a complete hydrologic cycle.

The boundary conditions at the surface are the same in all the simulations: they include orography and land-sea distribution, sea surface temperature, annual ice-floe evolution, these last values being fixed by present climatology. Indeed, geological studies show that these conditions were not very different in this time. The model has been validated in comparison with present climatology in experiments of perpetual January and July and also by simulating complete annual cycles (Royer (1980), Royer et al. (1981), Déqué and Royer (1982))

We shall look only into the significance of differences for the summer months (June, July and August), where we can find the most important differences of insolation in comparison with the present time.

We study 29 variables, each of them being given on an horizontal grid of 640 points (20×32). They are listed in Table 1 (see next page).

2. USE OF THE CHERVIN AND SCHNEIDER'S ALGORITHM

In this part, the significance of the observed change is separately computed for each variable on each grid point. The comparisons between control and perturbed experiments are made only on time averages, that is to say on the average for each year and each grid point of daily values during the three months June, July and August. The computed deviation is the difference between the ensemble averages, separately computed in the control and perturbed cases. By dividing this deviation by an estimation of the standard deviation of the time averages, we obtain a signal-to-noise ratio. We multiply this ratio by a coefficient depending only on the available number of years

Table 1. Listing of the variables.

n°1	Zonal speed at 850 mbar
n°2	Zonal speed at 500 mbar
n°3	Zonal speed at 200 mbar
n°4	Meridian speed at 850 mbar
n°5	Meridian speed at 500 mbar
n°6	Meridian speed at 200 mbar
n°7	Vertical speed at 850 mbar
n°8	Vertical speed at 500 mbar
n°9	Vertical speed at 200 mbar
n°10	Temperature at 850 mbar
n°11	Temperature at 500 mbar
n°12	Temperature at 200 mbar
n°13	Saturation deficit at 850 mbar
n°14	Saturation deficit at 500 mbar
n°15	Saturation deficit at 200 mbar
n°16	Geopotential at 850 mbar
n°17	Geopotential at 500 mbar
n°18	Geopotential at 200 mbar
n°19	Sea level pressure
n°20	Snow cover
n°21	Soil water content
n°22	Surface temperature
n°23	Evaporation
n°24	Sensible heat
n°25	Solar radiation absorbed at the surface
n°26	Infrared radiation
n°27	Cloudiness
n°28	Precipitations
n°29	Thermal amplitude

in each case, so we obtain a variable that follows a Student law in the case where there is no significant change on the observed point.

The results are presented by means of maps visualizing the values of the signal-to-noise ratio. For all these maps, the critical value of the ratio at the significance level 5% is ± 3.04: the areas of values greater than $+3$ or smaller than -3 will be considered as significant at the level 5%.

But in a field that doesn't present any global significant variation, it is possible to find areas where signal-to-noise ratio is greater than the critical value (Storch (1982)). So it is interesting to test also the significance for a field as a whole. In the assumption that the field doesn't change as a whole, by using the previous method we computed at each point a variable drawn from a Student variate. Now we test if the aggregate of these variables can be considered as a sample drawn from a Student variate. Assuring these variables are statistically independant, we could use for exemple a Kolmogorov test: we compute the most important deviation between the sample cumulative distribution function and the Student cumulative distribution function. We multiply it by the square root of the number of points. The asymptotic distribution of this variate is known and can be applied to a number of points greater than 80 (see for exemple Kendall and Stuart (1979))

Unfortunately, the different points are not independent, so we can't directly apply a Kolmogorov test. But it is possible to have an opinion on the significance of a field by replacing the real number of points by an effective number of uncorrelated points. To compute this number, we use the method exposed by Der Megreditchian (1979) based on the equality of the distributions of the "anomalicity coefficients" of two fields. We define the effective number by:

$$n_{eff} = \frac{(tr(R))^2}{tr(R^2)} = \frac{N^2}{tr(R^2)} \qquad (1)$$

where N is the real number of points and R is the covariance matrix of the points.

Therefore, we have to compute an estimation of the matrix R, when we have only one realization of the field. Then, we assume that the field is statio nary and we estimate the spacial correlation between two points as a function of their distance and of the direction with regard to the north of the segment joining them on the Earth's surface. But sometimes we find estimations of correlation coefficients smaller than -1 : it can mean that the field is notstationary and so that there are some points where the variations are significant, but also it can mean that for remote points the precision of the covariances computing is insufficient. But this case is seldom seen.

So we have a test giving an idea of the global significance of a field, but there is no theoretical result rigorously justifying the replace of the real number by an effective number for not independent values. As a checking of the reliability of the results, we took one of the five years of present climate simulation as "perturbed experiment" and we compared it with the

four other years as "control experiments".

Here we don't go into all the details of the maps obtained in every case for each of the 29 variables;their interpretation will be made in a subsequent publication.Generally,we observe more significant areas 125,000 years ago than 115,000 years ago,especially for some variables as solar radiation or geopotential at 500 mbar.Maps corresponding to the geopotential at 500 mbar and histograms of the values shown on these maps are given in figures 1 and 2.

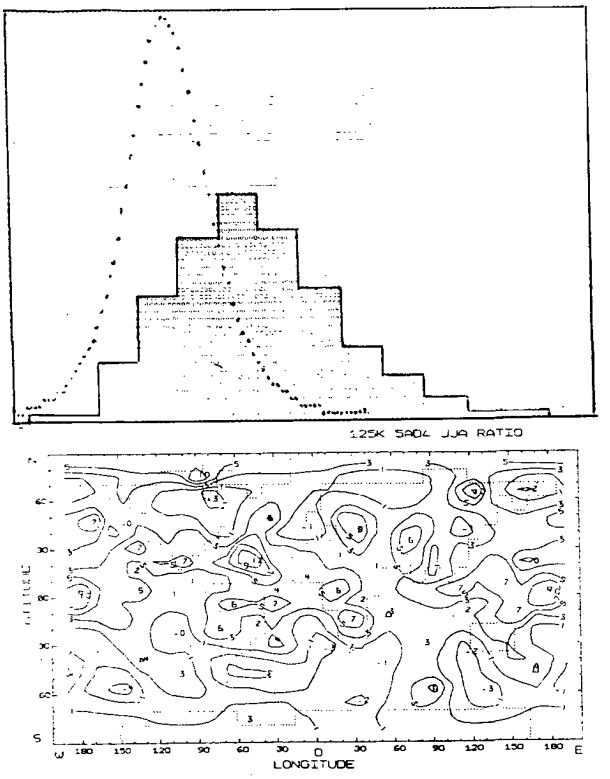


Fig. 1. Histogram and map of the values computed at each grid point by the Chervin and Schneider's method for the geopotential at 500 mbar 125,000 years ago.

For the year of present climate simulation,the maps show also some significant areas,as expected.Results for the geopotential at 500 mbar will be found on figure 3.

Moreover,we find areas with very great values for snow cover and soil water content.This is because the assumption that these variables are drawn from normal populations is not fulfilled,as we can realize observing their numerical values.So we can't apply the Student test here.Map and histogram of snow cover are shown on figure 4.

On the other hand,the surface temperature and the thermal amplitude being fixed to their climatological values,the

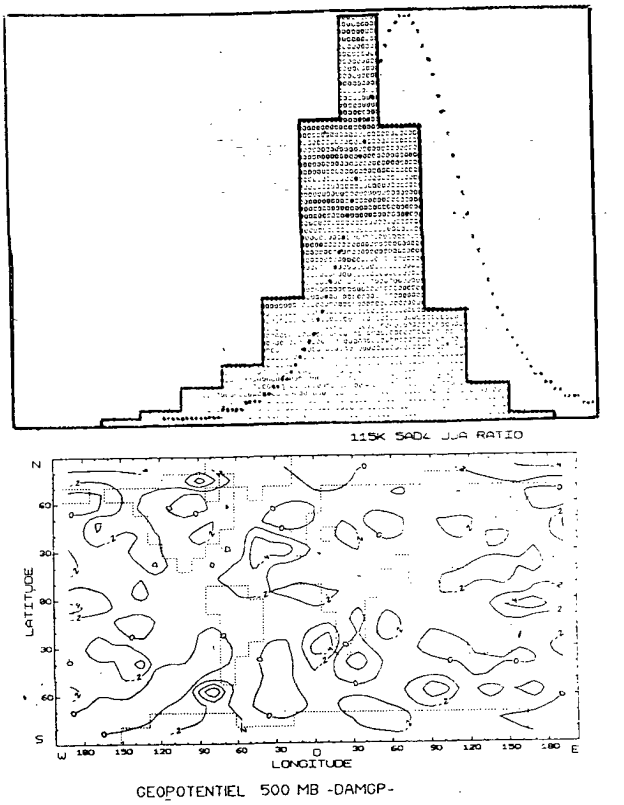


Fig. 2. Same as figure 1, 115,000 years ago.

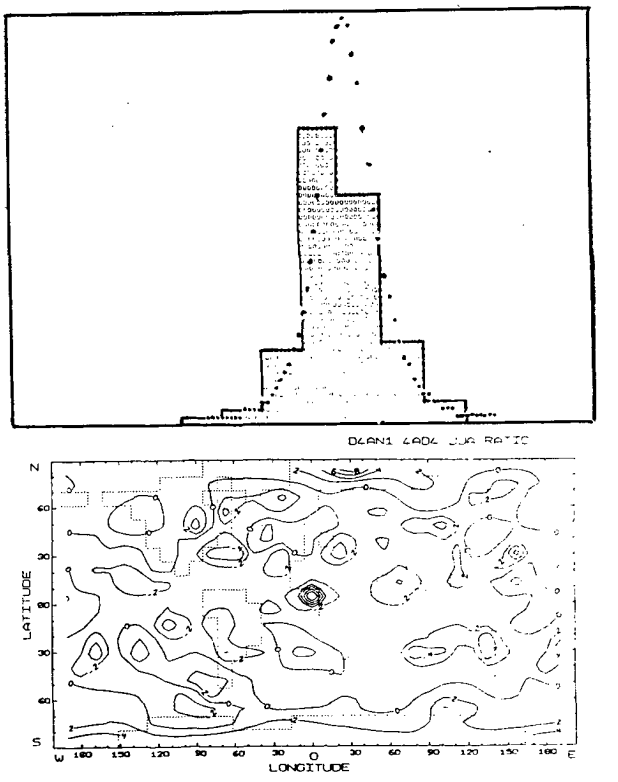


Fig. 3. Same as figures 1 and 2,for the present climate simulation.

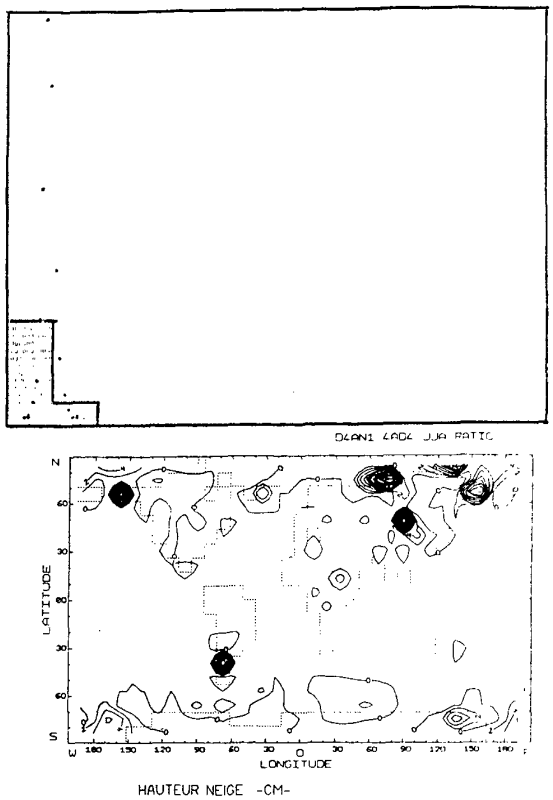


Fig. 4. Histogram and map for the snow cover in the present climate simulation.

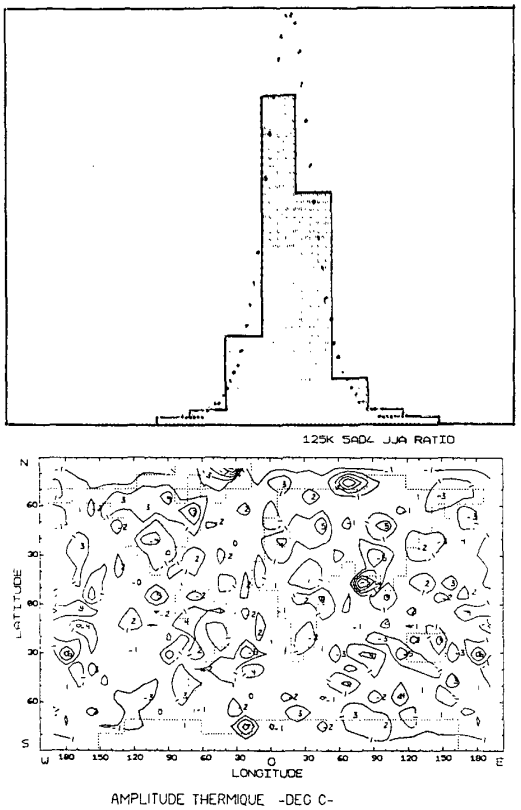


Fig. 5. Histogram and map for the thermal amplitude 125,000 years ago.

observed differences come only from the imprecision of the computation.

The study of global significance corroborate the previous conclusions.For the Kolmogorov test without effective number: for the simulation 125,000 years ago,one field only,the thermal amplitude (see figure 5),is not significant at the 5% significance level;for the simulation 115,000 years ago,13 are not significant;and for the present climate 12 are not significant. This last result clearly shows that we can't apply this test just as it is.Using the effective number method,we obtain:

6 fields are significant in the simulation of 125,000 years ago:
-temperature at 850 mbar,
-temperature at 500 mbar,
-saturation deficit at 850 mbar,
-geopotential at 500 mbar,
-geopotential at 200 mbar,
-solar radiation absorbed at the surface.

4 fields are significant in the simulation of 115,000 years ago:
-temperature at 500 mbar,
-geopotential at 500 mbar,
-geopotential at 200 mbar,
-solar radiation absorbed at the surface.

There is no significant field for the present climate.

But,we have to note that sometimes, the effective numbers were smaller than 80 and that correlation coefficients were occasionally smaller than -1 (for 2 fields 125,000 years ago,6 fields 115,000 years ago and 4 fields under the present climate).

In conclusion,the Chervin and Schneider method can't be applied for two variables, and the Kolmogorov test completed by the effective numbers computation gives interesting indications of the global significance of the fields.

3. TIME SERIES' STUDY

In the previous section,we compared only trimestrial time averages,so we didn't use the information included in the time series simulated by the model,as Katz (1982) showed off,proposing a parametric time series modeling approach.Here,we'll use at each point an estimation of the autocorrelations for a particular time series: the difference between perturbed daily values and the average of the values for all the control experiments.Thus we obtain a series near stationarity(Laurmann and Gates (1977)). Its mean is zero when there is no significant difference.To estimate this mean,we

use a linear unbiased minimal variance estimator because a mean different from zero is more especially significant that the sample variance is smaller. Generally, this estimator is different from the arithmetic average (see for instance Anderson (1971)). Its expression is:

$$\hat{m} = (U'V^{-1}U)^{-1}(U'V^{-1}X) \qquad (2)$$

where $X=\{x_1,\ldots,x_n\}$ is the observed sample of the time series, U is the n-dimensional vector $\{1,\ldots,1\}$ and V is the autocovariance matrix of X.

The variance of $\hat{m}$ is $(U'V^{-1}U)^{-1}$

Here, the difficulty is to estimate V. Assuming the series is stationary, we take as autocovariance estimation:

$$c(k)=\frac{1}{N}\sum_{t=1}^{N-k}(x_t-\bar{x})(x_{t+k}-\bar{x}) \qquad (3)$$

where N is the total number of points,

$\bar{x}=\frac{1}{N}\sum_{t=1}^{N}x_t$ is the sample mean,

and the lag k varies from 0 to N-1. This estimation gives a definite positive matrix.

So the tested quantity is the ratio between $\hat{m}$ and its standard estimation :

$$q=(U'\hat{V}^{-1}U)^{-\frac{1}{2}}(U'\hat{V}^{-1}X) \qquad (4)$$

We don't know the distribution of q, but if the series expectation is zero, the distribution of q is centred because $\hat{m}$ is an unbiased estimator. Nevertheless, to obtain conclusions for the significance of the changes, we made a simulation of 1500 values of q, each of them being computed from a drawing of 92 values of a normal distribution with mean zero and variance one (92 is the total number of days of June, July and August). The simulated sample is centred (average of -0.024) with standard deviation of 3.226, the extreme values being -24.497 and 18.251. It can't be considered as Gaussian.

The quantiles at 5% and 1% are about 7 and 11. The punctual test we used is then to consider as significant at the 5% level the values of q greater than 7 or smaller than -7. For the global significance, we compared the cumulative distribution function of all the values of q in a field with the simulated sample one. The Kolmogorov test is now replaced by a Smirnov test, and the effective numbers are estimated as previously.

The results look like the previous section ones, the modifications being generally more significant 125,000 years ago than 115,000 years ago, and the present climate simulation showing significant areas.

For the global significance and the Kolmogorov test without effective number: for the simulation 125,000 years ago, 2 fields are not significant at the 5% level: the saturation deficit at 200 mbar and the thermal amplitude; for the simulation 115,000 years ago, 15 are not significant; and for the present climate 12 are not significant. If we introduce the effective numbers:

5 fields are significant 125,000 years ago:
- -temperature at 850 mbar,
- -temperature at 500 mbar,
- -temperature at 200 mbar,
- -geopotential at 200 mbar,
- -solar radiation absorbed at the surface.

5 fields are significant 115,000 years ago:
- -temperature at 500 mbar,
- -saturation deficit at 850 mbar,
- -geopotential at 500 mbar,
- -geopotential at 200 mbar,
- -soil water content.

2 fields are significant for the present climate:
- -temperature at 200 mbar,
- -soil water content.

Nevertheless, in the previous section, we saw that it was difficult to obtain a valid conclusion for the soil water content, and anyhow it is not surprising to find one or two erroneous conclusions at the level 5% for a series of 29 values.

Here also, the effective numbers can be inferior to 80 and the correlation coefficients inferior to -1 (for 4 fields 125,000 years ago, 9 fields 115,000 years ago and 1 in the present climate).

We can see that the obtained results are rather similar to the Chervin and Schneider method ones. There is no tangible increase of the number of significant fields and of the areas of significant change on the maps. Probably the imprecision of the covariance matrix estimation compensate the profit that we thought to obtain using the time series and a minimal variance estimator.

As an example, figure 6 shows the maps of the values of q for the geopotential at 500 mbar. We can compare these maps with the maps on figures 1, 2 and 3, keeping in mind that the critical values are not the same: for the significance level 5%, they are represented by the isolines ± 3 in the figures 1, 2 and 3 and by the isolines ± 7 in the figure 6.

4. CONCLUSIONS

The direct use of the time series, without estimating a parametric model, doesn't appreciably reduce the signal-to-noise ratio, because of the deficiency of precision

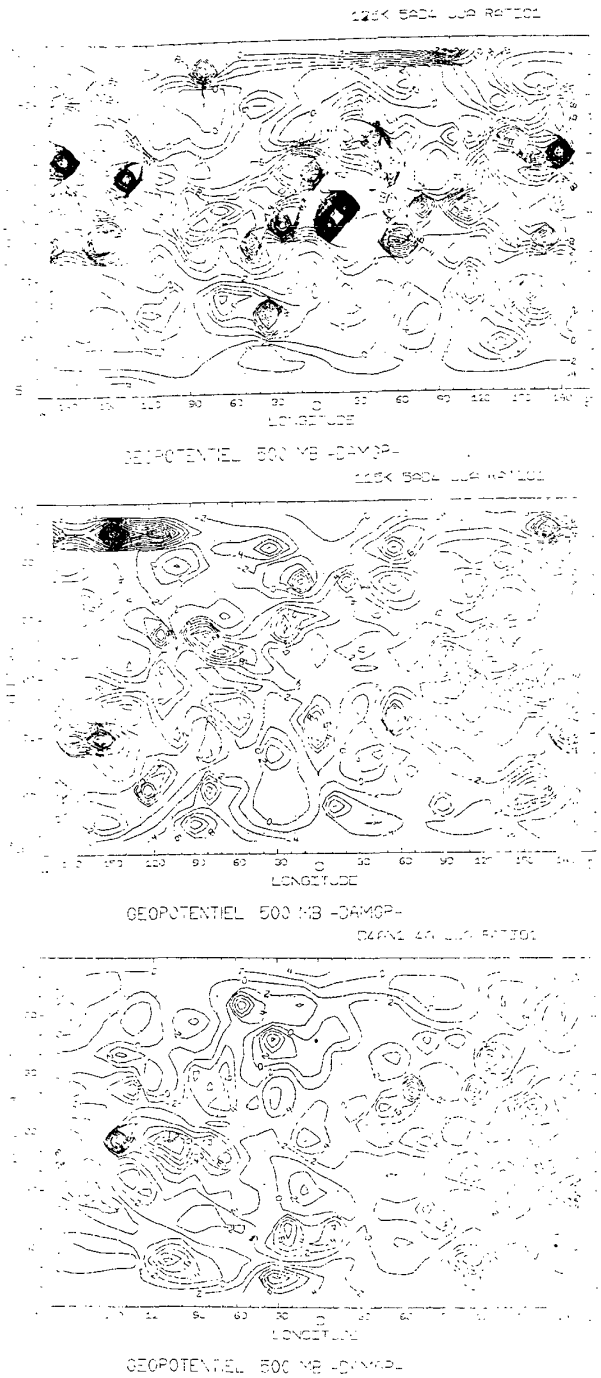


Fig. 6. Maps of the values at each grid point of the minimal variance estimator of the mean difference between perturbed and control experiments, divided by its standard deviation.
Upper: 125,000 years ago.
Middle: 115,000 years ago.
Lower: present climate.

in the time series covariances estimation. So it seems necessary to use a model adjustement, in order to reduce the number of estimated parameters.

On the other hand, the application of the Kolmogorov test, completed by an estimation of effective numbers, often permits to pass judgment on the global significance of a field.

5. REFERENCES

Anderson, T. W., 1971: The statistical analysis of time series. John Wiley and Sons.

Chervin, R. M., 1980: On the simulation of climate and climate change with general circulation models. Journal of the Atmospheric Sciences, 37, 1903-1913.

Chervin, R. M. and S. H. Schneider, 1976: On determing the statistical significance of climate experiments with general circulation models. Journal of the Atmospheric Sciences, 33, 405-412.

Der Megreditchian, G., 1979: Méthodes statistiques en climatologie. Aspects théorique et pratique. Note technique de l'EERM, 38.

Katz, R. W., 1982: Statistical evaluation of climate experiments with general circulation models: a parametric time series modeling approach. Journal of the Atmospheric Sciences, 39, 1446-1455.

Kendall, M. and A. Stuart, 1979: The advanced theory of statistics. Volume 2, Inference and relationship. Charles Griffin and Co. Ltd.

Laurmann, J., A. and W. L. gares, 1977: Statistical considerations in the evaluation of climatic experiments with atmospheric general circulation models. Journal of the Atmospheric Sciences, 34, 1187-1199.

Royer, J.-F., 1980: Présentation d'un modèle spectral de circulation générale à faible résolution. Expériences préliminaires avec un chauffage newtonien. Note technique de l'EERM, 83.

Royer, J.-F., M. Deque, H. Canetti, and M. Boulanger, 1981: Présentation d'un modèle spectral de circulation générale à faible résolution. Simulation du climat de janvier. Note de travail de l'EERM, 16.

Storch, V. S., 1982: A remark on Chervin-Schneider's algorithm to test significance of climate experiments with GCM's. Journal of the Atmospheric Sciences, 39, 187-189.

ON THE INTERPRETATION AND ESTIMATION OF EFFECTIVE SAMPLE SIZE

H.J. Thiébaux

Department of Mathematics, Statistics,
and Computing Science
Dalhousie University
Halifax, Nova Scotia
B3H 4H8 Canada

F.W. Zwiers

Department of Mathematics
University of Saskatchewan
Saskatoon, Saskatchewan
S7N OWO Canada

1. INTRODUCTION

Statistical and dynamical relationships between observed values of a geophysical system or model effectively reduce the number of independent data. The reduction is expressible in terms of the covariance structure of the process and, in some instances, it is reasonable to devise a measure of the "effective sample size" in terms of sample statistics. Here we discuss the concept of "effective sample size"; and, having settled upon one of several possible definitions, examine various methods of estimating this quantity. It is found that "effective sample size" is quite difficult to estimate reliably. However a procedure is described which the authors feel could be used successfully; and it is noted that the concept can be extended to spatial arrays of data.

The concept of "effective number of degrees of freedom" or "effective sample size" (ESS) was discussed in a paper by Laurmann and Gates (1977). However, there and in its use elsewhere, the quantity is not uniquely defined; and this has lead to some confusion about its correct interpretation. We believe that a careful examination of this concept and of alternative formulations for its evaluation and some key statistical properties of these estimates, may be helpful to researchers who wish to use it in the analysis of time series or spatially coherent processes.

In Section 2 we derive a measure of effective sample size, as it has previously been defined. In Section 3 we discuss phraseology which has been used to describe the ESS and argue that the usual formulation for its evaluation is only one of a family of possible estimators. Section 4 describes three methods of estimating ESS and Section 5 compares these methods via the results of Monte Carlo experiments. The final section considers and rejects the use of ESS in testing for the mean of a time series, and proposes an alternative test.

2. DERIVATION OF "EFFECTIVE SAMPLE SIZE" FOR A TIME SERIES.

Non-negligible correlations between successive values of a stochastic process preclude independence of the components of a vector of any N sequential observations. This is a concomitant of predictability and its measure, either in terms of Leith's (1973) "time between effectively independent samples" or the reciprocal "effective number of independent samples" in a fixed time span; and the latter provide useful means of quantifying the degree of time coherence in the process. The "time between effectively independent samples" or simply the "effective sample size" which we denote by ESS, is derived by equating the ensemble mean square of a time-averaged mean, say $\sigma^2(\bar{X})$, to the standard formula for the variance of the mean of independent samples. The solution for the so-called number of independent samples is one measure of ESS. In this derivation the ensemble mean square is determined by the autocovariance function (acf), which thus gives explicit expression to time coherence of successive values, as shown in the following.

For N time-sequential observations, say $X_1,\ldots,X_n$, of a statistically stationary stochastic-dynamic process with variance σ^2, the ensemble mean square of the observed mean maybe written in terms of the variance and autocorrelation function. Following Anderson (1971, p. 448)

$$\sigma^2(X) = \langle[\bar{X}-\mu]^2\rangle = \langle[\sum_{i=1}^{N}(X_i-\mu)/N]^2\rangle$$

$$= \sum_{i,j=1}^{N}\langle(X_i-\mu)(X_j-\mu)\rangle/N^2$$

$$= \sum_{i,j=1}^{N} C(i-j)/N^2$$

$$= \sum_{\tau=-(N-1)}^{+(N-1)}(N-|\tau|)C(\tau)/N^2$$

$$= \sigma^2\sum_{\tau=-(N-1)}^{+(N-1)}(1-\frac{|\tau|}{N})\rho(\tau)/N$$

where $C(\tau)$ is the symmetric covariance between observations lagged by a time interval of length τ, and $\rho(\tau)$ is the corresponding lag-correlation function. If in fact $\bar{X}$ were the mean

of N' independent sample values, this would be σ^2/N'. Hence by equating $\sigma^2(\bar{X})$ and σ^2/N' we obtain a measure of the effective number of independent samples:

$$N' = \sigma^2/\sigma^2(\bar{X}) = N/\sum_{\tau=-(N-1)}^{+(N-1)} (1 - \frac{|\tau|}{N})\rho(\tau). \quad (2.1)$$

3. INTERPRETATION OF "EFFECTIVE SAMPLE SIZE" FOR A TIME SERIES.

One expression in particular which has been used to describe the ESS is misleading; and it is the phrase "equivalent number of degrees of freedom" which is sometimes associated with N'. This is misleading because, regardless of the extent of time coherence, a sample of N observations from any nonsingular Gaussian stochastic process has N degrees of freedom; and a suitable linear transformation of the observation vector, for which the transformation is determined by the covariance structure of the process, has N independent components. The phrase "effective sample size" is less troublesome; and it manages to convey the notion that the N pieces of information in the sample are "smeared" across the N observations by the time-duration of their influences.

Phraseology used in the literature may appear to imply that the ESS is uniquely defined. However there are several equally valid functional definitions which we will examine here. Expression (2.1) above was derived by equating the variance of the time average of a sample of N observations from a correlated time series with the variance of a sample of N' observations from an uncorrelated series. However there are other statistical characteristics of sampled sequences whose comparison could be used analogously to generate a measure of "effective sample size". For example, one could equate the expected number of runs above and below the mean, in a sample of length N from a correlated time series, with that for a sample of length N' from an uncorrelated series. In the latter case, given that the sample contains N_1' observations above the mean it can be shown that the expected number of runs R is

$$E(R) = 2N_1'(N'-N_1')/N' + 1$$
$$= 2N'\alpha(1-\alpha) + 1$$

where α is the proportion of observations above the mean. For large samples this can be simplified to

$$E(R) = N'/2 + 1$$

since α converges to 0.5 with probability one, as N' goes to infinity. By equating this expression with the expected number of runs for a sample of length N from a correlated time series and solving for N', we get a different expression for ESS.

Since it is not possible to derive the expected number of runs for samples from correlated time series analytically, even for the simple case of a Markov process, we illustrate the non-equivalence of the two definitions of ESS with the results of a Monte Carlo experiment. Ensembles of 5000 samples of length 90 and 180 were created from a simulated Markov process, with a lag-one correlation coefficient of 0.75. These samples were used to estimate the expected number of runs above and below the mean. The results are summarized in Table 1; and it is clear that the two techniques for evaluating ESS yield substantially different results. Thus the technique for equating characteristics of sample statistics for samples from correlated time series with those for samples from uncorrelated series does not yield a unique measure. Indeed, there are still other characteristics which could reasonably be compared to generate measures of "effective sample size". In the remainder of this paper we consider only the quantity defined by (2.1) which we arrived at by equating variances, and we reserve the term "effective sample size" and its abbreviation ESS to denote it.

4. ESTIMATING THE 'EFFECTIVE SAMPLE SIZE"

In practice the parameters σ^2 and $C(\tau)$, and therefore $\rho(\tau)$ as well, are generally not known. Various approaches to the estimation of these parameters, and hence of N', are available. However, as demonstrated in the following, none of them is entirely satisfactory.

The most straight forward approach is to estimate σ^2 and $C(\tau)$ directly from the data. Since the ensemble mean of the stochastic process will generally be unknown, the sample mean $\bar{X}$ is substituted for μ in the computational formulas for covariance estimation, as follows:

$$\hat{C}(\tau) = \frac{1}{N-\tau} \sum^{N-\tau} (x_i - \bar{x})(x_{i+\tau} - \bar{x})$$

and for the sample variance $S^2 = \hat{C}(0)$. However these estimates of lag-covariance $C(\tau)$ and of variance σ^2 are not unbiased. Following Anderson (1971, pp. 438-464) the statistical ensemble average of the lag-τ covariance estimate is

$$\langle C(\tau)\rangle = C(\tau) - \frac{1}{n}\{C(0) + 2\sum_{r=1}^{\tau}[1 - \frac{r\tau}{N(N-\tau)}]C(r)$$
$$+ 2\sum_{r=\tau+1}^{N-\tau-1}[1 - \frac{r\tau}{N(N-\tau)} - \frac{r-\tau}{N-\tau}]C(r)$$
$$+ 2\sum_{r=N-\tau}^{N-1}\frac{(N-r)\tau}{N(N-\tau)}C(\tau)\}$$

for τ in the interval $(1, N-\tau-1)$, and a comparable expression for $\tau > (N-1)/2$. If it were possible to increase the length of the observed sequence indefinitely, the ensemble means of the sample covariances would converge to the corresponding values of the autocovariance function, provided the series $\sum_{-\infty}^{\infty} C(\tau)$ converges. Thus $\hat{C}(\tau)$

is asymptotically an unbiased estimate of $C(\tau)$ for all time lags. However, for an observation period which is likely to be significantly non-infinite with respect to the time coherence of the observed process, it is not realistic to assume that the bias factor is near zero. Accordingly, it appears that $\sigma^2(X)$ and σ^2, and hence N', cannot be reliably estimated by sample covariance and variance values.

A second possible approach is through the use of a time series model. For example, an autoregressive moving average (ARMA) model such a Box and Jenkins (1976) describe, may be fitted to the sequence of N observations and the auto-covariance function of the fitted model used to estimate N'. However, once again, the estimates of the coefficients of the ARMA model are only asymptotically unbiased. For finite observation sequences, the estimated autocovariance function will be biased and, again, N' will be difficult to estimate.

Finally, we consider estimating ESS from the power spectrum of the observed sequence. To see how this may be achieved we note that for large sample sizes we have the following approximation for the variance of the sample mean:

$$\sigma^2(\bar{X}) \approx 2\pi\, f_{XX}(0)/N$$

where $f_{XX}(\lambda)$ is the spectral density function of the observed time series. Thus, if we equate the right-hand-side with the expression for the variance of N' independent observations, and solve for N', we obtain

$$N' \approx N\sigma^2/2\pi f_{XX}(0)$$

and the problem of estimating N' reduces to estimating the spectrum of the process at the origin. Unfortunately there will be biases in this estimation; and a tradeoff must be made between the inherent variability of the spectral estimator and its bias (See Jones, 1975, or Koopmans, 1976, for discussions of the consequences of using smoothed periodogram-, or similar spectral-estimators). When the true spectrum has a relative maximum at the origin, such as that of a red noise process, we should expect the estimator of $f_{XX}(0)$ to be negatively biased and hence estimates of N' to be positively biased. On the other hand, when the true spectrum has a trough at the origin, such as those of many second-order processes which are important for explaining spatial interactions in the atmosphere (Thiebaux, 1976 & 1981), we should expect the estimator of $f_{XX}(0)$ to be positively biased and hence the resulting estimates of N' to be negatively biased.

It is clear that the problem of estimating ESS is embedded in the problem of estimating the variance of the time averaged mean, which is considered by Jones (1975). In this connection Jones discusses various methods of fitting time series models and estimating the spectrum at the origin. Although any method which yields acceptable estimates of the variance of the time averaged mean might also be expected to yield acceptable estimates of ESS, this bridge must be approached with caution: Amounts of bias which would be relatively inconsequential when making estimates of the variance, result in considerable bias in estimates of ESS because they are proportional to the inverses of those variance estimates.

5. MONTE CARLO EXPERIMENTS.

Monte Carlo experiments were carried out with simulated first-and second-order autoregressive processes for which the corresponding spectral density functions have a relative maximum and a relative minimum at the origin, respectively. The models chosen were the following first-order autoregression

$$x_t = 0.75X_{t-1} + \epsilon_t \qquad (5.1)$$

which has a "time between effectively independent samples" of about 7 time units, and the second-order autoregression

$$s_t = 1.2X_{t-1} - 0.6X_{t-2} + \epsilon_t \qquad (5.2)$$

for which the "time between effectively independent samples" is 1.76 time units, by rigorous computation from the (true) model properties. In these expressions the ϵ_t represent sequences of uncorrelated random variables. The spectral density functions for these models are shown in Figure 1.

The experiments were conducted by generating 1000 sequences of length N, for each model, for various values of N. ESS was estimated for each sequence by the three approaches discussed above; and Tables 2.6 describe the outcomes in terms of parameters of the Monte Carlo distributions thus generated. In each table the true "effective sample sizes", analytically derived with the corresponding models and lengths of sequences, are given in the column headed "ESS"; and estimated ESS are compared for sequence lengths: N = 30, 60, 90, 120, and 240. Table 2 presents model values for the 1000 sequences generated, Table 3 the medians, Table 4 the means of the distributions, Table 5 standard deviation values, and Table 6 RMS errors.

The estimator of ESS based on direct estimated of the autocovariance function is denoted by DIRECT.

The estimator utilizing the autocovariance functions of autoregressive moving average models which have been fitted to generated sequences, are denoted by ARMA. In these evaluations, autoregressive models were fitted to sample sequences using the method of least squares and Akaike's minimum AIC criterion (Akaike 1973, 1976) was used to determine the order of the process for which observations were simulated. The AIC criterion is described by

$$AIC = N \ln \hat{\sigma}_\epsilon + 2p$$

where p is the order of the autoregressive model which is fitted and $\hat{\sigma}_\epsilon^2$ is the estimated variance of the random uncorrelated innovations of the model. Thus the fitted model essentially minimizes the sum of squared errors, except that a penalty is effectively imposed for "extravagent use of parameters". The program UNIMAR from the TIMSAC-78 package (Akaike, 1979) was used to find the best fitting model satisfying the AIC criterion. For both types of sequence generation, i.e. whether the observations were simulated for an AR(1) or an AR(2) process, we allowed the UNIMAR program to select amongst models of orders up to five.

Estimates of ESS based on estimates of the power spectrum at the origin are denoted by POWER SPECTRUM. We present four sets of values for this approach to the estimation of ESS, corresponding to four different methods of estimating the spectrum at the origin: "SPEC1" and SPEC5" are based on the Daniel, or smoothed periodogram estimator of the spectrum. In this case $f_{XX}(0)$ is estimated simply as

$$f_{XX}^{(N)}(0) = \frac{1}{m} \sum_{j=1}^{m} I_{XX}^{(N)}(2\pi j/N)$$

where $I_{XX}^{(N)}(2)$ is the periodogram, namely, the squared modulus of the finite fourier transform of the observations. m is the number of ordinates of the periodogram which are averaged together, with $m = 1$ for SPEC1 and $m = 5$ for SPEC5. At the expense of a small increase in variance, the bias of the Daniel estimator can be reduced by employing a data window or "taper". Estimator "SPECTS" used the Daniel estimator with 20 percent of the observations tapered with a cosine taper. This is to say that the X_t observations are transformed as follows:

$$Y_t = b_t X_t \ , \ t = 1,\ldots,N \ ,$$

where $b_t = b(t/N)$ for

$$b(\nu) = \begin{cases} \frac{1}{2}\left[1-\cos(\pi\nu/a)\right] & , \ 0 \le \nu \le a \\ 1 & , \ a \le \nu \le 1-a \\ \frac{1}{2}\left[1-\cos(\pi(1-\nu)/a)\right] & , \ 1-a \le \nu \le 1 \\ 0 & , \ \text{otherwise} \end{cases}$$

The method designated by "BART" uses a weighted covariance estimator with a Bartlett lag-window, for which the estimator of $f_{XX}(0)$ is given by

$$f_{XX}^{(N)}(0) = \frac{1}{2\pi} \sum_{\tau=-(M-1)}^{(M-1)} (1-|\tau|/M) C_\tau \ .$$

Here M determines the "bandwidth" or resolution of the estimator. Although a number of different lag-windows might have been used, all those with essentially the same bandwidth would give essentially the same estimates of $f_{XX}(0)$ (See Jenkins and Watts, 1968). The bandwidth for the BART estimator is

$$\beta = 3\pi/M \ .$$

In our experiments we choose M to be the square root of sequence length N, thereby insuring that the bandwidth of the estimator and therefore its bias, decrease with increasing N. Since the bandwidth of weighted covariance estimators is generally relatively wide compared to that of Daniel estimators, there is only slight reduction in bias when a taper is used (See Koopmans, 1974). Consequently we did not employ a taper with this estimator.

As can be seen from Table 2 - 5, the DIRECT estimates of ESS are extremely unreliable in the sense that they have large standard deviations for both first- and second-order generating processes and all sequence lengths; and their medians and means differ substantially from the true ESS values. Although the mode values are not greatly in error, their comparison with median and mean values show the distributions of DIRECT estimates to be strongly skewed to the right.

SPEC1 estimates behave similarly to DIRECT estimates: their distributions are skewed to the right and have large variances. Here this behavior can be explained in terms of the method of estimation, since an unsmoothed (m=1) periodogram estimate of the spectrum at the origin is used. The periodogram ordinate at frequency $2\pi/N$, which is the one used to estimate $f_{XX}(0)$ when $m=1$, is asymptotically proportional to a random variable with exponential distribution (See Brillinger, 1981). Correspondingly estimates of $f_{XX}(0)$ near zero and, thus also, very large estimates of ESS, are relatively frequent. SPEC5 estimates which are based on the smooth periodogram with $m=5$, are better behaved than the two sets of estimates just described. Nonetheless SPEC5 does not provide highly reliable estimates: In most cases the RMS errors shown in Table 6 are more than 50 percent of true ESS values.

NB. In Tables 3 and 4 we see evidence of the effects of the biases in the estimates of $f_{XX}(0)$: In the AR(1) case we see that the bias is positive (and quite large, in relative terms) for small sequence lengths, while in the AR(2) case the bias is negative for short sequences.

The SPECT5 estimator which employs the smoothed periodogram estimator with $m=5$ and a cosine taper is disappointing in that any reduction in bias in the estimates of $f_{XX}(0)$ seems to have been more than compensated by an increase in variance. The BART estimator, which uses the weighted covariance estimator of $f_{XX}(0)$, has a mixed performance in the senses that: (1) when the observations were generated from the AR(2) process, the estimator produced estimates of ESS with very low biases and low variances, and (2) RMS errors of estimation ranged from 20 to 29 percent of the true ESS. However, the estimator did not perform as well when observations were

generated from the AR(1) process. As can be seen from Figure 1, this is attributable to the sharp peak of the AR(1) spectrum at the origin, which makes it difficult to estimate $f_{XX}(0)$ accurately.

The ARMA estimator which employs an auto-covariance function derived from fitting time series models appears to be the best of those considered. The biases are about the same as those of the SPEC5, SPECT5, and BART methods, as can be seen by comparing the entries of Table 3 and 4. Table 5 shows that the variances of the estimates remain fairly constant as a function of sequence length: a desirable behavior which can be attributed to the fact that fitting time series models provides consistent estimators of the true models used to generate the data.

6. A TEST FOR THE MEAN OF A TIME SERIES.

One place in which the ESS has been employed or suggested for use is in connection with the evaluation of climate change experiments with general circulation models. Clearly, in order to evaluate the outcomes of such experiments it is necessary to have a statistical procedure for testing hypotheses concerning means of sequences of time-correlated observations. In particular, Laurmann and Gates (1977) proposed the statistic

$$Y = \frac{\sqrt{\hat{N}'}\,(\bar{X}-\mu_0)}{S} \qquad (6.1)$$

to test an hypothesis

$$H_0 : \mu = \mu_0 .$$

In (6.1) $\hat{N}'$ is an estimate of the ESS, $\bar{X}$ is the sample mean, and S^2 is the sample variance which is given by

$$S^2 = \sum_{j=1}^{N} (X_j-\bar{X})/N .$$

Terminology which has been used to describe ESS, suggests that a sample of length N contains just N' independent pieces of information and, therefore, that Y should be distributed as a Student's t variate with N'-1 degrees of freedom when $\mu = \mu_0$. However, the distribution of the statistic given by (6.1) can be markedly different from the Student's t distribution. To illustrate this we conducted a Monte-Carlo experiment by generating samples of 1000 Y values using SPEC5 estimate of ESS. Values of the test statistic were computed from samples of various lengths taken from both the AR(1) and AR(2) models described previously. To check the t-distribution conjecture we estimated the probability of rejecting null hypothesis H_0 , when true, with a test conducted at the NOMINAL 5% significance level. That is, we estimated the true significance level of the test which is given by

$$\alpha' = P(|Y| > t_{(\hat{N}'-1);0.975})$$

where $t_{(\hat{N}'-1),0.975}$ is the point on the Student's t distribution with N'-1 degrees of freedom, which is exceeded with probability 0.025 . That is,

$$P(|t| > t_{(\hat{N}'-1),0.975}) = 0.05$$

when t truely does have Student's t distribution with N'-1 degrees of freedom. The results of the experiment are summarised in Table 7. Similar results can be obtained if other estimators of ESS are employed. In general we see that the distribution of statistic (6.1) is different from that conjectured and that its distribution depends upon how ESS is estimated. Regardless of the method of estimation, it would be preferable to compare the value of the test statistic against a critical value chosen from its <u>true</u> asymptotic distribution, where the latter may be determined from the N' estimation equation. In the case of the SPEC5 estimator this is

$$N' = N\sigma^2/2\pi f_{XX}^{(m)}(0) \quad \text{with} \quad m=5$$

so that

$$Y = \frac{\sqrt{N}(\bar{X}-\mu_0)}{\sqrt{2\pi f_{XX}^{(m)}(0)}}$$

which is asymptotically distributed as Student's t with <u>2m</u> degrees of freedom (Brillinger, 1981). With this approach we may be certain of having the test's limiting characteristics correct (at least).

7. IN CONSIDERING A SPATIAL ANALOGUE

We would like to point out that the concept of "effective sample size" is easily extended to spatial arrays of observations. For example, a numerical forecast model operates coherently on a global or hemispheric scale so that, for simultaneous output at several points of a region, the dynamical coherence of the model translates into spatial coherence of climate-outcome-values and of statistics used to assess their significance. By analogy with the time-coherent outcomes treated above, spatial coherence of the process effectively reduces the number of "independent" values at the grid-points of a geographically designated grid. When the collective impact of spatially arrayed climate experiment outcomes is to be assessed by averaging over a region rather than through time, determination of the effective number of independent values in a spatial array is relevant.

There is essentially nothing in the foregoing that restricts its applications to outcomes of an experiment which are sequential in time. Again, since the usual methods of estimating the spatial correlation structure of an observed field involve removal of the mean field, the bias problems discussed above will be present. The results of the experiments described in Section 5 suggest that much better estimates of ESS for

spatial arrays could be made by modelling the spatial correlation structure of the observed field as has been done by Thiébaux (1976).

7. REFERENCES.

Akaike, H., 1973: Maximum likelihood identification of Gaussian autoregressive moving average models. Biometrika, 60, 255-265.

Akaike, H., 1976: Canonical correlation analysis of time series and the use of an information criterion. System Identification: Advances and case studies. Edited by R.K. Mehrs and O.G. Lainiotis, Academic Press.

Akaike, H., G. Kitasawa, E. Arahata and F. Tada, 1979: TIMSAC-78, Computer Science Monographs, No. 11, Inst. Stat. Math.

Anderson, T.W., 1971: The sample mean, covariences and spectral density. Ch. 8 in The Statistical Analysis of Time Series. Wiley, 438-500.

Box, G.E.P., and G.M. Jenkins, 1976: Time series analysis: forecasting and control. Revised Edition. Holden-Day.

Brillinger, D.R., 1981: Time series: data analysis and theory. Expanded Edition. Holden-Day.

Jenkins, G.M., and D.G. Watts, 1968: Spectral Analysis and its Applications. Holden-Day.

Jones, R.H., 1975: Estimating the variance of time averages. J. Appl. Met., 14, 159-163.

Jones, R.H., 1976: On estimating the variance of time averages. J. Appl. Met., 15, 514-515.

Koopmans, L.H., 1974: The Spectral Analysis of Time Series. Academic Press.

Laurmann, J.A. and W.L. Gates, 1977: Statistical considerations in the evaluation of climatic experiments with atmospheric general circulation models. J. Atmos. Sci., 34, 1187-1199.

Leith, C.E., 1973: The standard error of time-average estimates of climate means. J. Appl. Meteor.,12, 1066-1069.

Thiébaux, H.J., 1976: Anisotropic correlation functions for objective analysis. Mon. Wea. Rev., 104, 994-1002.

Thiébaux, H.J., 1981: The kinetic energy spectrum visavis a statistical model for geopotential. Tellus, 33, 417-427.

Thiébaux, H.J. and F.W. Zwiers, 1981: Some alternative approaches to testing for the significance of differences in climate model outcomes. Preprint, Seventh Conference on Probability and Statistics in Atmospheric Science, Nov. 2-6, Monterey, California.

Zwiers, F.W., 1981: Simulation experiments with a test for the mean of a time series for use in climate in climate change experiments. Preprint, Seventh Conference on Probability and Statistics in Atmospheric Science, Nov. 2-6, Monterey, California.

ACKNOWLEDGEMENTS.

Research by H.J. Thiébaux was supported by Natural Sciences and Engineering Research Council (NSERC) of Canada Grant A9182. Research by F.W. Zwiers was supported by NSERC Grant A5448 and by the Atmospheric Environment Service of Canada.

N	E(R)	95% C.I.	N	N'(V)
90	22.46	(22.32 , 22.6)	43	13
180	42.95	(42.76 , 43.14)	84	26

Table 1

Point and interval estimates of E(R) based on 5000 samples of length N from a Markov process with lag-one correlation coefficient 0.75. The "runs comparison" value of the ESS is given in column 4 and the "variance comparison" value, derived analytically using expression (2.1) is given in column 5 .

			ESTIMATED ESS					
Model	N	ESS	DIRECT	ARMA	POWER SPECTRUM			
					SPEC1	SPEC5	SPECTS	BART
AR(1)	30	5	8	7	3	12	12	11
	60	9	11	10	6	14	13	--
	90	13	14	12	6	17	15	--
	120	18	16	18	9	19	17	29
	240	35	29	38	--	30	35	44
	30	16	14	19	8	11	11	13
	60	34	36	31	15	20	21	26
	90	51	--	47	38	37	41	43
	120	68	70	--	61	49	70	56
	240	136	--	149	--	101	184	--

Table 2

Mode of 1000 estimates of the "effective sample size" made with six different estimators. The true ESS is given in the column headed "ESS".

			ESTIMATED ESS					
Model	N	ESS	DIRECT	ARMA	POWER SPECTRUM			
					SPEC1	SPEC5	SPECTS	BART
AR(1)	30	5	19	8	7	13	13	12
	60	9	33	12	11	15	15	18
	90	13	49	16	17	19	18	24
	120	18	62	20	22	22	22	29
	240	35	117	37	46	39	38	48
	30	16	39	30	17	11	11	13
	60	34	102	38	41	23	24	28
	90	51	151	55	72	43	45	47
	120	68	210	73	94	63	65	64
	240	136	448	139	190	141	147	133

Table 3

Median of 1000 estimates of the "effective sample size" made with six different estimators. The true ESS is given in the column headed "ESS".

			ESTIMATED ESS					
Model	N	ESS	DIRECT	ARMA	POWER SPECTRUM			
					SPEC1	SPEC5	SPECTS	BART
AR(1)	30	5	48	15	80	13	13	13
	60	9	73	14	80	16	16	19
	90	13	89	17	117	20	20	25
	120	18	114	21	88	25	25	30
	240	35	213	38	222	43	43	50
AR(2)	30	16	67	26	55	11	12	14
	60	34	150	43	201	27	27	30
	90	51	228	60	309	50	52	50
	120	68	314	78	382	73	75	68
	240	136	513	145	1196	162	171	138

Table 4

Mean of 1000 estimates of the "effective sample size" made with six different estimators. The true ESS is given in the column headed "ESS".

			ESTIMATED ESS					
					POWER SPECTRUM			
Model	N	ESS	DIRECT	ARMA	SPEC1	SPEC5	SPECTS	BART
AR(1)	30	5	82.3	63.6	1402	2.7	3.1	4.0
	60	9	117.5	9.3	1342	4.4	5.2	4.8
	90	14	121.0	7.8	882.2	6.5	7.5	6.3
	120	16	160.8	8.3	355.4	9.5	10.5	6.8
	240	29	290.3	9.7	867.0	19.7	20.0	11.6
AR(2)	30	16	103.8	24.9	162	1.1	2.5	2.5
	60	34	140.4	24.0	939	15.0	11.6	7.0
	90	51	224.2	26.2	1698	28.7	29.6	15.0
	120	68	323.6	28.7	1866	38.2	38.2	19.8
	240	136	692.5	37.5	1476	87.0	99.7	32.6

Table 5

Standard Deviation of 1000 estimates of the "effective sample size" made with six different estimators. The true ESS is given by the column headed "ESS".

			ESTIMATED ESS					
					POWER SPECTRUM			
Model	N	ESS	DIRECT	ARMA	SPEC1	SPEC5	SPECTS	BART
AR(1)	30	5	18.6	12.88	281	1.69	1.92	1.79
	60	9	14.9	1.17	149	0.92	0.97	1.23
	90	14	11.0	0.67	68.3	0.73	0.79	1.04
	120	16	10.4	0.49	20.1	0.66	0.70	0.77
	240	29	9.73	0.29	25.4	0.61	0.62	0.54
AR(2)	30	16	7.23	1.68	10.4	0.32	0.29	0.20
	60	34	5.34	0.75	28.1	0.49	0.40	0.24
	90	51	5.60	0.54	33.7	0.56	0.58	0.29
	120	68	5.98	0.45	27.8	0.57	0.57	0.29
	240	136	5.90	0.28	13.4	0.67	0.78	0.24

Table 6

RMS Error of 1000 estimates of the "effective sample size" computed as a proportion of the time ESS. The true ESS is given by the column headed "ESS".

Sample Length	Data Source AR(1)	AR(2)
30	.288	.042
60	.175	.040
90	.140	.070
120	.108	.063
240	.104	.061

Table 7

Proportion of 1,000 samples for which $|y| > t_{\hat{N}'-1, 0.975}$ where $\hat{N}'$ is estimated using the SPEC5 estimator of the ESS.

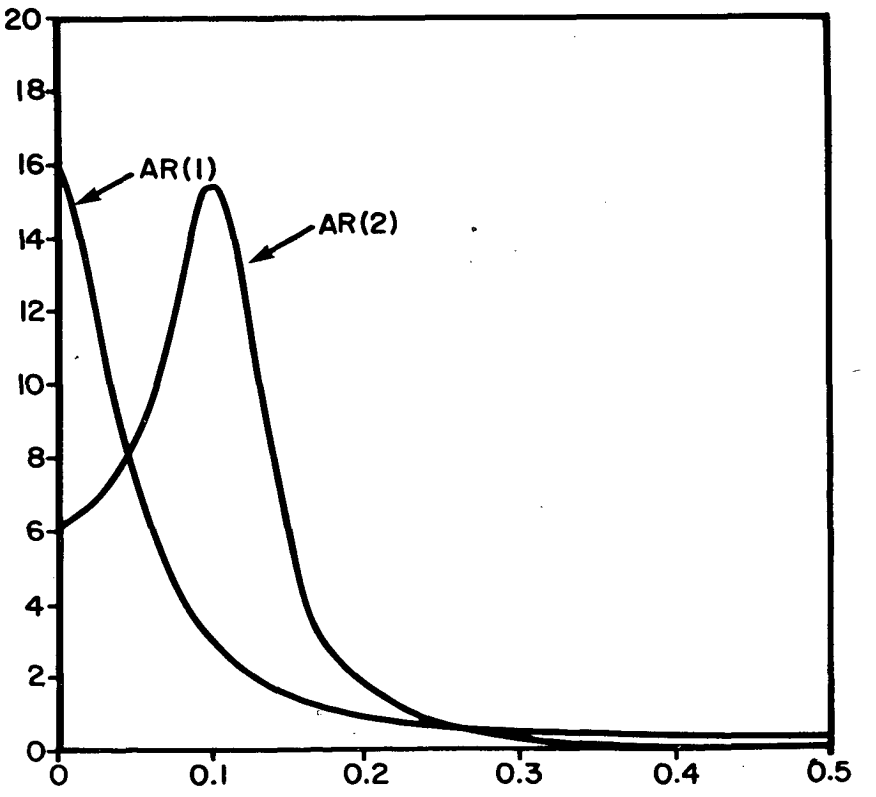


Figure 1.1.

Power spectra of the AR(1) process: $x_t = 0.75x_{t-1} + \epsilon_t$, and the AR(2) process: $x_t = 1.2x_{t-1} - 0.6x_{t-1} + \epsilon_t$, with $\sigma^2_\epsilon = 1.0$. Note the relative maximum at the origin of the s.d.f. for the AR(1) process and the relative minimum at the origin for the AR(2) process.

EVALUATION OF THE STATISTICAL SIGNIFICANCE OF CLIMATIC CHANGES

Chr. GOOSSENS and A. BERGER

Institut d'Astronomie et de Géophysique
Georges Lemaitre
Université Catholique de Louvain
2 Chemin du Cyclotron
B-1348 Louvain-la-Neuve
Belgique

1. INTRODUCTION

Climatic change is difficult to define, let alone to detect and to measure. Clearly, one must trace a climatic element with a continuing trend, or a discontinuity in the average value or in the variability of some climatic element. Because climate is inherently variable on all time scales, it is not possible to ascribe univocally any observed variation to some specific cause, or to classify it, with complete confidence, as a manifestation of the natural variability. The question as to whether a given variation in climate is a statistically significant change as contrasted with a mnifestation of the normal "random" variability will often depend on the time and space scales under consideration.

Over the past 30 years, many authors (Mitchell, 1961, 63; Callendar, 1961; Jones et al., 1982) have analysed long term average temperature series based on instrumental data, using graphical analysis and performing a moving average (for example 5 or 10 years running means) in order to avoid short-term variations. These studies showed a warming between 1881 and 1940 and a subsequent cooling to mid or late 1960's with a period of little trend following.

In this paper, climatic change will be detected by means of appropriate statistical tests. These tests, first introduced by Sneyers (1957, 1963) are based on the fact that, under the hypothesis of a stable climate, the climatological series must behave has a simple random series, i.e. that the succesion of values must be independent and that the probability distribution must remain always the same. In order to check this hypothesis, two kinds of non-parametric tests will have to be applied : (i) the Sneyers' serial correlation coefficient (independance of the successive values) and (ii) a test of trend (the Mann-Kendall rank statistics).

The following data will be used : the long-term series of average Northern Hemisphere temperature, the very long European reconstructed series (De Bilt, Central England, Paris) and most of the instrumental Western European long term series.

2. DESCRIPTION OF THE STATISTICAL METHODS

In the study of inhomogeneities in climatological series, the various alternatives to randomness that are likely to be of concern have one property in common. Namely, their spectrum display significant relatively low frequencies which introduce positive serial correlations at small lags. In general, the values of the serial correlation at lag one will be more influenced by trend, periodicities, persistence or other forms of non-randomness, than the values of serial correlation at larger lags. This circumstance suggests the great potential of the lag-one serial correlation coefficient as a measure of non-randomness in general.

If we have the series of N terms X_i (i=1 ... N) which are deviate variates, the statistics of the test (Sneyers, 1975; Goossens, 1983) is given by :

$$r = \sum_{i=1}^{N} X_i X_{i+1} \Big/ \sum_{i=1}^{N} X_i^2$$

For large N, and if the series is random, r is normally distributed with an expected value of -1/N-1 and a variance of 1/N-1.

In testing the statistical significance of r for the null hypothesis of randomness, the one-tailed significance test will be used because most alternatives to randomness which are of concern in climatological series, would increase the value of r. The null hypothesis is thus rejected for large value of r. We should first determine by means of a table for a normal reduced variable, the value of r such as $\alpha_1 = P(u > u(r))$ where $u(r) = [(N-1)r+1] / \sqrt{N-1}$. The null hypothesis of randomness being accepted or rejected at the selected significance level α_0 if respectively $\alpha_1 > \alpha_0$ or $\alpha_1 < \alpha_0$.

It often happens that the most likely alternative to randomness in a climatological series is some form of trend. As a rule, there is no reason to assume that the trend in question is linear, so that it is advised to use tests of randomness that are nearly uniform power for alternatives of non-linear trends.

In order to use the Mann-Kendall rank statistics, we need to consider only the relative values of all terms in the analysed series. For each term X_i the number n_i of term X_j preceeding it ($X_j < X_i$) is computed and the sum t is performed (Mann, 1945; Sneyers, 1975; Goossens, 1983) :

$$t = \sum_{i=1}^{N} n_i \qquad (2)$$

For large N, under the hypothesis of no trend, t is normally distributed with an expected value of N(N-1)/4 and a variance of N(N-1)(2N+5)/72. In testing the statistical significance of t for the null hypothesis the two-tailed significance test is used, the null hypothesis being rejected for large value of $|u(t)|$. The statistic u(t) used to test the statistical significance of t is given by: $u(t)=t-E(t)/\sqrt{\text{var } t}$. The value of α_1 is then computed by means of a table for a normal reduced variable such as $\alpha_1=P(u > |u(t)|)$, the null hypothesis being accepted or rejected at the desired significance level α_0 for respectively $\alpha_1 > \alpha_0$ or $\alpha_1 < \alpha_0$. The null hypothesis is thus rejected for large value of $|u(t)|$. If u(t) is significant, we have an increasing or a decreasing trend for respectively $u(t) > 0$ or $u(t) < 0$. A sequential version of this test has been proposed by Sneyers (1963, 1975) and allows to detect the approximate beginning of the trend. It is based on the fact that the value of t for the i first term may be given by

$$t_i = n_1 + n_2 + \ldots + n_i$$

The corresponding value of $u(t_i)$ may be obtained using N=i in formula (2).

This principle may also be applied to the retrograde series. In this case for each term X_i the number n'_i of term X_i such as $X_i < X_j$ $(i<j)$ is computed. If we put $i'=(N+1)-i$ and $n'_i=n_{i'}$, the value of $u'_i=-u(t_{i'})$ with $u'_1=u_N$.

If the graphical representation of the values of the standardized values of $u(t_i)$ (t_1) and $u(t_{i'})(t_2)$ shows intersection of the curves of t_1 and t_2, it may conventionally be considered as the beginning of a climatic change.

3. NORTHERN HEMISPHERE TEMPERATURES

The long term series of average Northern Hemisphere temperatures were then investigated. The first one is the reconstructed series by Jones et al. (1982) based on monthly mean station data gridded on a 5° latitude by 10° longitude grid. The basic data set being monthly station data originally published in the World Weather Record and updated by NOAA in Monthly Climatic Data for the world. The others series used are the averaged temperatures series of Borzenkova et al. (1976), updated by Gruza and Ran'kova (1979) for selected latitudes (20°N, 40°N, 55°N, 65°N, 80°N and 20-80°N). The Russian data are apparently based on three series of charts. Data for 1881-1960 were processed by a team at the Main Geophysical Observatory (GGO) and these were published as monthly average temperature. Charts for the period 1881-1935 (Sokhrina et al., 1959) and as anomaly charts for the period 1881-1960 (Sharon, 1960-7). For 1961-75, anomaly charts have been published in the Synoptic Bulletin of the USSR Hydrometeorological Center.

Test of inhomogeneity shows that most of the series are inhomogeneous, the null hypothesis of the Sneyers' serial correlation test being rejected at the 0.05 significant level (Table 1).

The test of Mann is then applied to each series. Both of them, except the 20°N latitudes average temperature series which does not have any trend, present a general warming trend (u(t) being significant and positive).

Fig. 1 gives the graphical representation of the monthly temperature anomalies for the Northern Hemisphere, relative to the 1946-60 reference period, from 1811 to 1981 (lower part) and the values of the standardized values of u(t) (upper part), t_1 for the partial series beginning in 1881 and ending from year to year (1882,1883,...,1981), t_2 for the partial series ending in 1981 and beginning from year to year. Fig. 2 a to f has been constructed in the same manner as Fig. 1 for selected latitudes. For the average Northern Hemisphere temperatures series, the beginning of this long term warming trend which is continuous up to now, is observed around 1900. Superposed to it, we have long term oscillations which display a maximum around 1950 and minima around 1885 and 1975. Now, the average N.H. temperatures seems to increase again. This curve shows also short term oscillations, which seems to have a mean period of 5 years.

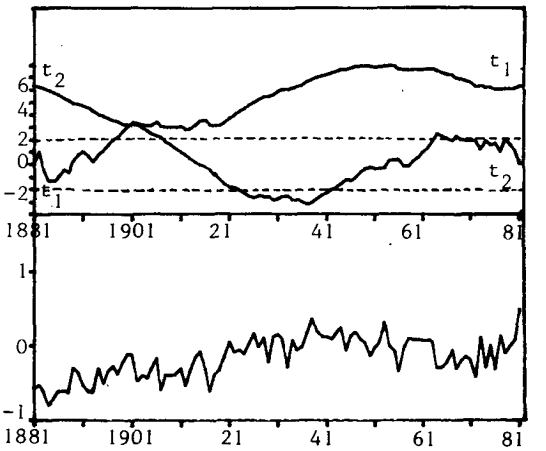


Fig. 1. Monthly temperature anomalies (°C for the Northern Hemisphere relative to the 1946-60 reference period (Jones et al., 1982). Values based on the trend test for partial series.

The middle latitudes series undergo this long term change around 1895, for higher latitudes (80°N) this change only appears around 1907. These series also present oscillations between minima and maxima values. For high latitudes, after a short fluctuation up to 1916, the average annual temperature rises until 1940 and then become nearly constant. For the middle latitudes we have a stable period from 1895 to 1920 followed by an increase of the temperatture up to 1940. The low latitudes present only short fluctuations over the whole period of observations. Each series (except for 20°N) shows a cooling since 1950.

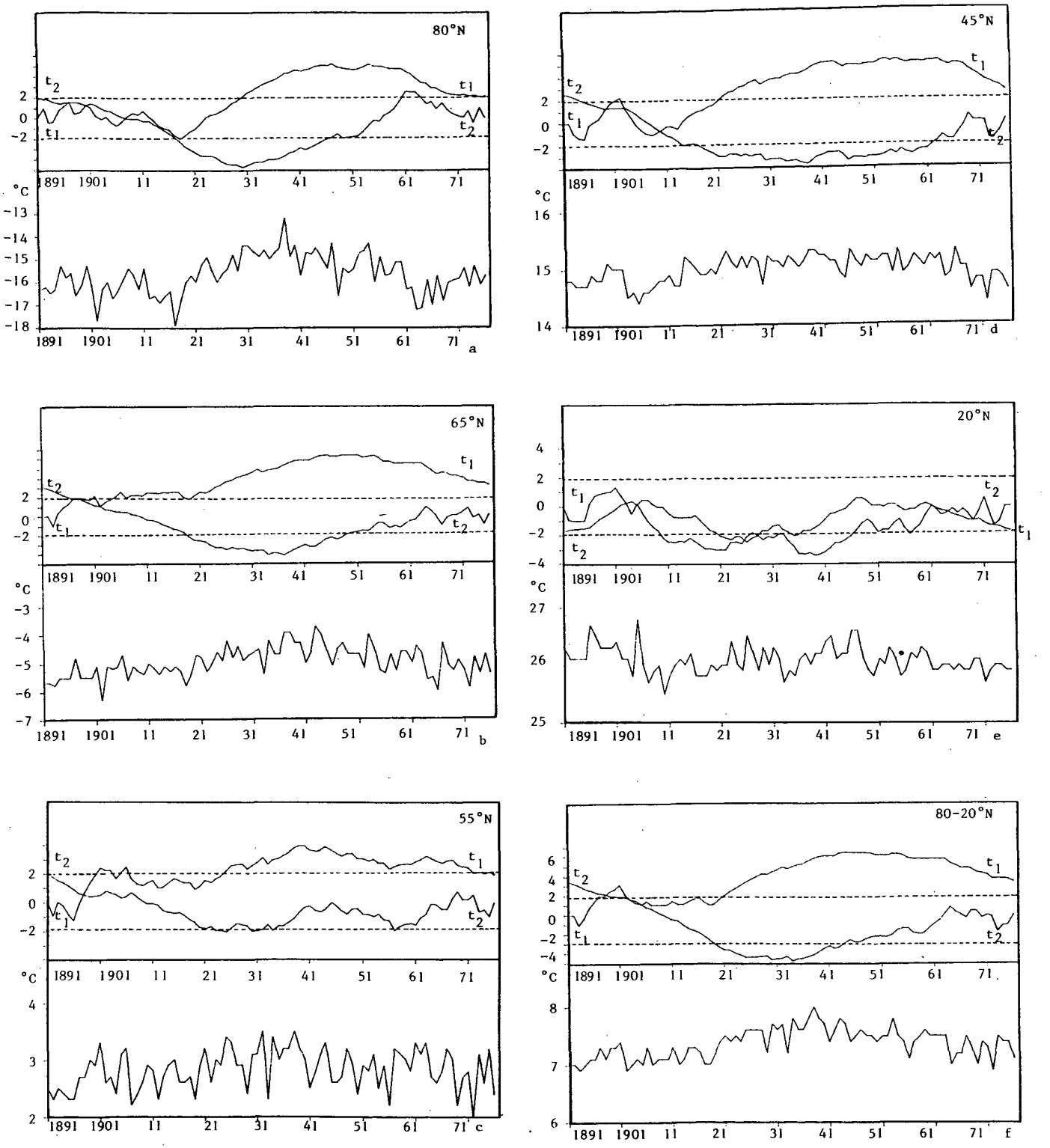


Fig. 2. a to f Mean annual air temperature values at selected latitudes for the period 1891-1976. Values based on the record test for partial series.

Table 1. Sneyers' serial correlation test. Probability of error α_i when H_0 : randomness is rejected for the analysed temperature series.

Station	r	α_i
Reconstructed series		
Central England	0.27	0.00
Paris	0.22	0.00
De Bilt	0.11	0.04
Average N.H. temperature series		
80°N	0.20	0.02
65°N	0.14	0.08
55°N	0.03	0.34
45°N	0.45	0.00
20°N	0.33	0.00
20-80°N	0.45	0.00
Western European temperature series		
Wien	0.29	0.00
Innsbruck	0.01	0.40
Paris	0.48	0.00
Nantes	0.23	0.09
Marseille	0.66	0.00
Potsdam	-0.02	0.48
Berlin	0.09	0.08
Frankfurt	0.00	0.46
Hohenpeissenberg	0.13	0.04
Athens	0.34	0.00
Milano	0.96	0.00
Bodo	0.16	0.04
Lisboa	0.59	0.00
Madrid	-0.28	0.07
Palma	0.21	0.09
Aberdeen	0.05	0.27
Uppsala	0.03	0.32
Bâle	0.74	0.00
Zurich	0.34	0.00
Santis	0.06	0.23
Genève	0.31	0.00

Similar studies have been performed for January and July. For January, the high polar latitudes show a warming trend since 1920, the 65°N latitudes since 1910, the 55°N latitudes since 1897. The 45°N and 20°N latitudes do not show any pronounced revearsal trend. For July, the high polar latitudes show a warming trend around 1960, 65°N latitudes around 1930, 55°N and 45°N around 1912, while temperature in low latitudes remains nearly constant.

4. CENTRAL ENGLAND, PARIS AND DE BILT SERIES

Three of the longest series of instrumental temperature records existing in the world are from Central England (1659-1973), Paris-France (1680-1980) and De Bilt, the Netherlands (1735-1981). The reconstruction of these time series are comprehensively described in papers, respectively, by Manley (1956, 1973), Dettwiller (1981) and Labrijn (1944). There is no use in trying to summarize these papers, except for the information considered in Fig. 3 a to c.

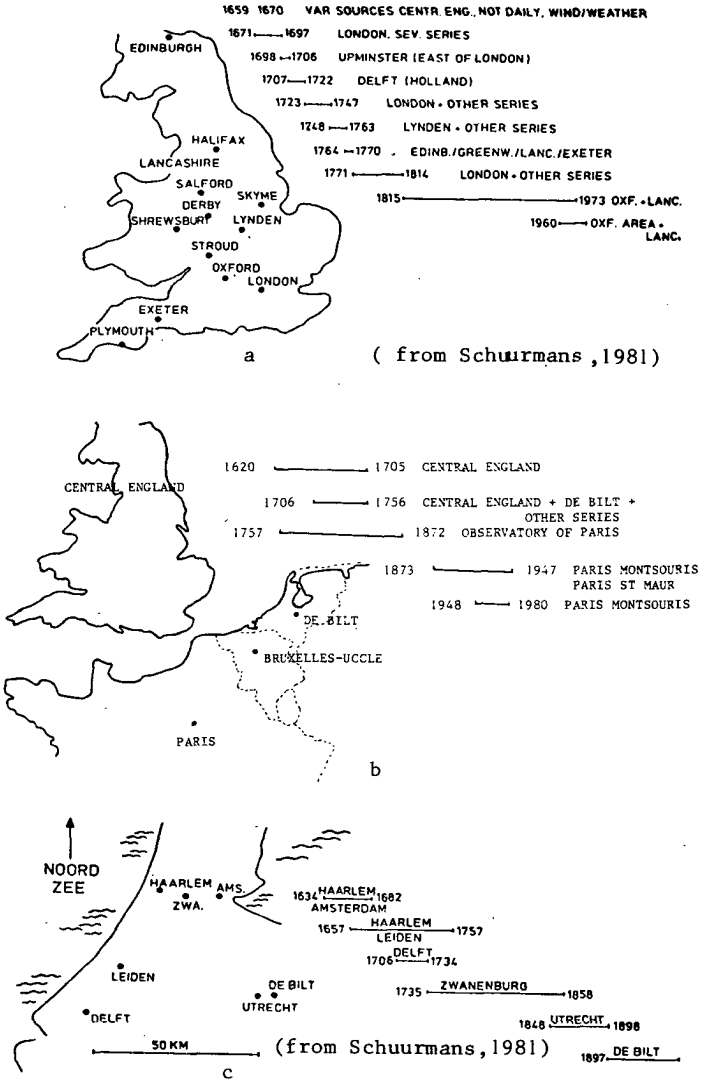


Fig. 3. a to c Location of stations, from which records are used in compiling the Central England (a), the De Bilt (b) and the Paris (c) temperature series. Horizontal bars indicate the time interval for which records (or collection of records) from other places contribute to each series.

The application of the Sneyers' serial correlation test (Table 1) shows that at the level of 0.05, each series present inhomogeneities. This may probably be due to the complicated structure of these series, especially for the earlier years.

The Mann test was also applied to each series. The value of u(t) being significant and positive, both series present a general warming trend. Fig. 4 gives the graphical representation of the annual temperature series of Central England from 1659 to 1973 and Fig. 5 the one for Paris (1680 to 1981). The intersection of the two curves shows the existence of a change around 1898 for Central England and 1920 for Paris. These series present also oscillations between maxima

and minima values : a very cold period at the end of the 18th century (end of the "Little Ice Age"), followed by a rapid warming up to 1720. From 1720 to the beginning of the 20th century, the climate is stable. A warning trend is then observed around 1920 in Paris and around 1890 in Central England followed by a stable period up to 1920. Such a climatic change, around 1916, was also observed by Sneyers (1974) (Fig. 7) in the annual mean temperature series at Uccle (1833-1969) and is observed at De Bilt around 1905 (Fig. 6). This warming is always observed. It must however be underlined that the series of De Bilt presents much more variability during the 18th and 19th century that the other series.

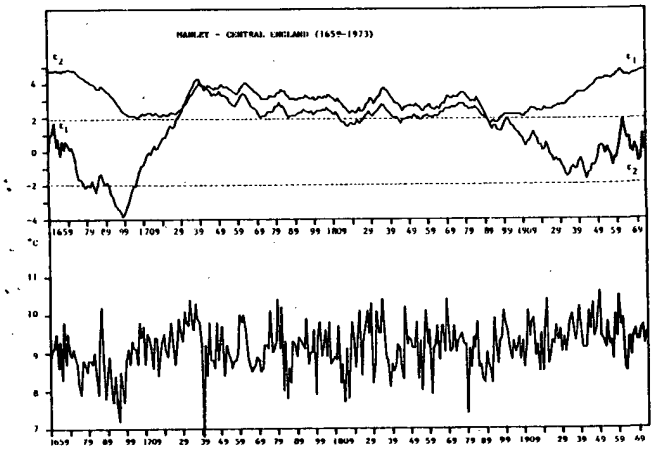


Fig. 4. Mean annual air temperature values in Central England for the period 1659-1973. Values based on the trend test for partial series.

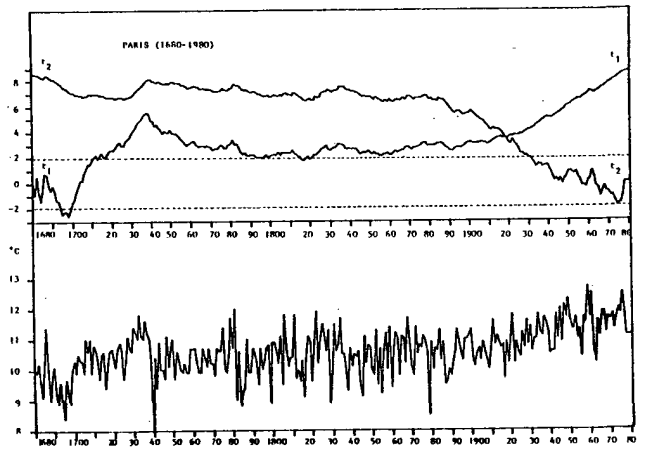


Fig. 5. Mean annual air temperature values in Paris for the period 1680-1980. Values based on the trend test for partial series.

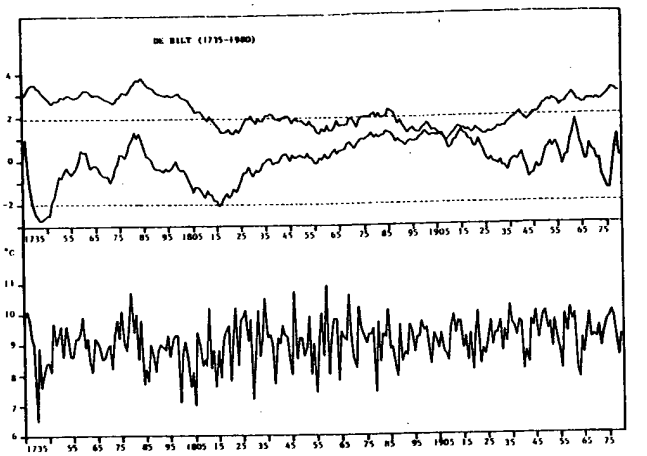


Fig. 6. Mean annual air temperature values in De Bilt for the period 1735-1980. Values based on the trend test for partial series.

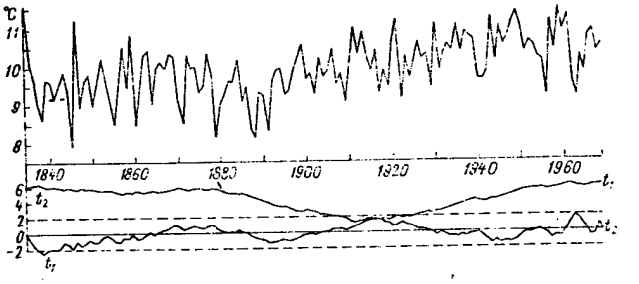


Fig. 7. Mean annual air temperature values in Brussels-Uccle for the period 1834-1969. Values based on the trend test for partial series. (Sneyers, 1974)

5. INSTRUMENTAL TEMPERATURE SERIES FOR WESTERN EUROPEAN COUNTRIES

Our basic data sets are the monthly station data originally published by the Smithsonian Institution in various volumes of the World Weather Records up to 1960, digitized by the National Center for Atmospheric Research (Jenne, 1975) and updated by NOAA in Monthly Climatic Data for the World (MCDW).

The Sneyers' serial correlation coefficient (Table 1) shows that, at the level of 0.05, many of these stations present inhomogeneities in their series. This may be easily explained by the fact that the location of the stations or the instruments have changed over the years. The effect of increasing urbanisation on climatic records should also be noted (Mitchell, 1961; Landsberg, 1981).

The test of Mann is then applied to most of the instrumental Western European series (Table 2). Most of the stations present a general warming trend except Hohenpeissenberg, Wien and Madrid which do not present any trend and Palma which has a cooling trend.

Table 2. Instrumental temperatures for Western European countries series, location of the station, period of observation, year of the beginning of the climatic change.

Country	Station	Lat.	Long.	Height	Period of observation	Beginning of the climatic change
Austria	Wien	48°15N	16°22E	203 m	1775-1981	1810 (⁙)
	Innsbrück	47°16N	11°24E	579 m	1906-1977	no trend
France	Paris	48°49N	2°20E	50 m	1874-1960	1925
	Nantes	47°10N	1°37E	27 m	1881-1981	1921
	Marseille	43°27N	5°13E	8 m	1871-1981	1950
Germany	Potsdam	52°23N	13°04E	85 m	1893-1980	no trend
	Berlin	52°28N	13°18E	58 m	1769-1970	1856
	Frankfurt	50°03N	08°35E	113 m	1835-1970	1855
	Hohenpeissenberg	47°48N	11°01E	986 m	1781-1970	1810 (⁙)
Greece	Athens	38°00N	23°42E	107 m	1858-1976	1915
Italy	Milano	45°30N	9°12O	103 m	1866-1978	1950
Norway	Bodo	67°16N	14°22E	13 m	1868-1981	1900
Portugal	Lisboa	38°42N	9°060	95 m	1864-1977	1935
Spain	Madrid	40°24N	3°420	657 m	1860-1979	1955
	Palma	39°36N	2°42E	45 m	1866-1979	1887 (⁙)
United Kingdom	Aberdeen	57°12N	2°060	59 m	1872-1981	1890
Sweden	Uppsala	59°52N	17°38E	14 m	1855-1970	1900
Switzerland	Bâle	47°36N	7°36E	318 m	1755-1970	1918
	Zurich	47°24N	8°36E	569 m	1864-1981	1910
	Santis	47°18N	9°18E	2496 m	1883-1981	1930
	Geneve	46°12N	6°12E	405 m	1768-1970	1936

⁙ means a decreasing trend

For the stations of Hohenpeissenberg and Wien (Fig. 8), intersection of curves t_1 and t_2 occurs around 1810 (Table 2). For Uppsala (Fig. 9), Bodo and Aberdeen, the warming trend begin around 1900. The behaviour of stations in Switzerland (Fig. 10), France, Portugal and Greece is similar with a change starting a little bit later, around 1925 (Table 2). Berlin (Fig. 11) and Frankfurt already present this change around 1855. Both of these stations present oscillations inside the series. Wien and Hohenpeissenberg present a cooling from 1810 up to 1900, then a warming up to now, with a stable period from 1900 to 1925 for Wien. Stations having a change around 1900 present a maximum around 1950. Most of the stations having a change around 1925 present a warming up to now.

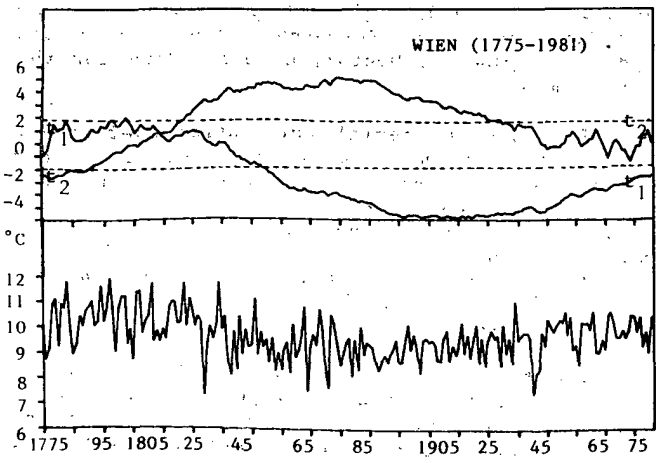


Fig. 8. Mean annual air temperature values in Wien for the period 1775-1981. Values based on the trend test for partial series.

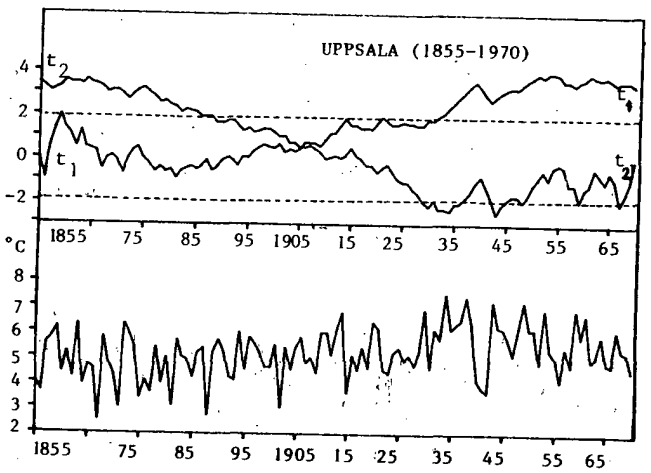


Fig. 9. Mean annual air temperature values in Uppsala for the period 1855-1970. Values based on the trend test for partial series.

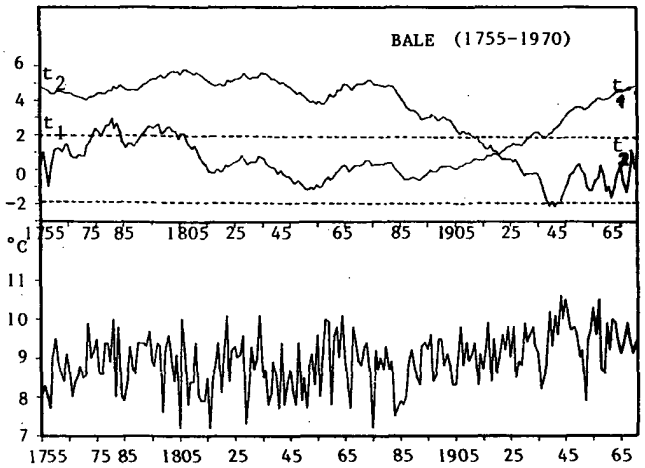


Fig. 10. Mean annual air temperature values in Bale for the period 1755-1970. Values based on the trend test for partial series.

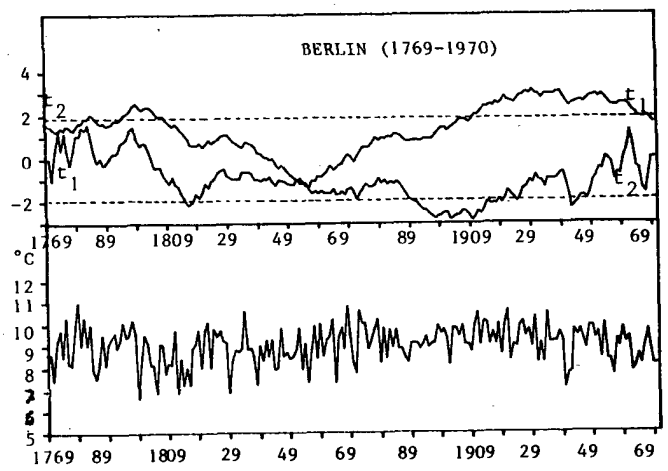


Fig. 11. Mean annual air temperature values in Berlin for the period 1769-1970. Values based on the trend test for partial series.

6. CONCLUSION

Our aim has been to applied statistical tests in order to detect significant changes in climatic time series. The European reconstructed series (England, De Bilt, Paris, Uccle), the average Northern Hemisphere temperatures curve and the instrumental records of temperature for Western European countries have been used.

The test of Mann shows that most of these series present a general warming which support the reality of the early 20th century northern hemispheric warming noted by many authors. However, it clearly appears from the detailed statistical analysis of all curves that this warming trend is continuous up to now and started generally around the turn of our century ± 15 years. It is only the climatic variations superposed on it which display minima around 1885 and 1975 and a maximum around 1945 (as best noted on the Northern Hemisphere temperature series).

For stations situated between 50 and 60°N and for the average N.H. temperature, the beginning of the long term warming trend is observed around 1900, with superposed to it a warming up to around 1945, followed by a cooling. Stations situated between 40° and 50°N present a long term change later on, around 1925 and a warming again since that time, while for higher latitudes (80°N), this change appears as soon as 1907. The low latitudes present only short fluctuations.

These parallel trends may be explained by the fact that climatic variations arise primarly from variations in the larger-scale components of the general circulation, which in turn are dynamically interrelated over great distances. Consequently, climatic variations tend to be conherent in areas many hundred (or even thousands) of kilometer in diameter. It is no exageration to say and this is statistically demonstrated here that many climatic variations, especially, the long term variation that span several decades in time, are world wide in geographical extent.

7. ACKNOWLEDGEMENTS

Thanks are due to the "Institut de Recherches Appliquées à l'Industrie et à l'Agriculture" for financial help to one of the authors (C.G.), to R. Sneyers (Institut Royal Météorologique, Bruxelles) for his helpful suggestions, to the Climatic Research Unit of East Anglia for providing the European data and to C. Schuurmans for providing the data of De Bilt.

8. REFERENCES

Borzenkova I.I., K. Ya. Vinnikov, L.P. Spirina and D.I. Stekhnovskii (1976). Change in the air temperature of the Northern Hemisphere for the period 1881-1975. Meteor. Gidrol., 7, 27-35 (in Russian).

Callendar G.S. (1961). Temperature fluctuations and trend over the earth. Quar. J. Meteor. Soc., 87, 1-12.

Dettwiller J. (1981). Les températures annuelles à Paris durant les 300 dernières années. La Météorologie, VI, 25, 103-110.

Goossens Chr. (1983). Etude de l'homogénéité et de la stabilité des séries climatiques. Scientific Report 1983/1, Institut d'Astronomie et de Géophysique, Université Catholique de Louvain-la-Neuve.

Gruza G.V. and E.Ya. Ran'kova (1979). Data on the structure and variability of climate-air temperature sea level. Northern Hemisphere. Vsesoyuze'e Naucho-Issledovatelovatyel' Skeenstitut Gidrometeorologicheskoe Informatsii-Mirovoe Tsentr Dann'ch (Central Institute for the Collection of Hydrometeorological Information - World Data Centre), Obninsk, USRR, 204pp. (in Russian).

Jones P.D., T.M.L. Wigley and P.M. Kelly (1982). Variations in surface air temperatures : Part

I. Northern Hemisphere, 1881-1980. Mon. Wea. Rev., 110, 2, 59-70.

Labrijn A. (1945). The climate of Netherlands during the past 2 1/2 centuries. Med. Verh. Kon. Ned. Met. Inst., 49, 102.

Landsberg H.E. (1981). The urban climate. International Geophysics Series, N.Y. Acad. Press, 28, 275pp.

Jenne R. (1975). Data sets for meteorological research.. NCAR-TN/IA-III, National Centre for Atmospheric Research, Boulder, 194pp.

Manley G. (1953). The mean annual temperature of Central England. Quar. J. Meteor. Soc. 79, 24, 242-261.

Manley G. (1974). Central England temperature : monthly means 1659-1973. Quar. J. Meteor. Soc. 100, 387-405.

Mann H.B. (1945). Non-parametric test against trend. Econometrika, 13, 245-259.

Mitchell J.M. (1961). Recent secular changes of global temperature. Ann. N.Y. Acad. Sci., 95, 235-250.

Mitchell J.M. (1963). On the world-wide pattern of secular temperature change. Proc. UNESCO/-WMO Symposium of Changes of Climate, UNESCO Arid Zone Research Series, 20, UNESCO, Paris 161-181.

Schuurmans C. (1981). Central England and De Bilt series. In : Climatic Variations and Variability : Facts and Theories, A. Berger (Ed.), Reidel.

Sharon V. Ya. (Ed.). 1960-67 : Charts of Air Temperature Deviations from the Multiyear Northern Hemisphere Averages. N°s 1-4. Main Geophysical Observatory (GGO), Leningrad.

Sneyers R. (1957). Sur la détermination de l'homogénéité des séries climatologiques. J. Scientifique de la Météorologie, 7, 28, 359-372.

Sneyers R. (1963). Sur la détermination de la stabilité des séries climatologiques. Proc. UNESCO/WMO Symposium of Changes of Climate, UNESCO Arid Zone Research Series, 20, UNESCO, Paris, 37-44.

Sneyers R. (1974). Statistical analysis of temperature and rainfall data at Brussels-Uccle from 1833 to 1969. Proceedings of the symposium of physical and dynamic climatology Leningrad, WMO n° 347.

Sneyers R. (1975). Sur l'analyse statistique des séries d'observations. WMO, Genève, n° 415, TN143.

Sokhrina R.F., O.M. Chelpanova and V. Ya. Sharova (1959). Air temperature and atmospheric precipitation in the Northern Hemisphere. Atlas of Charts. Gidrometeoizdat, Leningrad.

On the verification of January GCM simulations

Hans von Storch and Erich Roeckner
Universität Hamburg, Bundesstr. 55, 2000 Hamburg 13, FRG

1. INTRODUCTION

The objective of this paper is to present a technique how to compare a model generated (small) sample of certain climate variables with the ensemble of respective observed values and to apply it exemplarily to January simulations performed with the Hamburg University GCM. Of course, the suggested procedure may also be used to compare different model generated climates.

For this purpose, a multivariate, nonparametric test to compare e.g. 1st and 2nd moments of two multivariate random variables is proposed. It is based on a drastical reduction of degrees of freedom by means of an EOF-expansion and on the comparison of a statistic of the original data with those obtained under all permutations of the same data record. If the number of permutations is too large, a randomized design is possible. (Cf. Boyett and Shuster, 1977). A similar approach was proposed by Preisendorfer and Barnett in a paper recently submitted to J.Atm. Sc.. The statistic is chosen according to what is to be tested: for 1. moments the standard t-or the rank statistic proposed by Mann and Whitney may be appropriate, for 2nd moments the standard F- or the rank-statistic of Siegel and Tukey (see e.g. Conover, 1971).

Since the test procedure can yield only the results "the nullhypothesis is to be rejected, i.e. the model differs from nature" and "the validity of the alternative is not sufficiently sure", we can not give positive statements on the model but only negative ones. Therefore, we will emphasize the deficiencies of the model under concern. This does not mean that it is worse than other models, but only that most modellers like to show the bright sides of their GCM instead of their unsatisfying aspects. We believe that after 30 years of development GCMs have become so good that we can start to investigate their drawbacks.

The procedure is applied to assess the quality of January simulations performed with the Hamburg University GCM (Roeckner, 1979). Thus, one of the random variables is the observed climate and the other the model generated circulation. The parameters taken into account are the zonal monthly means of the geopotential height at 300, 500 and 850 mb, the zonally averaged variance of the transient and of the stationary geopotential-eddies at the same levels and frequency-wavenumber spectra (MEM-estimates; Storch and Fischer, 1983) of geopotential at 300 mb along 50N.

This work is a direct continuation of Storch and Roeckner (1983) (in the following abbreviated as SR). In that earlier article the task was treated how to decide whether an individual model generated January may originate from observed climate (with the aid of the statistic (2.7) mentioned in Section 2.7 of this paper).

After having diagnosed certain model-deficiencies, the question arises, how to eliminate them. However, this question generally cannot be answered without further studies. By means of a simplified model, we try to relate partly the insufficient modelled zonal average of the variance of stationary eddies to the simulated zonal current. Surely, such an approach cannot prove any hypothesis. But, at least, it yields an earnest hint.

In Section 2, the nonparametric technique and the competitive parametric procedure to test the equality of vectors of e.g. expectations or variances is described und discussed. The modell and the data to which the test is applied is given in Section 3. The results concerning the stationary features and the discussion by means of the simplified modell is contained in Section 4. The quality of the simulation of the transient features is treated in Section 5.

2. A CLASS OF NONPARAMETRIC TWO-SAMPLE TESTS

The following procedure is the generalization of a permutation technique originally proposed by Fisher (1966) and further developed by Boyett and Shuster (1977).

2.1. Univariate approach

Let x and y be two random variables (r.v.) and $x_1, \ldots x_n$ and $y_1, \ldots y_m$ samples of x and y, respectively, $n \geq m$ and z the combined vector $(x_1, \ldots x_n, y_1, \ldots y_m)$. We want to test whether some parameter P (e.g. expectation or variance) of x and y is different. For this purpose we need a function t(z) with large values if $P(x) \neq P(y)$ and small values otherwise.

Furtheron, t shall fulfill:
(*) If Υ is the set of all permutations σ of $\{1, \ldots n+m\}$ and if the nullhypothesis $P(x)=P(y)$ is true, for all $\sigma \in \Upsilon$ the random variable $t(z_\sigma) := t(z_{\sigma(1)}, \ldots z_{\sigma(n+m)})$ is distributed identical to t(z) itself.

Then, the decision "$P(x) \neq P(y)$" is right at the 5% level of significance if

$$(2.1) \quad N := |\{\sigma \in \Upsilon, t(z_\sigma) > t(z)\}| / |\Upsilon| \leq 5\%$$

Since no assumptions are made on the type of the distribution of the random variables, the procedure may be called "nonparametric". This aspect is the main advantage compared to standard normal techniques, because these tests may be applied to all kind of distributions, e.g. to nonsymmetric ones.

If the numbers n and m are large, it may be impossible to compute $t(z_{\sigma})$ for all permutations. To overcome this difficulty, one can restrict oneself to a randomly formed subset of Υ ("randomized design"). We apply a subset of 2000 permutations.

2.2. Multivariate generalization

If instead of a univariate observation a p-dimensional vector is to be treated, one can compute the statistic t^j for each component, j=1,...p, and form some appropriate vector norm, e.g. the least square norm:

$$(2.2)\quad T := \sum_{j=1}^{p} |t^j|^2$$

Then, the same procedure as outlined above is applied to T instead of t. (If (*) is true for t, then it holds for T also.)

2.3. Expectation: choice of t

If the coincidence of the expectations of the (multivariate) random variables x and y is to be tested (i.e. P(x) := E(x)), at least two functions t are suitable:

(T1) the standard Student statistic.

(T2) the rank-statistic proposed by Mann and Whitney:

$$(2.3)\quad t := \left| \sum_{i=1}^{n} R_i - \frac{n}{2}(n+m+1) \right|$$

(The sum is formed just for the first n components of z.) The rank R_i is assigned to the i.th observation if R_i-1 observations are greater than z_i and n+m-1-R_i are smaller. Thus, rank 1 is given to the largest z_i, rank 2 to the second largest and n+m to the smallest value.

T2 results in a fully nonparametric procedure: no estimation of the parameters which are to be tested is necessary. This aspect is favorable in case of small samples.

The condition (*) is strictly spoken not fulfilled if E(x) = E(y) but only if x and y are identically distributed, e.g. in case of normality if additionally Var(x) = Var(y). Thus, a rejection of the nullhypothesis may be caused not only by $E(x) \neq E(y)$ but also, in extreme cases, by $Var(x) \neq Var(y)$. In order to overcome this problem practically the data are transformed according to

$$(2.4)\quad (x-\hat{x})/s(x) + \hat{x} \quad \text{and} \quad (y-\hat{y})/s(y) + \hat{y}$$

where s(x) is the estimated standard deviation of x and $\hat{x}$ the sample mean. By this transformation, the test procedure may have become radical or conservative. To exclude this possibility, a series of 2000 Monte Carlo trials with p=1, n=4, m=6 and normal r.v. x and y with E(x) = E(y) = 0 and different, random variances was performed. Since E(x)=E(y), each rejection of the nullhypothesis is false: statistic (T1) gave 4.6% of such erroneous rejections and (T2) only 3.4%. These rates are acceptable.

2.4. Variance: choice of t

If the coincidence of the trace of the covariance matrix of x and y is under consideration, the following expressions may be used as t:

(T3) the standard Fisher-statistic F

(T4) the nonparametric rank statistic proposed by Siegel and Tukey:

$$(2.5)\quad t := \left| \sum_{i=1} \left[R_i - \frac{n+m+1}{2}\right]^2 - \frac{n(n+m+1)(n+m-1)}{12} \right|$$

The ranks are assigned to the components of z such that rank 1 is given to the largest component, rank to to the smallest, rank 3 to the second largest, rank 4 to the second smallest and so on.

T4 yields as T2 a fully nonparametric procedure.

The problem concerning the validity of (*) occurs again: it is treated by the simple transformation

$$x-\hat{x} \quad \text{and} \quad y-\hat{y}$$

2000 Monte Carlo trials with x and y identical besides of random mean values resulted in 4.4% (T3) and 5.3% (T4) false rejections of the nullhypothesis Var(x)=Var(y).

2.5. EOF expansion

Generally, vectors with many components are to be handled, which lessens the chances to detect dissimilarities at the given level of significance. Therefore, a drastical reduction of the number of components of the vectors under consideration is often desirable. Mathematically, this reduction is performed by the projection of the originally given vector into a lowdimensional subspace spanned by some appropriate orthogonal vectors. These vectors can be the first few empirical orthogonal functions.

In this case, the test procedure is not applied to the original data set but to the first few coefficients of an EOF-expansion. The EOFs are obtained from a subset of the larger sample $\{x_1, \ldots x_n\}$. Since we don't subtract the sample mean, the EOF belonging to the largest eigenvalue shows

the structure of the long year mean. Thus it is labeled as "interannual mean EOF". To renounce the subtraction of the interannual mean is crucial because it often grants a better chance to detect an overall change, which is hardly presentable by a series of centered EOFs which have generally a mean close to zero (try to expand 1 into a trigonometric series without cos(0)).

Problems around the EOF-expansion are discussed by SR, Sec. 2.4 and by Storch (1983).

2.6. Univariate analysis

If the multivariate tests yields a significant dissimilarity, the application of the 1-dimensional version of the test either to the EOF-coefficients or to the grid point values sometimes shows which components differ clear.

2.7. A comparable parametric test

In order to assess the power of the procedure outlined above, a standard normal technique is available (cf. Jackson, 1981): Provided x is a normal random variable with covariance-matrix S and expectation E(x), the random variable

$$(2.6)\quad G_0(x_1,\ldots,x_m) := m * g\left(\frac{1}{m}\sum_{i=1}^{m} x_i\right)$$

with m independent samples x of x and

$$(2.7)\quad g(x) := (x - E(x))' S^{-1}(x - E(x))$$

is χ^2 distributed with pm degrees of freedom. Thus, if G_o is greater than the 5%-quantile of the respective χ^2 distribution, the null-hypothesis that x is identical to y may be rejected at the 5% level of significance. Generally this rejection is caused by distinct mean values. Therefore, this statistic is applied to check (2.1) with the T's according to (T1) and (T2).

The statistic g itself may be used to investigate whether an individual state y_i originates from the ensemble x (see SR).

A statistic sensitive to differences of the variances is

$$(2.8)\quad G_D(x_1,\ldots x_m) := \sum_{i=1}^{m} g(x_i) - G_0(x_1,\ldots x_m)$$

which has p(m-1) d.f.. G_D is employed to value (T3) and (T4).

The main drawbacks of these parametric tests are:
(1) G_o and G_D are quite sensitive to stray shots - just one extraordinary observation x_i may cause a rejection even if all others are absolutely unconspicious. (An example occures in Sec. 5.2.)
(2) Strictly spoken, only normal distributed r.v. may be treated. (See Sec. 4.2)
(3) Since the true expectation and covariance matrix S are unknown, these quantities have to be estimated. Therefore, the quantiles are valid asymptotically only.

3. APPLICATION

The procedure described in the preceding Section is applied to compare some circulation statistics obtained by observation (= random variable x) and by simulation (= r.v. y): The simulation was done in the perpetual January mode with the Hamburg University model (Roeckner, 1979).

3.1. The GCM

The Hamburg University GCM was run up to 300 d. The first 30 d are needed to reach stationarity and are thus not considered. The remaining 270 d are subdivided into 30d-intervalls. Each of these 9 intervals are regarded as one simulated January. Thus, m = 9.
Figure 1 shows the subdivision of the total range of 300d into 30d-intervalls and displays the stationarity beyond day 30 in terms of the hemispherically averaged zonal and eddy kinetic energy.

The model is identical to that described in SR except for modifications in the planetary boundary layer (PBL)-treatment and for the use of more reliable surface boundary data.

One of the major shortcomings of the simulations used in SR was a gradual cooling up to a mean value of ~3K below the January climatological mean, caused mainly by downward directed surface heat fluxes over the low-latitude oceans. This problem is associated with the coarse vertical resolution of the model (3 layers) with the lowest level at 850 mbar - quite often above the top of the PBL and not below as necessary for the application of Ekman-layer theory.

In the present model the cooling trend has been avoided by introducing a vertical "countergradient" in the virtual potential temperature, depending on the estimated height of the PBL. The procedure effectively reduces the tendency of downward directed surface heat fluxes over low-latitude oceans especially.

The surface properties as compared to those used in SR are listed in Table 1. One important difference between both data sets is that the new data used for the present study contain much more spatial variance which is reflected also in an increased stationary eddy variance of e.g. geopotential height (not shown).

The effective roughness length z_o was derived according to a procedure of Fiedler and Panofsky (1972) who obtained a relationship between z_o and mean terrain height $\bar{h}_s$ from aircraft measurements over different terrain. In our method these data were reinterpreted in a way that we obtain a z_o

-dependence on the variance $\overline{h_s'^2}$ which is calculated from the US-Navy fine-mesh terrain heights in a 1/6°-grid (see Table 1).

Table 1

Surface data
z_o : surface roughness length
h_s : terrain height in the fine mesh (1/6°)
r_s : surface relative humidity
$\bar{(\ )}$: average over the model's grid square (2.8125°)
ECMWF: European Centre for Medium Range Weather Forecasts

Surface property	old model (SR)	present model
orography	smoothed atlas data	unsmoothed grid-square averages of fine-mesh (1/6°) US-Navy terrain data created at ECMWF.
surface roughness over land and ice	z_o = 0.2 m everywhere	$z_o \sim \overline{h_s'^2}$ with $(h_s' = h_s - \bar{h}_s)$
sea surface temperature and sea-ice extent for January	interpolated from 5° longitude 4° latiude data of Schutz and Gates (1971)	interpolated from 1° data of Alexander and Mobley (1976)
surface humidity over land	from assumption of r_s = 70% everywhere	from climatological soil moisture data for January (created by Dittmann at ECMWF)

3.2. Interpreted parameters

As in SR the geopotential height H is interpreted: this variable is besides of the temperature the only dynamic 3-dimensional one measured and analyzed carefully for a number of years on the Northern Hemisphere. Our data record consists of the daily DWD analyses of the 16 Januaries 1967 - 1982. Thus, n = 16. The data observed in the Januaries 1967-73 are used to compute the EOFs (cf. SR, Section 2.4).

The daily series of observed and simulated geopotential height fields is decomposed into stationary and transient parts:

$$H(\lambda,\varphi,t) = H'(\lambda,\varphi,t) + \bar{H}(\lambda,\varphi)$$

$$H'(\lambda,\varphi,t) = H'^{*}(\lambda,\varphi,t) + [H'](\varphi,t) \qquad (3.1)$$

$$\bar{H}(\lambda,\varphi) = \bar{H}^{*}(\lambda,\varphi) + [\bar{H}](\varphi)$$

The overbar (bracket) denotes a monthly (zonal) mean, a prime (star) the deviation from a monthly (zonal) mean. For technical reasons, the hemispheric monthly means are substracted from the overbared values, additionally. The parameters studied are the zonally averaged $[\bar{H}]$, $[H'^{*2}]\cos\varphi$ and $[\bar{H}^{*2}]\cos\varphi$ at all model levels, i.e. 300, 500 and 850 mbar.
Furtheron, the monthly one-sided frequency wavenumber spectra along 50N at 300 mbar are handled to value the simulated time scales. The spectra are estimated by the maximum entropy method (Storch and Fischer, 1983).

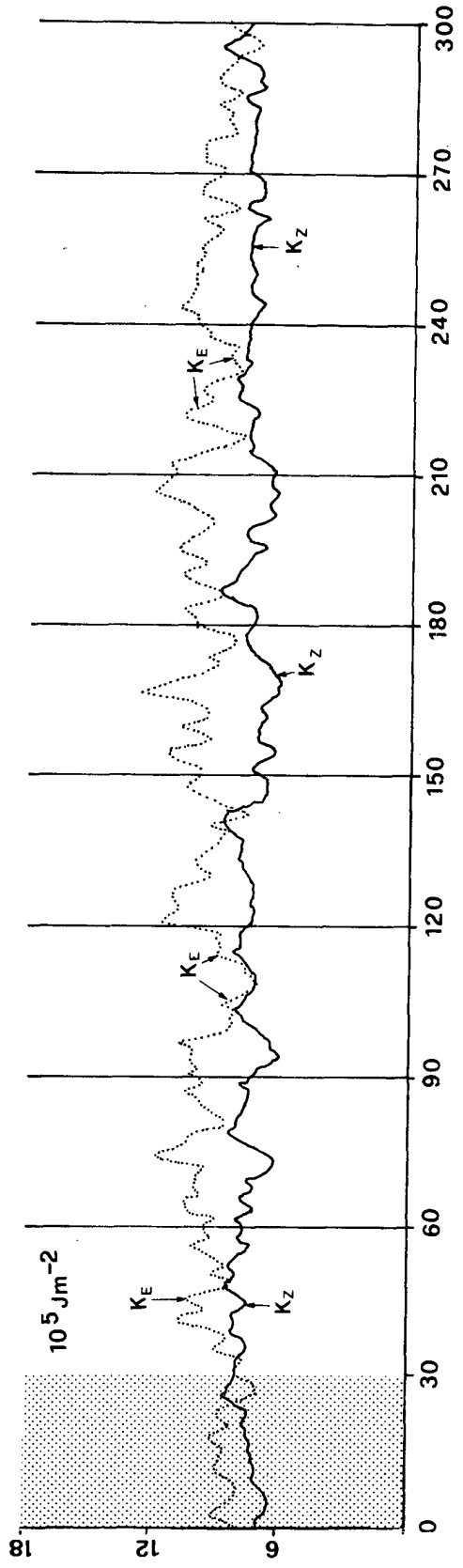


Figure 1
Temporal developement of zonal and eddy kinetic energy, K_Z and K_E, and subdivision of the total 300d-range into 30d-intervalls.

4. STATIONARY FEATURES

In Table 2, the results of the nonparametric tests T1-4 applied to the stationary $[\bar{H}]$ and $[\bar{H}^{*2}]\cos\varphi$ are given.

Table 2

Results of the nonparametric test statistic N (2.1) with T1 - T4 applied to the stationary features (Section 4). If N $\leq$5%, the model generated state is significantly different from the respective observed one. A star is placed in the $G_{\bar{o}}$ or $G_{\bar{D}}$ -column, if the respective tests classifies the model generated state as significant.

Circulation statistic	Expectations T1	T2	$G_{\bar{o}}$	Variances T3	T4	$G_{\bar{D}}$
850mb						
$[\bar{H}]$	0.1%	0.1%	*	49%	57%	*
$[\bar{H}^{*2}]\cos\varphi$	0.1%	0.1%		57%	47%	
500mb						
$[\bar{H}]$	0.1%	0.1%	*	6.9%	8.1%	
$[\bar{H}^{*2}]\cos\varphi$	3.0%	4.2%	*	0.2%	1.0%	*
300mb						
$[\bar{H}]$	0.1%	0.1%	*	2.5%	18%	
$[\bar{H}^{*2}]\cos\varphi$	1.3%	0.2%	*	5.3%	4.5%	*

4.1 Ensemble means

The ensemble averages of the zonal monthly mean is discernibly insufficient simulated at all levels, which is mainly described by a too small coefficient of the interannual mean EOF: the north-south-contrast is weakened and the maximum of the zonal current is shifted northward for about 10 degrees. (1) In fact, this deficiency is not only found in the ensemble average but was diagnosed in each individual model generated January by the g-statistic (2.7). At 850 mb, even more EOF-coefficients are remarkable pointing at additional unrealistic patterns: $[\bar{H}]$ is lifted in the polar region for about 100 m. In Figure 2, the situation is shown for 300 and 850 mb.

The ensemble averages of the model generated zonally averaged variance of stationary eddies, $[\bar{H}^{*2}]\cos\varphi$ differ significantly from the respective observed ones at all levels, also. The univariate analysis yields an underestimated amplitude of the respective interannual mean EOF in the "free atmosphere", i.e. at 300 and 500 mb. In opposition, the interannual mean EOF is too strong at 850 mbar. Different from $[\bar{H}]$, the individual $[\bar{H}^{*2}]\cos\varphi$ are not rejected by the g-statistic (2.7). In Figure 3, the situation is shown for 300 and 850 mb.

---

1) This problem is also observed in more advanced and global GCMs, e.g. the ECMWF model. (Savijärvi, 1978)

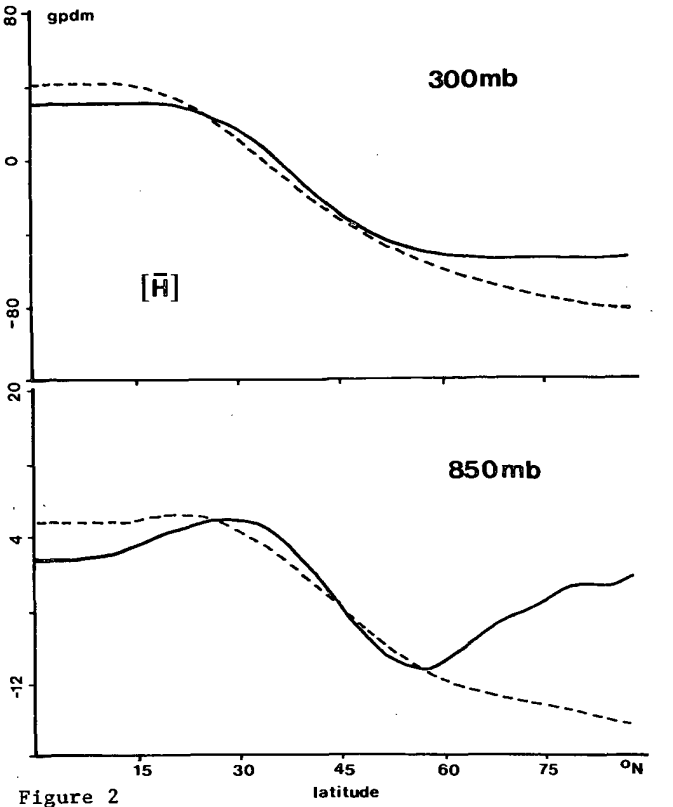


Figure 2
Estimated ensemble means of $[\bar{H}]$ at 300 mbar (above) and 850 mbar (below) for the observed climate (dashed line) and the model generated one (solid line).

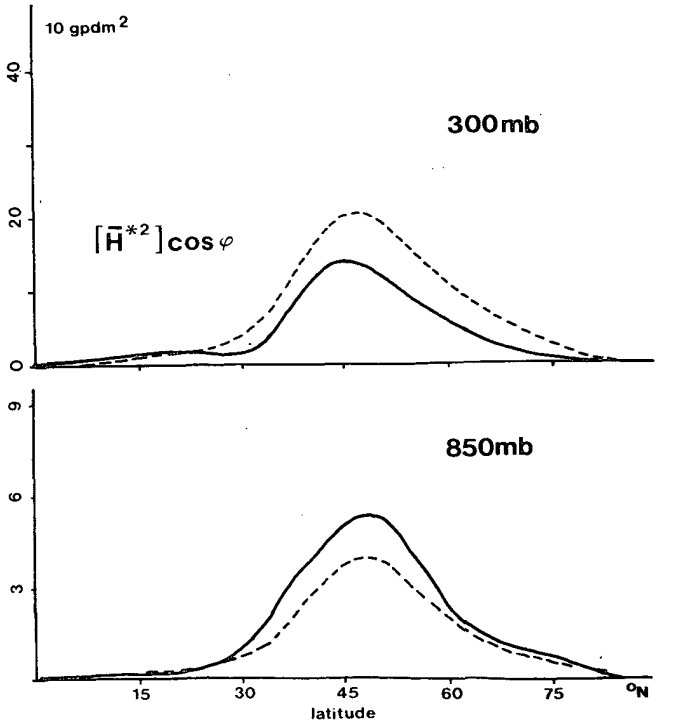


Figure 3
Same as Figure 1, but showing $[\bar{H}^{*2}]\cdot\cos\varphi$, the zonal average of the variance of stationary eddies.

A comparison of the results of T1-T2 with the respective ones obtained by G_o gives a fair agreement (Table 2).

4.2 Interannual variability

The interannual variability is unconspicious at 850 mbar. Var($[\bar{H}]$) is reduced insignificantly at 500 mbar and significantly (according to T3) at 300 mbar. Furtheron, the interannual variability of the variance of the stationary eddies, Var($[\bar{H}^{*2}]$), is distinctly decreased in the free atmosphere. In Figure 4, the situation is displayed at 300 mb showing the 95% tubes of both climates. (The 95%-tube is defined pointwise in the Euclidean space: it is the symmetric interval around the mean containing on an average 95% of all realizations of the respective climate variable in case of normality; cf. SR.)

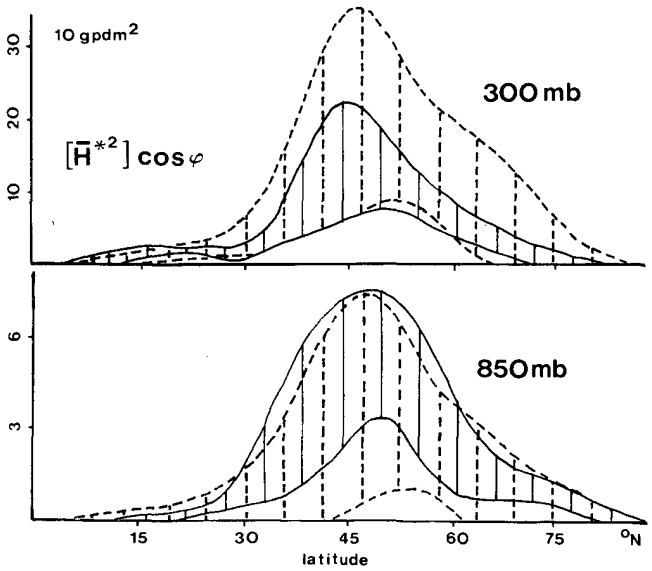


Figure 4
95%-tubes for the observed (dashed) and simulated (solid) zonally averaged variance of stationary eddies at 300 mbar (above) and 850 mbar (below).
A 95%-tube is defined for each latitude individually as that symmetric intervall around the mean containing on an average 95% of all states of its ensemble.

The results of the nonparametric tests T3-4 coincide quite well with the G_D one (see Table 2), except for Var$[\bar{H}]$ at 850 mb. This may be due to the assumption of normality for the application of the parametric test: at least one of the EOFs used to expand $[\bar{H}]$ at that level has significantly nonnormal distributed coefficients (Lillifors, 1967). The statistic (2.6) characterized just that EOF-coefficient as extremely.

4.3 Discussion

With the aid of a linear, stationary and barotropic model based on the potential vorticity equation, the dissimilarities of $[\bar{H}^{*2}]\cos\varphi$ may in a crude way be related to the insufficiently modelled zonal current, i.e. to $[\bar{H}]$.

The model is forced by a realistic topography (same as in the GCM) and a prescribed zonal current U. It is run with the 16 at 500 mbar observed U's and the 9 simulated U's. The resulting stationary eddies are labeled with a ℓ: H_ℓ .

We restrict ourselves to the zonally averaged variance of the (stationary) eddies, since the phases of the linear response are not so good.

To start with the interannual mean: The ensemble average of $[\bar{H}^{*2}]$ along 50N at 500 mbar computed with the GCM is 81 gpdm² compared to 110 gpdm² observed. Thus, the variance of GCM generated stationary eddies equals 74% of observation. On the other hand, the linear model predicts a response of 67 gpdm² to the simulated 500 mbar-zonal current and 97 gpdm² to the observed U's: The ratio is 69%.

To have additionally a hemispheric measure, the amplitude of the interannual mean EOF, c_M, is computed for the GCM generated and the observed $[\bar{H}^{*2}]$. The ratio is 59%. A comparable number is the meridional integral of $[\bar{H}^{*}_\ell]$ resulting in a ratio of 49% for the simulated versus observed U.

Thus, the deficiency of the simulated mean intensity of stationary eddies is expected to become eased if a more realistic zonal current could be simulated.

The same holds for the interannual variability: the ratio of simulated to observed Var(c_M) is 18%, whereas the linear approach predicts a relation of 26%.

5. TRANSIENT FEATURES

In Table 3 the results of the nonparametric tests T1-4 applied to $[\bar{H}^{*2}]\cos\varphi$ at all levels, and to the one-sided frequencey wavenumber spectra at 300 mb along 50N are given.

5.1 Ensemble means

All ensemble means of the zonally averaged variance of transient eddies are insufficiently modelled. According to the univariate EOF-analysis this remarkability is caused by additional unrealistic patterns (a maximum at about 42°N at 300 and 500 mbar, and at about 48°N at 850 mb) as well as an extraordinary amplitude of the interannual mean EOF: in the free atmosphere the interannual mean is too weak whereas at 850 mb it is overestimated. The situation is sketched in Figure 5 for 300 and 850 mbar. These results

hold for about half of all individual model generated Januaries according to (2.7), also.

The frequency-wavenumber spectra at 300 mbar along 50°N are unconspicious which may be partly due to the adequately modelled zonal current at that specific latitude (cf. Storch and Fischer, 1983).

Table 3

Results of the nonparametric test statistic N (2.1) with T1 - T4 applied to the transient features (Section 5). If N $\leq$5%, the model generated state is significantly different from the respective observed one. A star is placed in the G_σ or G_D -column, if the respective tests classifies the model generated state as significant.

Circulation statistics	Expectation T1	T2	G_σ	Variances T3	T4	G_D
850mb						
$\overline{[H^{*2}]}\cos\varphi$ 500mb	0.2%	0.5%	*	72%	48%	*
$\overline{[H'^{*2}]}\cos\varphi$	0.5%	0.1%	*	26%	24%	
300mb						
$\overline{[H'^{*2}]}\cos\varphi$	0.1%	0.1%	*	58%	30%	
Frequency-wavenumber spectra	82%	57%		1.3%	4.4%	*

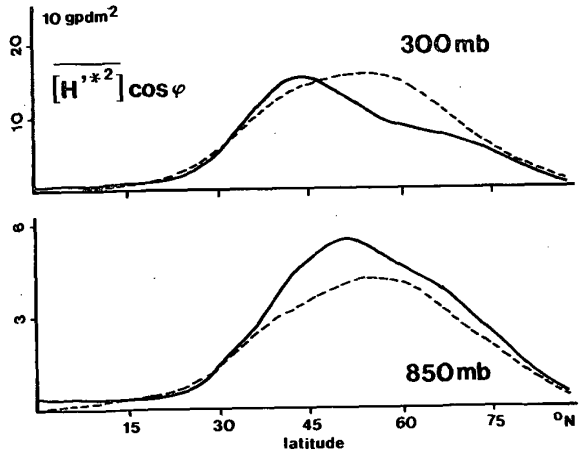


Figure 5
Estimated ensemble means of the zonally averaged variance of transient eddies at 300 mbar (above) and 850 mbar (below). The observed state is dashed, the simulated solid.

5.2. Interannual variability

The interannual variability of $\overline{[H'^{*2}]}\cdot\cos\varphi$ is not remarkable at any level. To display this fact, the respective 95%-tubes of the observed and modelled climates are shown in Figure 6 at 300 and 850 mbar.

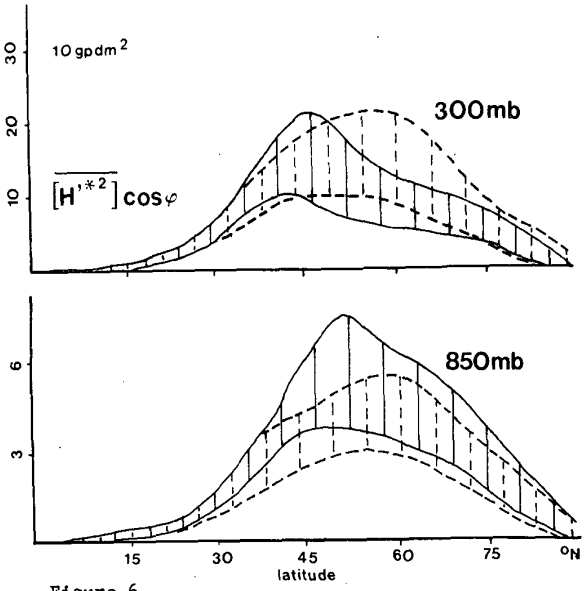


Figure 6
95%-tubes for the zonally averaged variance of transient eddies. Same as Figure 4.

Only the interannual variability of the frequency wavenumber spectra are significant at the 95% level. The univariate analysis yields an underestimation of the variance of the coefficients of all considered EOFs. This fact may partly be related by the aid of the simplest Rossby phase relation to the underestimated Var[H] at that level (cf. Storch and Fischer, 1983).

Again, the results of the T- and G-tests conincide fair expect for the scattering of the variance of the transient eddies at 850 mbar (see Table 3) . The rejection by G_D is caused by the transient activity between the days 150 and 180 of the model run. This single interval yields an extraordinary $g([\overline{H'^{*2}}])$ resulting in a significant value of (2.8). Thus, here we have an example of the drawback (1) of the G-tests mentioned in Section 2.7

6. CONCLUSIONS

A multivariate nonparametric two-sample test was presented and discussed. To examine whether the proposed procedure possesses some power, it is applied to long term integration performed with the Hamburg University GCM, exemplarily.

The geopotential height at 300, 500 and 850 mbar was examined. At the 5% level of significance, the suggested test diagnosed the following

deficiencies:

(1) 1. moments, i.e. interannual means
All zonally averaged model generated quantities are significantly different from observations. Especially the zonal monthly means are remarkable, since they exhibit a reduced difference between the northpole and the equator and a jet stream with appropriate magnitude but a location shifted nearly 10 degrees northward. The stationary eddies are overestimated at the model's lowest level, but underestimated at 300 and 500 mb. The latter can partially be connected to the inappropriate zonal flow by a simple linear, barotropic, stationary, orographically forced model. The transient eddies show an erroneous maximum at about 40N at all levels and a general overestimation at 850 mbar. The frequency wavenumber spectra at 50N are unconspicious, which may be understood from the fact that the zonal mean flow is simulated realistically at that specific latitude (cf. Storch and Fischer, 1983).

(2) 2. moments, i.e. interannual variability
The interannual scattering at 850 mb and of the transient eddies are unconspicious. That of the zonal monthly mean is insignificantly weak at 500 mb and significantly underestimated at 300 mb. The variability of the stationary eddies is also significantly too small at 300 and 500 mb, which can partly be related to the variability of the zonal current by the above mentioned simple linear model. Furtheron, the spectra vary less than the observed ones, which reflects the corresponding behavior of the mean zonal flow at that latitude.

A comparison of the presented scheme with a standard normal technique showed in most cases an equivalent power but also some drawbacks of the parametric method.

References

Boyett,J.M.; J.J.Shuster,1977, Nonparametric one-sided tests in multivariate analysis with medical applications, J.Amer.Stat. Ass. 72, 665-668
Conover,W.J., 1971, Practical nonparametric statistics, Wiley, New York - London - Sydney - Toronto
Fiedler, F. and H.A. Panofsky, 1972, The geostrophic drag coefficient and the effective roughness length, Quart.J.Met.Soc. 98, 213-220
Fisher,R.A., 1966, Design of experiments, 20th edition, Edinburgh: Oliver and Boyd
Jackson,Edward, 1981, Principal components and factor analysis: part III - additional topics relted to principal components, J.Quality Techn. 13, 46-58
Lillifors,H.W., 1967, On the Kolmogorov-Smirnov test for normality with mean and variance unknown, J.Am.Stat. Ass. 62, 399-402
Roeckner,E., 1979, A hemispheric model for short range numerical weather prediction and general circulation studies, Beitr.Phys.Atm. 52, 262-286
Storch,H.v., I.Fischer, 1983, An analysis of geopotential at 300 mbar in the frequency wavenumber domain along 50n in observed and modelled January climate, Beitr.Phys.Atm. 56, (in press)
Storch,H.v.; Roeckner,E., 1983, A strategy to assess the quality of GCM circulations with applications to a January simulation with the Hamburg University model, submitted to Mon. Wea.Rev.
Storch,H.v.; 1983, Statistical aspects of empirical orthogonal functions based on small sample sizes; this issue

BAYESIAN MODEL SELECTION

A. Ramachandra Rao and Rosalinda A. Cantiller

School of Civil Engineering
Purdue University
West Lafayette, IN 47907, U.S.A.

1. INTRODUCTION

Markovian models have been frequently fitted to climatological data. A class of models called the fractional Gaussian noise (FN) models have been recently proposed for modeling hydrologic and climatological data (Mandelbrot and Wallis (1968)). The main advantage of FN models is that they can preserve the statistic "Hurst's H" which is of considerable importance in reservoir storage design. Recently it has been shown that the statistic H of time series may also result from non-stationarity of time series (Klemeš (1974), Potter (1976)). Consequently, H would be an indicator of climatic change with climatological time series.

In view of these considerations, the question whether Markovian models to which the well known autoregressive (AR) and autoregressive moving average (ARMA) models belong or the FN models are appropriate for modeling climatological data is an important question. The basic objective of research reported herein is to compare FN and ARMA models by using Bayesian decision rules so that the appropriate model may be selected.

The Bayesian decision rule is formulated by first computing the posterior probabilities by using Bayes theorem and prior probabilities, under the assumption that the data are Gaussian distributed. These posterior probabilities are combined with loss functions to estimate average losses. The prior probabilities and loss functions may be varied to study the effect of changes in these decision variables on final decisions. The model which gives the least average loss is selected as the better model. Further details about formulation of Bayes decision rules are found in well known books such as that by Pratt et al. (1976).

The above procedure is used for comparing ARMA and FN models fitted to several observed or proxy climatological time series. The data analyzed include the annual central England temperature series, a varve thickness series and several tree ring series. The data are transformed whenever it is needed so that they are Gaussian and then used in the analysis. The data are also divided into segments and the segments are separately analyzed to study the variation in results which arise from different segments.

The rest of the paper is organized as follows. Data used in the study are discussed in Section 2. Some of the important aspects of the FN and ARMA models are discussed in Section 3. Theoretical aspects of the Bayesian decision rules used in the study are discussed in Section 4. The results from the Bayesian decision theory are discussed and a set of conclusions are presented in Section 5.

2. DATA USED IN THE STUDY

Six annual time series used in the present study are listed in Table 1. Four of these (Bristlecone, Bigcone, Navajo and Eaglecol) are tree ring series, and one (Espanola) is a varve thickness series. The tree ring and varve series are proxy climatologic data and were obtained from Dr. A.I. McLeod. The annual central England temperature data (England) are taken from Manley (1953). All of these series are modeled by Hipel and McLeod (1978) also. These data are transformed by using the Box and Cox (1964) transformations given in (1) so that the transformed data are Gaussian and homoscedastic. In eq. 1 C is a constant defined so that the z(i) values are positive and λ is the transformation parameter. After suitable λ values are estimated

$$\left.\begin{aligned} z(i)^{(\lambda)} &= \lambda^{-1}[(z(i)+c)^{\lambda}-1] \quad \lambda \neq 0 \\ z(i)^{(\lambda)} &= \ln[z(i)+c] \quad \lambda = 0 \end{aligned}\right\} \tag{1}$$

the data are standardized so that they would have zero mean and variance unity. The values of λ are also listed in Table 1.

The rescaled range parameter h shown in Table 1 would be 0.5 for white noise (uncorrelated) series. The fact that they are different from 0.5 indicates the extent of non-stationarity in the series. Three of the six series used in the study do not need transformation and C is zero for all the series.

Thus the data used are non-Gaussian in many cases and indicate effects of non-stationarity in all the cases. The transformed data are, however, approximately Gaussian.

3. FN AND ARMA MODELS

A zero mean Gaussian stochastic process x(·)

TABLE 1. Data Used in the Study.

Code	Type	Location	Period	N*	λ**	$\hat{h}$**	$\hat{\rho}_1$
Bristlecone	Pine	California	800-1963	1164	1.414	0.695	0.43
Bigcone	Spruce	California	1458-1966	509	1.00	0.611	0.52
Navajo	Douglas Fir	Arizona	1263-1962	700	1.00	0.653	0.47
Eaglecol	Douglas Fir	Colorado	1107-1964	858	0.624	0.761	0.38
Espanola	Mud Varves	Canada	-471--820	350	0.0	0.928	0.22
England	Temperature	Central England	1658-1952	255	1.0	0.694	0.5

* N = number of observations.
**λ and h values are taken from Hipel and McLeod (1978).

in continuous time is said to be a fractional Brownian motion process with parameter h if the increments of the process are stationary and (2) is satisfied.

$$E[(x(t+s) - x(t))^2] = c\, s^{2h} \qquad (2)$$

A formal derivative of the process $x(\cdot)$ is called the fractional noise (FN) process in continuous time (Mandelbrot and Van Ness (1968)). A discrete time approximation to the continuous time fractional noise process leads to the discrete fractional noise sequence used in this paper. One such approximation $z(\cdot)$ in the form of a very large order moving average process is given below (Mandelbrot and Wallis (1969)).

$$z(i) = \sum_{k=1}^{M} w_1(i-k)/k^{1.5-h}, \quad 0.5 \le h \le 1 \qquad (3)$$

In (3), M is a large integer (theoretically infinity). The sequence $w_1(\cdot)$ is zero mean and independent with a normal distribution $N(0,\rho_1)$. Climatological time series (with mean removed) can be approximated by the stationary sequence $\{z(i)\}$ with suitably chosen values of h and ρ_1, which are the two parameters in (3).

The correlation properties of the sequence $z(\cdot)$ are given by (4).

$$\left.\begin{aligned} E[z^2(i)] &= \sum_{j=1}^{M} \left(\rho_1/j^{3-2h}\right) \\ E[z(i)\, z(i-k)] &= \sum_{j=k+1}^{M} \rho_1/(j(j-k))^{1.5-h} \end{aligned}\right\}. \qquad (4)$$

Hence $\phi(k)$, the k^{th} correlation coefficient of the process $z(\cdot)$ is given in (5)

$$\phi(k) = \sum_{j=k+1}^{M} (j(j-k))^{1.5-h} \Big/ \sum_{j=1}^{M} j^{3-2h} \qquad (5)$$

The parameter h may be estimated by maximum likelihood (ML) method (Hipel and McLeod (1978)) or by graphical procedures (Mandelbrot and Wallis (1969)). The ML estimates $\hat{h}$ for the data used in the present study are listed in Table 1. By using these estimates $\hat{h}$, values of $\hat{\rho}_1$ may be estimated by (6) and these are also listed in Table 1. Thus the FN models M_1 have the structure given in

$$\hat{\rho}_1 = ((1/N) \sum_{i=1}^{N} z^2(i)) / \{ \sum_{j=1}^{M} (1/j^{3-2h}) \} \qquad (6)$$

7. In the present study M is arbitrarily chosen to be equal to 2N, where N is the number of observations.

$$M_1: z(i) = \sum_{j=1}^{M} w_1(i-j)/j^{1.5-h}; \; w_1 \sim N(o,\rho_1) \qquad (7)$$

The ARMA(p,q) models M_0 used in the study have the well known (Box and Jenkins (1970), Kashyap and Rao (1976)) form in (8). The parameter

$$M_0: z(i) = \sum_{j=1}^{p} \alpha(j)\, z(i-j) + \sum_{j=1}^{q} \beta(j)\, w(i-j) + w(i) \qquad (8)$$

estimates $\hat{\alpha}(\cdot)$ and $\hat{\beta}(\cdot)$ fitted to transformed data are listed in Table 2 along with the variance $\hat{\rho}$ of the residual sequence w(i). The mean of w(i) is zero and as z(i) are transformed so that they would be Gaussian, w(i) are normally distributed with mean zero and variance $\hat{\rho}$.

TABLE 2. Parameters of AMRA Models and Residual Variances.

Name	$\hat{\alpha}(1)$	$\hat{\alpha}(2)$	$\hat{\beta}(1)$	$\hat{\beta}(2)$	$\hat{\rho}$
Bristlecone	0.64		-0.40		0.92
Bigcone	0.38	0.16			0.78
Navajo	0.68		-0.42		0.89
Eaglecol	1.16	-0.24	-0.69		0.65
Espanola	0.96		-0.54		0.29
England			0.12	0.20	0.99

The FN and ARMA models in eqs. 3 and 8 may be simulated as the parameters of these models are known and Gaussian white noise sequences of prespecified means and variances are easily generated. These generated data sequences are also used in testing the suitability of models. The correlation coefficients $\theta(k)$ of ARMA models are estimated by using well known formulas (Box and Jenkins (1970)).

4. BAYESIAN DECISION THEORY

The models M_0 and M_1 which are compared are given below.

$$M_0: z(i) = \sum_{j=1}^{p} \alpha(j)\, z(i-j) + \sum_{j=1}^{q} \beta(j)\, w(i-j) + w(i)$$

$$M_1: z(i) = \sum_{j=1}^{M} w_1(i-j)/j^{1.5-h}$$

$w(\cdot)$ and $w_1(\cdot)$ are sequences of normal, independently distributed (NID) random variables with means zero and variances ρ and ρ_1. The parameters p, q, $\alpha(j)$, $\beta(j)$, ρ, h, and ρ_1 are estimated by using methods mentioned above. M is chosen to be 2N.

The empirical data for testing the validity of the models is the sequence z(1), ..., z(N). As mentioned earlier, the observations have been normalized such that

$$\frac{1}{N}\sum_{i=1}^{N} z(i) = 0, \quad \frac{1}{N}\sum_{i=1}^{N} z^2(i) = 1.$$

Let $\underset{\sim}{z}(N_1, N_2) \triangleq [z(N_1), z(N_1+1), \ldots, z(N_2)]^T$

The Bayesian decision theory has been well documented in many books such as that by Pratt et al. (1965). The key feature in the Bayesian analysis is the posterior probabilities $P(M_0|\underset{\sim}{z}(1,N))$ and $P(M_1|\underset{\sim}{z}(1,N))$ where $P(M_0|\underset{\sim}{z}(1,N))$ is the probability that the model M_0 is the correct model on the basis of the observations $\underset{\sim}{z}(1,N)$ and $P(M_1|\underset{\sim}{z}(1,N))$ is defined similarly.

Any decision involves errors and the corresponding losses under the various contingencies can be represented by the following matrix

		Action	
		M_0	M_1
True Model		d_{00}	d_{01}
		d_{10}	d_{11}

d_{ij} = Loss caused by choosing the model H_j to represent the given data when M_i is the correct model.

Usually: $d_{00}=0$, $d_{11}=0$, $d_{01}=1$, $d_{10}=1$ (9)

Clearly the decision rule which yields the least value of the average loss should be chosen. The average losses for the two possible decisions are given below.

J_0 = Average loss suffered by declaring the model M_0 to be correct on the basis of data $\underset{\sim}{z}(1,N)$

$$= \{d_{00}\, P(M_0|\underset{\sim}{z}(1,N)) + d_{10} P(M_1|\underset{\sim}{z}(1,N))\} \quad (P(\underset{\sim}{z}(1,N)) \quad (10)$$

J_1 = Average loss suffered by declaring the model M_1 to be correct on the basis of the data $\underset{\sim}{z}(1,N)$

$$= \{d_{01}\, P(M_0|\underset{\sim}{z}(1,N)) + d_{11}\, P(M_1|\underset{\sim}{z}(1,N))\} \quad P(\underset{\sim}{z}(1,N)) \quad (11)$$

Hence the optimal decision rule, i.e., the decision rule which leads to the least average loss is given below.

$$\left.\begin{array}{l} J_0 < J_1 = \text{ choose } M_0 \text{ as the model} \\ J_0 > J_1 = \text{ choose } M_1 \text{ as the model} \end{array}\right\} \quad (12)$$

By (9) and (10), we get (13).

$$\frac{J_0}{J_1} < 1 = \frac{P(M_0|\underset{\sim}{z}(1,N)}{P(M_1|\underset{\sim}{z}(1,N)} < \frac{d_{10} - d_{11}}{d_{01} - d_{00}} \quad (13)$$

Hence combining (12) and (13) with the loss matrix in (9), we get the following optimal decision rule

$$\left.\begin{array}{l} \dfrac{P(M_0|\underset{\sim}{z}(1,N))}{P(M_1|\underset{\sim}{z}(1,N))} < 1 = \text{ choose } M_1 \\ \dfrac{P(M_0|\underset{\sim}{z}(1,N))}{P(M_1|\underset{\sim}{z}(1,N))} > 1 = \text{ choose } M_0 \end{array}\right\} \quad (14)$$

The decision rule in (14) is *a priori* reasonable. It simply asks us to choose the model which has the greatest posterior probability.

The posterior probabilities are computed as follows. Let $P(M_i)$, i=0,1 be the prior probability that M_i is the correct model.

$$P(M_0) + P(M_1) = 1$$

The prior probabilities represent our degrees of belief in the validity of the two models before inspecting the data $\underset{\sim}{z}(1,N)$. Let $P(M_0)$ and $P(M_1)$ be 0.5, implying that we have no prior preference for either model.

Let $P(M_i|\underset{\sim}{z}(1,N))$ be the posterior probability that the model M_i is correct on the basis of the observations $\underset{\sim}{z}(1,N)$. The posterior probabilities are computed with the aid of the Bayes law given below

$$P(M_i|\underset{\sim}{z}(1,N)) = P(\underset{\sim}{z}(1,N)|M_i)\, P(M_i)/(P(\underset{\sim}{z}(1,N)|M_0)\, P(M_0) + P(\underset{\sim}{z}(1,N)|M_1)\, P(M_1))$$

By definition the sum $P(M_1|\underset{\sim}{z}(1,N)) + P(M_0|\underset{\sim}{z}(1,N))$ is equal to one. By definition of model M_0, we can write the expression for $P(\underset{\sim}{z}(1,N)|M_0)$.

$$P(\underset{\sim}{z}(1,N)|M_0) = \frac{1}{(2\pi)^{N/2}|\text{Det } \underset{\sim}{A}_N|^{1/2}} \exp[-\tfrac{1}{2}(\underset{\sim}{z}(1,N))^T \underset{\sim}{A}_N^{-1} \underset{\sim}{z}(1,N)] \quad (15)$$

$\underset{\sim}{A}_N$ is the covariance matrix of (z(1),...,z(N)) according to Model M_0.

$$\underset{\sim}{A}_N = E\left[\begin{pmatrix} z(1) \\ z(2) \\ \cdot \\ \cdot \\ \cdot \\ z(N) \end{pmatrix} (z(1), z(2), \ldots, z(N))^T \,\middle|\, M_0\right]$$

$$(A_N)_{ij} = \theta(|i - j|); \quad i, j = 1, \ldots, N$$

$\theta(k)$, k = 0,1,2,..., are the correlation coefficients of the model M_0.

Similarly, let $\underset{\sim}{B}_N$ be the covariance matrix of the vector (z(1),...,z(N)) according to the model M_1

$$(B_N)_{i,j} = \phi(i - j), \; i, j = 1, 2, \ldots, N.$$

$\phi(k)$, k = 0,1,..., are the correlation coefficients of the model M_1, which are computed by using eq. 5. The expression for $P(\underset{\sim}{z}(1,N)|M_1)$ is given below.

$$P(\underset{\sim}{z}(1,N)|M_1) = \frac{1}{(2\pi)^{N/2}|\text{Det } \underset{\sim}{B}_N|^{1/2}} \exp\left[-\frac{1}{2}\underset{\sim}{z}(1,N)^T \underset{\sim}{B}_N^{-1} \underset{\sim}{z}(1,N)\right] \quad (16)$$

Substitution of posterior probabilities (15) and (16) in (14) yields the required decision rule.

5. RESULTS FROM BAYESIAN TEST

Typical results from the Bayesian test are given in Table 3. Determinants of matrices $\underset{\sim}{A}_N$ and $\underset{\sim}{B}_N$ and the posterior probability ratios $P(M_0|\underset{\sim}{z})/P(M_1|\underset{\sim}{z})$ are also listed in Table 3. The prior probabilities $P(M_0)$ and $P(M_1)$ are assumed to be equal to 0.5. Data segments, each having 50 observations, are analyzed separately in order to estimate the variability in the results which is mainly due to choosing only a small portion of the data for analysis.

A significant feature of results presented in Table 3 is the variability in results from different sets of data. This is yet another demonstration of the well known fact that results from statistical analysis of short data sequences can be quite misleading.

The important aspect of the results given in Table 3 as far as the main objective of present investigation is concerned is that in most of the cases ARMA models are preferred to FN models. Average values of $P(M_0|\underset{\sim}{z})/P(M_1|\underset{\sim}{z})$ ratios for Bristlecone, Bigcone, Navajo, Eaglecol, Espanola and England are respectively 1.27, 40.27, 1.76, 6.59, 19.69 and 2.07. Although there is a stronger preference for ARMA models in some cases (Bigcone, Espanola) in none of the cases is the FN model preferred. Consequently on the basis of these results ARMA models are adequate to characterize climatic data such as those discussed herein.

The rescaled-range and correlation characteristics of data generated by using FN and ARMA models are also investigated and the results are graphically presented in Figures 1 and 2. Data sequences are generated by using FN and ARMA models M_1 and M_0 and their estimated parameters. Correlograms and rescaled range-lag plots of observed and generated data are shown in Figures 1 and 2 respectively for Bigcone, Espanola, Bristlecone and Navajo series. As discussed above, the Bayesian test results indicate that ARMA models are strongly preferred to FN models for Bigcone and Espanola and a weaker preference for Bristlecone and Navajo data. The results given in Figures 1 and 2 are needed to examine the correspondence between the results of the Bayesian test and the empirical characteristics of the data and are hence presented herein.

The correlograms of data generated by FN and ARMA models show characteristics which are similar to those of Bristlecone and Navajo data. For Bigcone and Espanola series, the correlograms of data generated by the ARMA model are closer to the correlograms of observed data than that computed by FN data. Thus the comparison of correlograms of generated and observed data support, in general, the decisions given by Bayesian test.

The rescaled range-lag characteristics of Bigcone and Espanola data, along with the data generated by ARMA models shown in Figure 2 also show similar general agreement. Even the curvilinear variation of rescaled range values of observed data and of data generated by ARMA models are in general agreement. A similar agreement between rescaled range values of data from ARMA models and of observed data is evident for Navajo and Bristlecone series.

In conclusion, ARMA models which are indicated by the Bayesian tests as being superior to FN models, have correlation and rescaled range-lag characteristics which are similar to those of observed data. This provides additional evidence supporting the use of ARMA models for climatological data.

TABLE 3. Bayesian Test Results.

	NAVAJO $\lvert A_N \rvert = 3.09\times10^{-3}$; $\lvert B_N \rvert = 1.38\times10^{-2}$; N = 700				EAGLECOL $\lvert A_N \rvert = 5.52\times10^{-10}$; $\lvert B_N \rvert = 4.37\times10^{-6}$; N = 858			
(N_1,N_2)	$z^T A_N^{-1} z$ x10	$z^T B_N^{-1} z$ x10	$P(M_1\|z)$ $\times10^{-1}$	$\frac{P(M_0\|z)}{P(M_1\|z)}$	$z^T A_N^{-1} z$ x10	$z^T B_N^{-1} z$ x10	$P(M_1\|z)$ $\times10^{-1}$	$\frac{P(M_0\|z)}{P(M_1\|z)}$
1,50	6.46	6.38	4.13	1.42	11.6	10.3	8.93	1.20×10^{-1}
51,100	5.04	4.72	7.04	4.21×10^{-1}	10.5	8.78	9.83	1.72×10^{-2}
101,150	2.43	2.59	1.78	4.60	8.92	7.28	9.76	2.45×10^{-2}
151,200	2.44	2.41	3.60	1.78	4.01	3.15	4.48	1.23
201,250	4.43	4.32	4.52	1.21	3.79	3.15	2.13	3.69
251,300	5.64	5.41	5.94	6.84×10^{-1}	5.35	4.93	0.85	10.76
301,350	3.07	3.01	3.95	1.53	3.31	2.73	1.69	4.92
351,400	5.47	5.38	4.25	1.35	4.39	4.00	0.73	12.71
401,450	5.09	4.98	4.52	1.21	5.54	4.80	3.10	2.23
451,500	6.22	6.12	4.43	1.26	5.16	4.52	2.16	3.62
501,550	6.16	6.08	4.20	1.38	2.41	1.99	0.87	10.50
551,600	6.09	5.98	4.51	1.22	3.60	3.20	0.75	12.43
601,650	5.91	5.74	5.21	9.20×10^{-1}	1.84	1.56	0.44	21.84
651,700	5.69	5.89	1.49	5.73	4.76	3.90	4.54	1.20
701,750					4.16	3.40	3.33	2.00
751,800					2.31	1.94	0.67	13.96
801,850					4.29	3.87	0.85	10.83

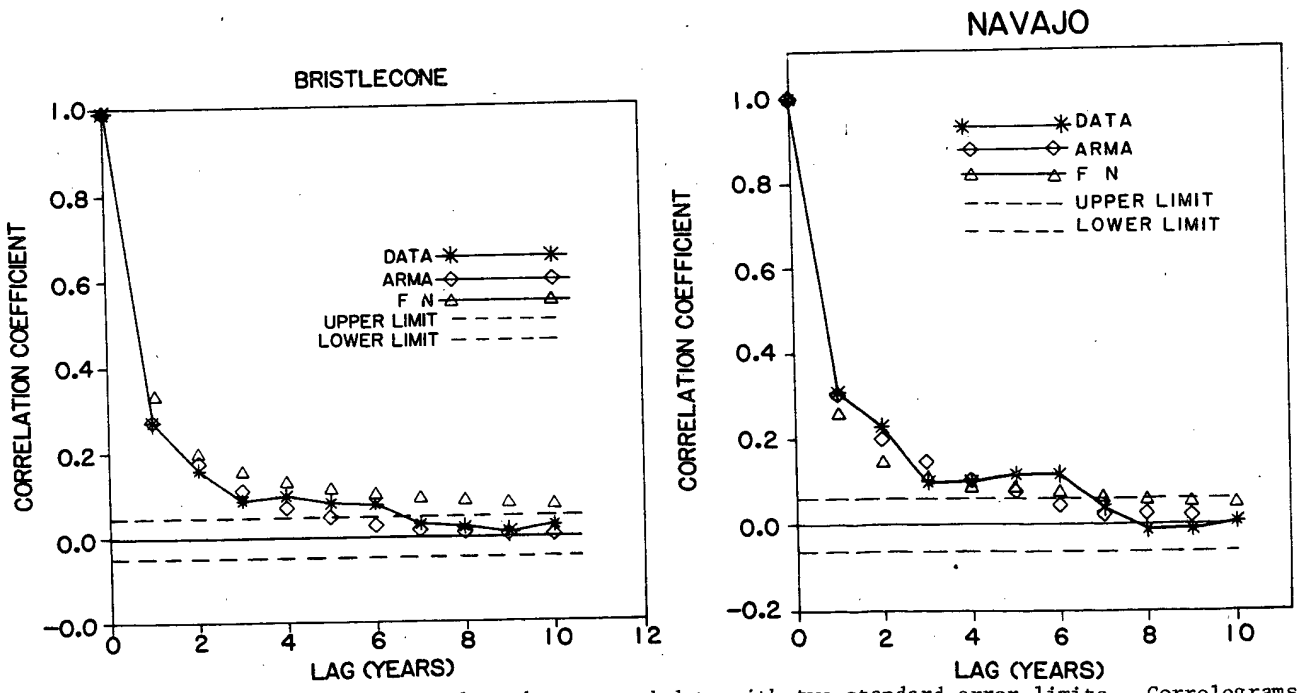


Figure 1. Correlograms of observed and generated data with two standard error limits. Correlograms of white noise would be within the error limits 95% of the time.

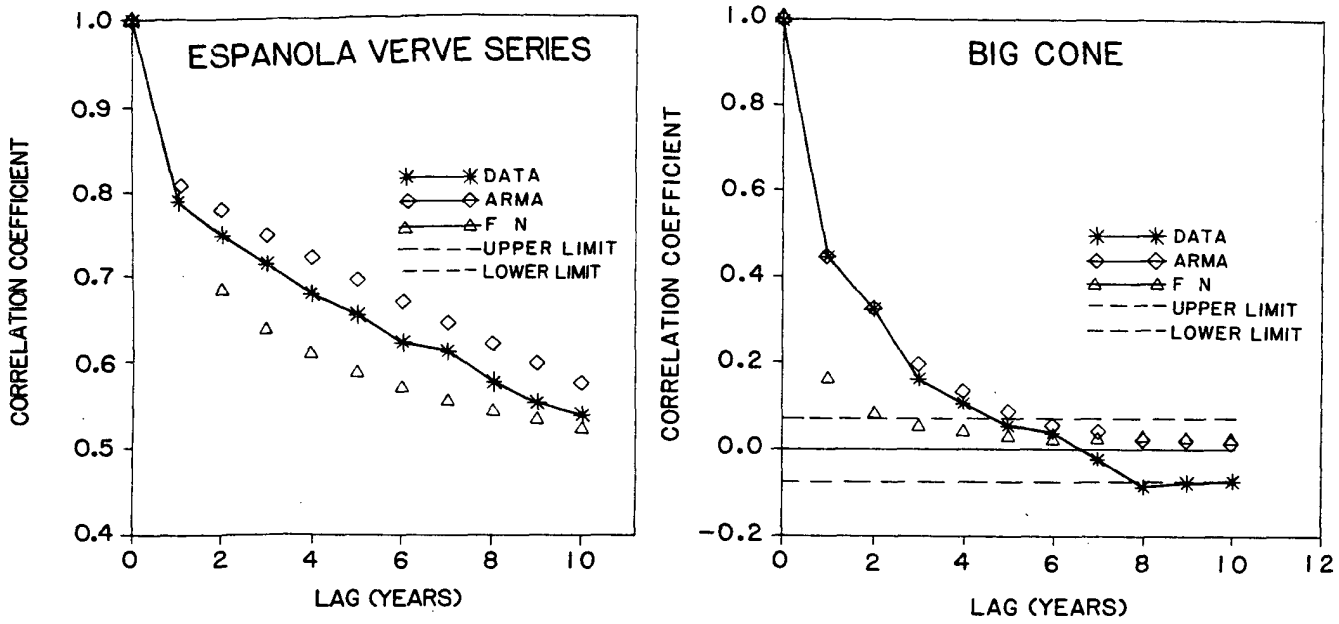


Figure 1. Continued

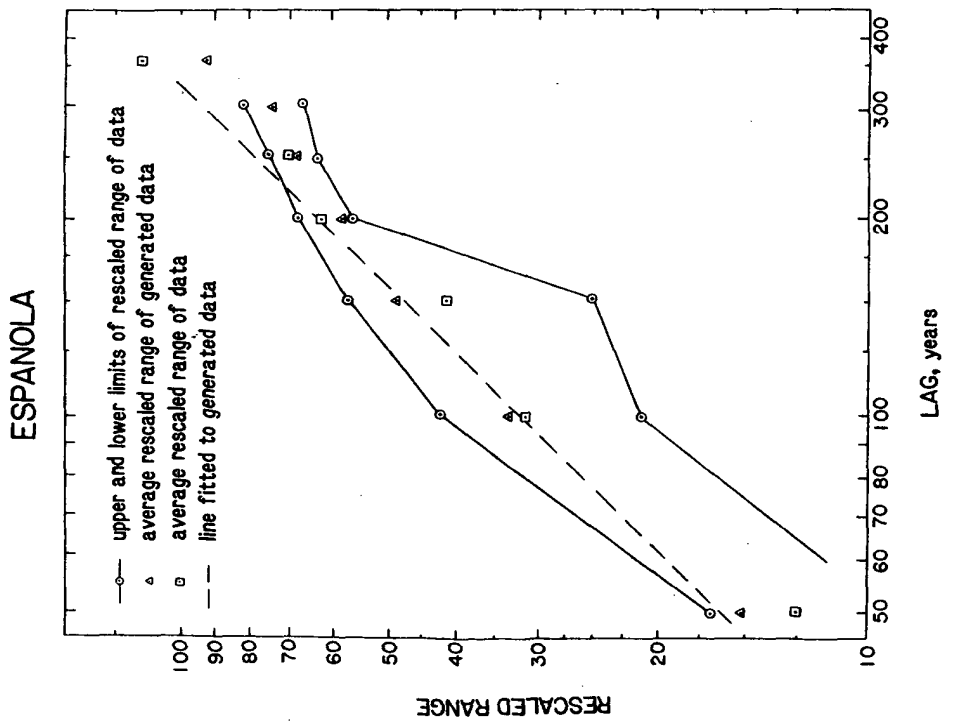


Figure 2. Rescaled range - lag characteristics of observed and generated data. Data are generated by ARMA models.. The upper and lower limits are estimated from observed data.

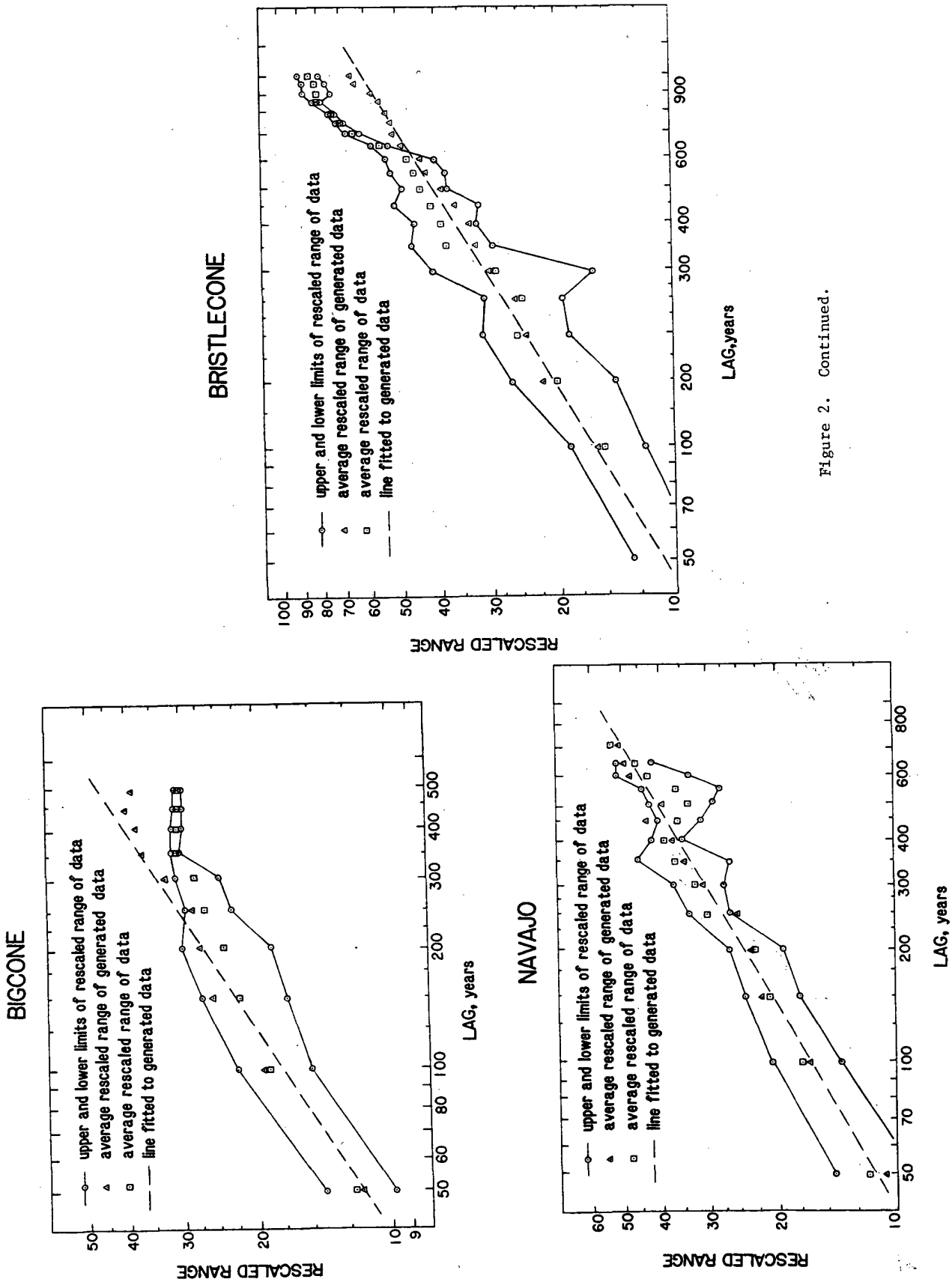


Figure 2. Continued.

6. REFERENCES

Box, G.E.P. and D.R. Cox, "An analysis of transformations", Jour. Royal Statist. Soc., Ser. B, 26, 211-252, 1964.

Box, G.E.P. and G.M. Jenkins, "Time series analysis: forecasting and control", Holden-Day, San Francisco, California, 1970.

Hipel, K.W. and A.I. McLeod, "Preservation of the Rescaled adjusted range 2. Simulation studies using Box-Jenkins models", Water Resources Research, 14(3), 509-516, 1978.

Kashyap, R.L. and A.R. Rao, "Dynamic stochastic models from empirical data", Academic Press, New York, 1976.

Klemeš, V., "The Hurst phenomenon: A puzzle?", Water Resources Research, 10(4), 675-688, 1974.

Mandelbrot, B.B. and J.R. Wallis, "Noah, Joseph and operational hydrology", Water Resources Research, Vol. 4, No. 5, pp. 900-918, 1968.

Mandelbrot, B.B. and J.W. Van Ness, "Fractional Brownian motions, fractional noises and applications", SIAM Rev., 4(10), 422-437, 1968.

Mandelbrot, B.B. and J.R. Wallis, "Computer experiments with fractional noises", Water Resources Research, 5(1), 228-267, 1969.

Manley, G., "The mean temperatures of Central England (1698-1952)", Quart. J. Roy. Meterol. Soc., 79, 242-261, 1953.

Potter, K.W., "Evidence for non-stationarity as a physical explanation of the Hurst phenomenon", Water Resources Research, 12(5), 1047-1052, (1976).

Pratt, J.W., H. Raiffa, and R. Schlaifer, "Introduction to statistical decision theory", McGraw-Hill, New York, 1965.

A MARKOV MODEL FOR VARIATIONS IN CLOUD COVER

Henrik Madsen, Henrik Spliid and Poul Thyregod

The Institute of Mathematical Statistics and Operations Research
The Technical University of Denmark
DK-2800 Lyngby, Denmark

1. INTRODUCTION

Within topics as building insulation and heating research, there are a great number of projects concerned with the construction of simulation models of the climate parameters of interest in connection with heat dynamics of buildings. Some of these models consider cloud cover as an independent variable, so that variation in radiation components and air temperature partly depends on cloud cover variations. In the present study we shall discuss estimation of models for cloud cover suitable for the above mentioned simulation.

An analysis of 15 years hourly observations of cloud cover shows, that a main part of the variations can be described by 1. order homogeneous Markov models. Models in discrete and in continuous time are considered, and the estimation of the parameters is described. Due to physical conditions for transitions of cloud cover, the matrix of transition rates belonging to the embedded Markov process in continuous time is restricted to have a very simple structure. A maximum likelihood estimate of the matrix of transition rates and an indication of the possible estimation error under the model are found numerically.

The main purpose of the present paper is to discuss the estimation of the matrix of transition rates, given a specific structure of the matrix. Furthermore the resulting models describe important parts of the variation of the cloud cover.

2. DATA DESCRIPTION

The research is based on 15 years of hourly cloud cover observations, taken from the synoptic observation station at Værløse Airport about 15 km north-west of Copenhagen, Denmark. The observations are total cloud cover in oktas and follow the WMO code (WMO, 1972, 1974). Thus the cloud cover is measured as an integer between 0 and 9, where 0 corresponds to completely clear sky, 8 to completely overcast sky and the observation 9 indicates that the cloud cover is unobservable, which usually is the case in foggy weather or heavy snow-fall. The observations are based on a subjective evaluation and as such they can be encumbered with judgemental errors.

3. DIURNAL AND ANNUAL VARIATION

Below the annual variation of the cloud cover mean is illustrated. The mean is calculated as the average over the 24 hours of the day and over the 15 years.

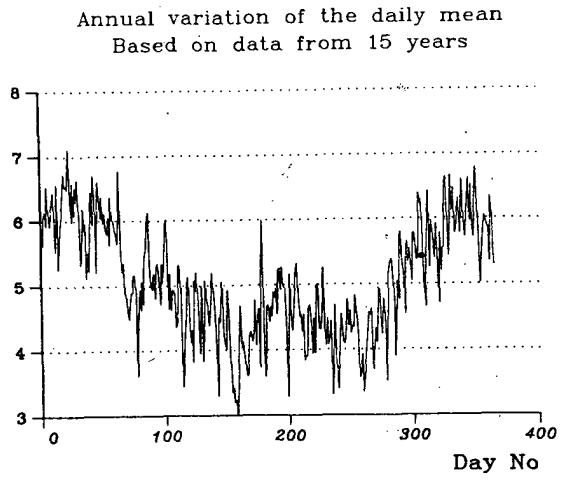


Fig. 1. Annual variation of the mean.

Figure 1 shows that the winter is far more cloudy than the summer. Further the figure indicates that it is more cloudy in July than in the beginning of September.

In order to illustrate the annual variation of the distribution of cloud cover, the next two figures show the distributions in January and June. These two distributions are extremes compared to the distributions of the rest of the months, which vary smoothly between the January and June distributions.

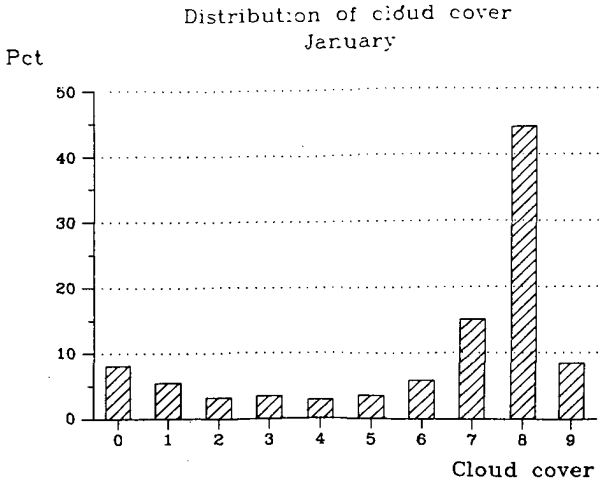


Fig. 2. Distribution of cloud cover. January.

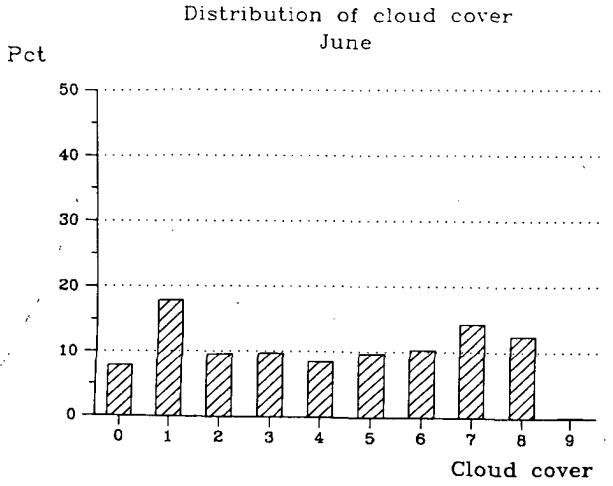


Fig. 3. Distribution of cloud cover. June.

The diurnal variation of the cloud cover mean is illustrated on fig. 4. The mean is based on all twelve months and the 15 years.

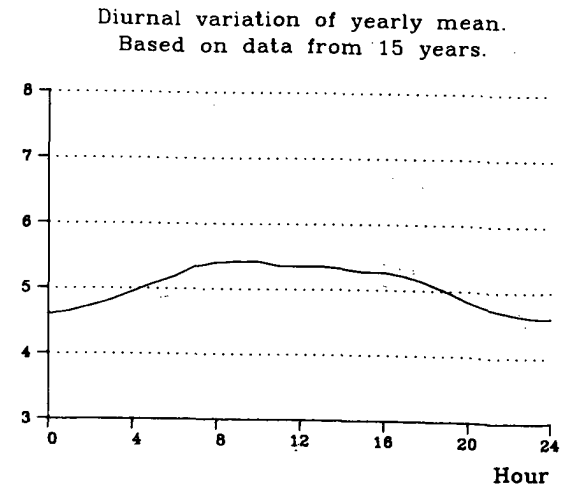


Fig. 4. Diurnal variation of the mean.

The figure of the diurnal variation of the cloud cover mean shows that it is more cloudy during the day than at night, but the variation is smaller than the yearly variation.

The diurnal variation of the distribution of cloud cover is not shown, but an examination has shown that the diurnal variation consists of a higher frequency of clear sky categories (0-1) at night than during the day. Conversely the frequency of middle categories are smaller at night than during the day. The frequency of overcast sky (7-8) is nearly constant during day and night. Finally foggy weather (9) is much more frequent at night than at daytime.

4. HOURLY CORRELATIONS

4.1 Autocorrelation functions

A useful tool for illustrating the hourly correlations are the autocorrelation functions (e.g. Box and Jenkins, 1976). The estimated autocorrelation and partial autocorrelation function based on all observations are stated in fig. 5 and fig. 6 respectively.

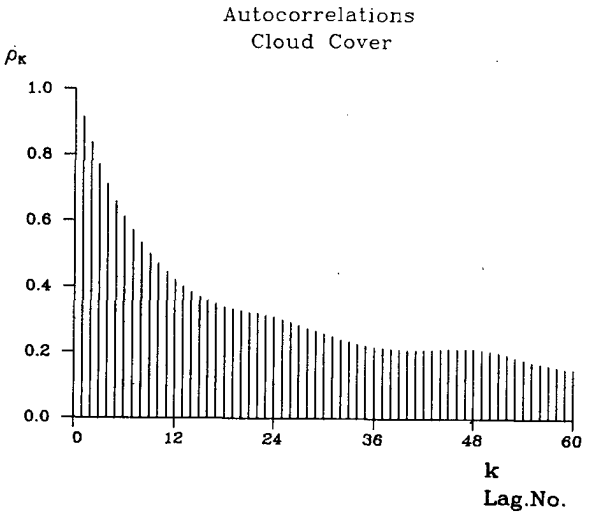


Fig. 5. Estimated autocorrelation function.

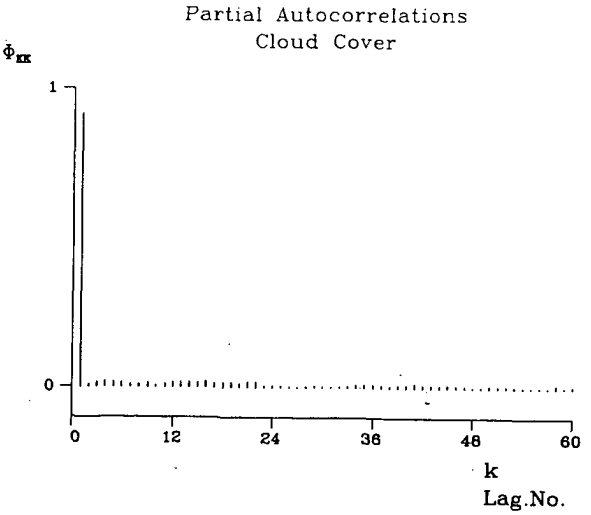


Fig. 6. Estimated partial autocorrelation function.

The regularity of the autocorrelation function is due to the great number of observations. The autocorrelations die out exponentially, disturbed only by a weak 24 hours oscillation. This oscillation is most likely due to the diurnal variation of cloud cover previously shown. In agreement with the almost exponentially dying out of the autocorrelation function, the partial autocorrelations $\hat{\phi}_{kk}$ are small for lags k greater than one. A simple test is performed in order to decide whether $\hat{\phi}_{kk}$ are small for lags greater than one. For the well-known AR(1)-process $\phi_{kk} = 0$ for $k \geq 2$. In this case an estimate of the variance of the estimator is (Box and Jenkins, 1976)

$$\hat{V}(\hat{\phi}_{kk}) \simeq \frac{1}{n} \qquad , \forall\ k \geq 2$$

where n is the total number of observations. The total number of observations is in this case 131496, thus

$$\hat{V}(\hat{\phi}_{kk}) \simeq (0.0028)^2 , \forall\ k \geq 2$$

For lags between 2 and 48 the number of $\hat{\phi}_{kk}$ for which $\hat{\phi}_{kk} \geq 0.006$ is 35. Thus we must conclude that probably not all the information about the future is contained in only the last observation.

But since the partial autocorrelations for lags greater than one still are small compared with $\hat{\phi}_{11}$, it is natural as a first approach to consider a first order model in the construction of a simulation model. When doing so we neglect the weak 24 hour oscillation in the autocorrelation function. Hence the class of possible processes are enclosed by first order homogeneous Markov processes.

Using common time series formulation one of the most simple processes in this class is the AR(1)-process

$$(X_t - \mu) = A_t + \varphi(X_{t-1} - \mu)$$

but this model does not qualify as a simulation model, since e.g. negative values of cloud cover will occur.

The diurnal variation is probably a main reason why the data show some dependence or partial correlation for lags greater than one. Thus an improved first order model must include time changing parameters. Due to this it could be useful to consider inhomogeneous Markov processes. In section 7 we shall briefly introduce such inhomogeneous Markov models for cloud cover. But, as mentioned, we shall restrict the attention to first order homogeneous Markov models in this paper. This approach also serves the main purpose of this paper, which is to describe the estimation of parameters of the embedded Markov process in continuous time given a specific structure of the process, and this description becomes more simple in the homogeneous case. The drawback of the homogeneous model is, as mentioned, that the slight diurnal variation and its effect on the hourly correlations is neglected.

5. HOMOGENEOUS MARKOV MODELS

5.1 Discrete time

Since the cloud cover is measured by the integers 0 through 9, it is reasonable to try to describe the variations by a first order Markov chain model. The resulting autocorrelation function of that model will die out exponentially, thus the partial autocorrelations will be zero for lags greater than one.

In the following the observations will be denoted

$$X_1, X_2, \ldots, X_N$$

where $X_t \in \{0,1,\ldots,9\}$ denotes the state or the cloud cover category at time t.

Considering the homogeneous Markov process the one-hour transition probabilities are independent of time and defined as (e.g. Cox and Miller, 1980)

$$p_{jk} = \text{prob}\{X_{t+1} = k \mid X_t = j\}$$

from which the matrix of transition probabilities is found as

$$P = \{p_{jk}\}$$

The conditional likelihood function becomes

$$L(P) = \prod_j \prod_k p_{jk}^{n_{jk}}$$

where n_{jk} is the number of jumps from j to k in the series of observations. The ML (Maximum Likelihood) estimate for p_{jk} is (Andersen and Goodman (1957))

$$\hat{p}_{jk} = \frac{n_{jk}}{n_{j\bullet}}$$

where $n_{j\bullet} = \sum_k n_{jk}$

Previously we mentioned the marked annual variation of the cloud cover, thus the transition probabilities depend on the season. As a compromise between a desirable amount of data and consideration to the annual variation, we have chosen to estimate a model for every month. Thus the ML-estimate of the matrix of transition probabilities is found for every month in the year. n_{jk}, $n_{j\bullet}$ and $\hat{p}_{jk}$ for January are shown in table 1 on the next page.

There is an obvious dependence between two successive observations. The probability of observing the same or a neighbouring cloud cover in successive observations is high, thus the transition probabilities tend to be largest in the diagonal. The larger the transition the smaller is the corresponding probability.

j \ k	0	1	2	3	4	5	6	7	8	9	
0	747 0.835	86 0.096	17 0.019	15 0.017	5 0.006	5 0.006	5 0.006	4 0.004	2 0.002	9 0.010	895
1	95 0.156	346 0.567	71 0.116	39 0.064	18 0.030	13 0.021	12 0.020	10 0.016	3 0.005	3 0.005	610
2	24 0.067	85 0.238	105 0.294	71 0.199	20 0.056	17 0.048	15 0.042	17 0.048	1 0.003	2 0.006	357
3	9 0.024	41 0.108	71 0.187	97 0.256	51 0.135	43 0.113	38 0.100	22 0.058	6 0.016	1 0.003	379
4	4 0.013	13 0.041	37 0.117	47 0.149	71 0.225	54 0.171	52 0.165	33 0.104	4 0.013	1 0.003	316
5	4 0.010	15 0.039	23 0.060	37 0.096	48 0.124	76 0.197	97 0.251	70 0.181	13 0.034	3 0.008	386
6	3 0.005	8 0.012	14 0.022	31 0.048	45 0.069	90 0.138	222 0.342	187 0.288	45 0.069	5 0.008	650
7	3 0.002	5 0.003	16 0.009	23 0.014	41 0.024	62 0.037	138 0.082	1001 0.594	386 0.229	11 0.007	1686
8	3 0.001	6 0.001	3 0.001	15 0.003	16 0.003	24 0.005	63 0.013	327 0.066	4295 0.872	173 0.035	4925
9	4 0.004	2 0.002	0 0.0	3 0.003	3 0.003	2 0.002	8 0.009	14 0.015	171 0.182	734 0.780	941

Table 1. One-hour transition probabilities. January.

In discrete time a simulation model is directly formulated on the monthly estimated matrices of transition probabilities. The disadvantage of the model in discrete time is the great number of parameters, namely 90 for each month. In addition it is not obvious how the annual variation of the estimates can be smoothed.

5.2 Continuous time

5.2.1 Model and assumptions

In continuous time the required number of parameters is reduced due to the fact that cloud cover does not jump momentarily, e.g. from state 2 to state 5. The cloud cover changes gradually. In general only jumps to neighbouring states are possible in a short time. But there are two exceptions. Regarding the matrix of transition probabilities in table 1 of the previous section we observe, that the transition probability from state 0 to state 9 is higher than the transition probability from state 0 to e.g. state 8. Contrary the probability for transition from state 9 to state 0 is higher than the probability for transition from state 9 to state 1. This can be explained physically by the fact, that foggy weather (state 9) can replace clear sky (state 0) after sunset due to radiative cooling of the lower air stratum. Transitions from state 9 to state 0 are possible due to radiative heating after sunrise.

Introducing

$$p_{i,j}(\Delta t) = \text{prob}\{X_t = j \,|\, X_{t-\Delta t} = i\}$$

the observations above lead to a model on the following assumptions for $\Delta t \to 0$:

$$p_{i,j}(\Delta t) = o(\Delta t) \qquad |i-j| \geq 2$$

$$p_{i,i}(\Delta t) = 1 - q_i \, \Delta t + o(\Delta t)$$

$$p_{i,i-1}(\Delta t) = w_i \, q_i \, \Delta t + o(\Delta t)$$

$$p_{i,i+1}(\Delta t) = (1 - w_i) q_i \, \Delta t + o(\Delta t)$$

$$i \in \{0,1,\ldots,9\}$$

where q_i is the transition rate (intensity for leaving), and w_i is the frequency of transitions to a lower cloud cover category. The notation include transitions from 0 to 9 and reverse, if we define state(-1) = state(9) and state(10) = state(0).

Due to the described assumptions the model is limited to be a first order homogeneous Markov process in continuous time, and the assumptions lead directly to Kolmogorov's differential equation (e.g. Cox and Miller, 1980)

$$\frac{d\,P(t)}{dt} = P(t)Q \quad ; \quad P(0) = I$$

where $P = \{p_{i,j}(t)\}$, and the matrix of transition rates

$$Q = \begin{bmatrix} -q_0 & (1-w_0)q_0 & 0 & \cdots & 0 & w_0 q_0 \\ w_1 q_1 & -q_1 & (1-w_1)q_1 & \cdots & 0 & 0 \\ & \ddots & \ddots & \ddots & & \\ 0 & \cdots & w_i q_i & -q_i & (1-w_i)q_i & 0 \\ & & & \ddots & \ddots & \\ (1-w_9)q_9 & 0 & 0 & & w_9 q_9 & -q_9 \end{bmatrix}$$

Given the matrix of transition rates, the matrix of transition probabilities is found by solving Kolmogorov's differential equation, thus

$$P(t) = e^{Qt} \, P(0) = e^{Qt}$$

Compared with the discrete time model the number of parameters is reduced from 90 to 20.

5.2.2 Simple estimates

Based on the monthly estimated transition probabilities some very simple estimates can be found.

A simple estimate of the transition rate q_i is based on the transition probabilities p_{ii}, as

$$\tilde{q}_i = -\ln \hat{p}_{ii}$$

This estimate does not take into account, that the cloud cover can leave a state and return to the same state within one hour, that is in the sampling interval. Thus, evidently, the simple estimate will underestimate q_i, but for $q_i \ll 1 \text{ h}^{-1}$ the error will be negligible.

Simple estimates for the frequency of transition to lower cloud cover are found as

$$\tilde{w}_i = \sum_{j=0}^{i-1} \hat{p}_{ij} / (1 - \hat{p}_{ii}) \qquad i = 1,2,\ldots,8$$

$$\tilde{w}_0 = (\hat{p}_{08} + \hat{p}_{09}) / (1 - \hat{p}_{00})$$

$$\tilde{w}_9 = (1 - \hat{p}_{90} - \hat{p}_{91}) / (1 - \hat{p}_{99})$$

For the same reason as above these estimates do not consider all the possible combinations of

transition within one hour.

The simple estimates are shown on fig. 7 and fig. 8.

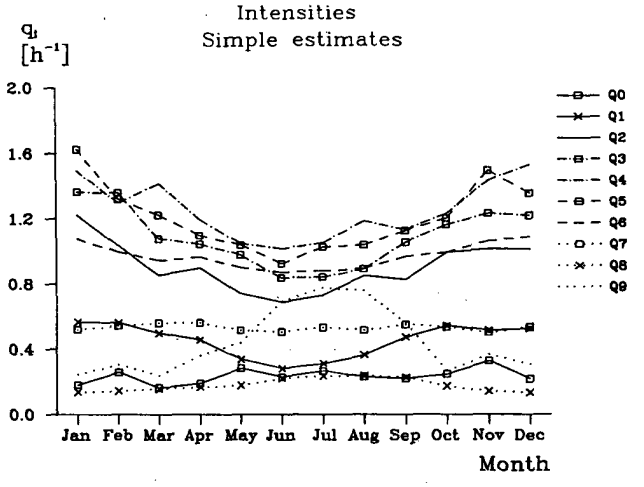


Fig. 7. Simple estimates of transition rates.

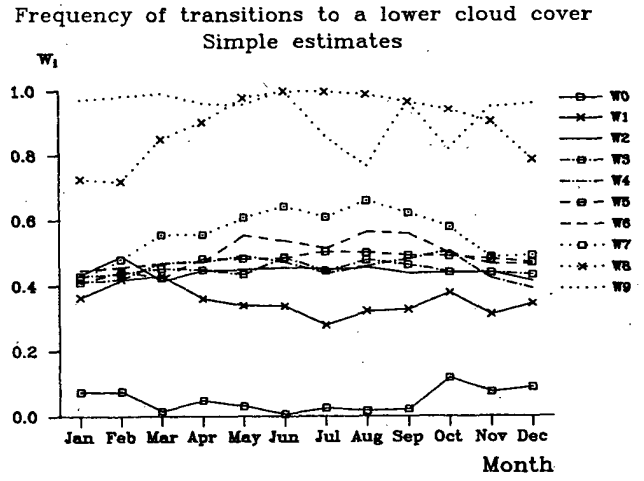


Fig. 8. Simple estimates of the frequency of transitions to a lower cloud cover category.

5.2.3 ML-estimates

Since the observations of the cloud cover are by the hour, the matrix of one-hour transition probabilities becomes (t = 1)

$$P = P(1) = e^{Q}$$

The logarithm of the conditional likelihood function is

$$\log L(P) = \sum_i \sum_j n_{ij} \log p_{ij}$$

where n_{ij} is the total number of jumps from i to j in the series of observations.

The ML-estimate is the set

$$\hat{\theta} = (\hat{q}_0, \ldots, \hat{q}_9, \hat{w}_0, \ldots, \hat{w}_9)$$

which maximizes the likelihood function. Further the ML-estimator $\hat{\theta}$ is asymptotically normally distributed with mean θ and variance

$$D = H^{-1}$$

where the matrix H is given by

$$\{h_{lk}\} = -E\left[\frac{\partial^2}{\partial\theta_l \, \partial\theta_k} \log L(P)\right]$$

An estimate of D is obtained by neglecting the expectation and applying

$$\{h_{lk}\} \simeq -\left(\frac{\partial^2}{\partial\theta_l \, \partial\theta_k} \log L(P)\right)\Big|_{\theta=\hat{\theta}}$$

Thus approximate confidence limits can be based on the normal distribution.

Since it has not been possible to determine an explicit expression for the ML-estimator $\hat{\theta}$, numerical methods have been used. Important problems in this connection are the computation of the exponential of the Q-matrix and the optimization of the likelihood function. The methods used for solving these problems are briefly described in section 6.

The ML-estimates are shown on fig. 9 and fig. 10. As expected a marked annual variation is observed. The estimates may be used to gain insight in the transitions between categories of cloud cover. For example it is seen that overcast sky (8) and clear sky (0) are the most stable states. The middle categories are seen to be most unstable.

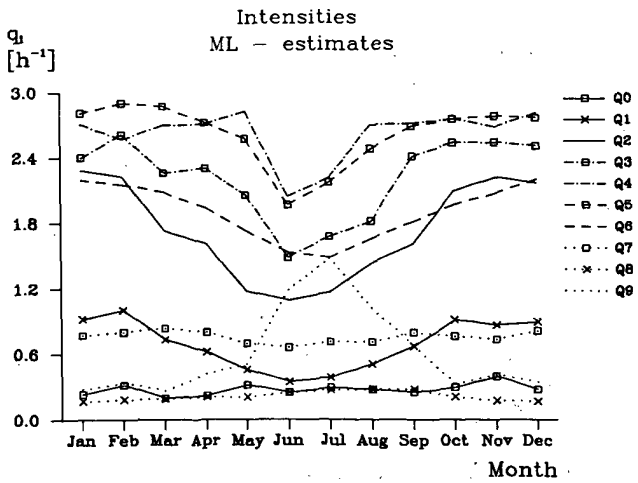


Fig. 9. ML-estimates of transition rates.

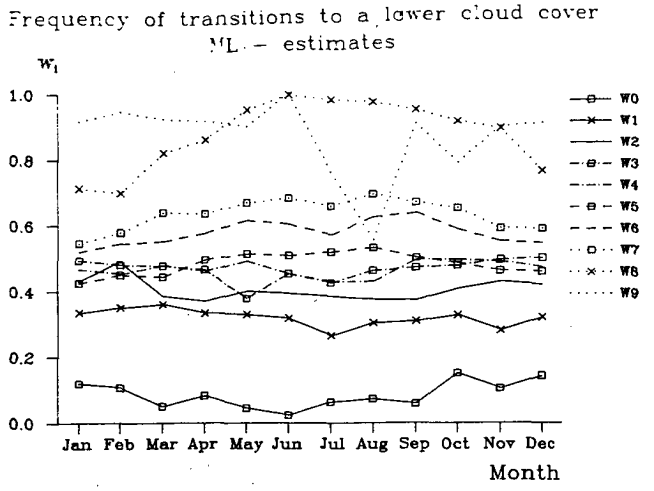


Fig. 10. ML-estimates of the frequency of transitions to lower cloud cover.

Comparing the ML-estimates of q_i with the simple estimates we note that the latter have been lifted and mostly for the greatest values of q_i - as expected. The ML-estimates of w_i do not deviate much from the simple estimates.

90% confidence bands for certain estimates are shown on fig. 11 and fig. 12. The uncertainty of the remaining transition rates is comparable with the uncertainty of $\hat{q}_1$. The coefficient of variation for $\hat{q}_0, \hat{q}_1, \ldots, \hat{q}_8$ are of the magnitude 3-8%. Due to the fact that fog is seldom present in the summer season the uncertainty of $\hat{q}_9$ is higher in these periods. The uncertainty of the other estimates of frequency of transitions are approximately as indicated by the uncertainty of $\hat{w}_1$.

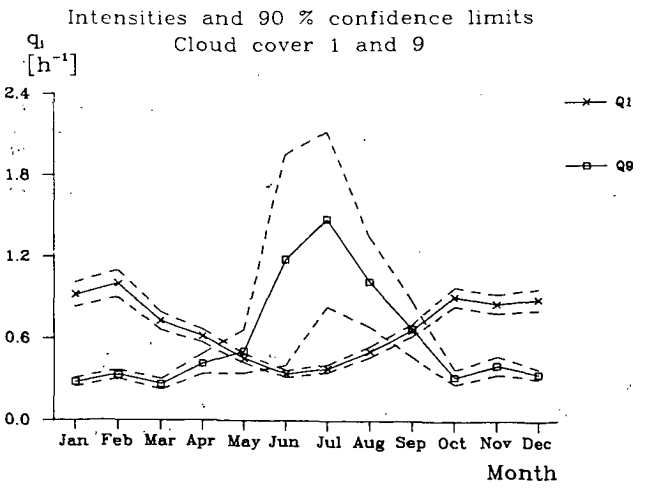


Fig. 11. 90% confidence band for $\hat{q}_1$ and $\hat{q}_9$.

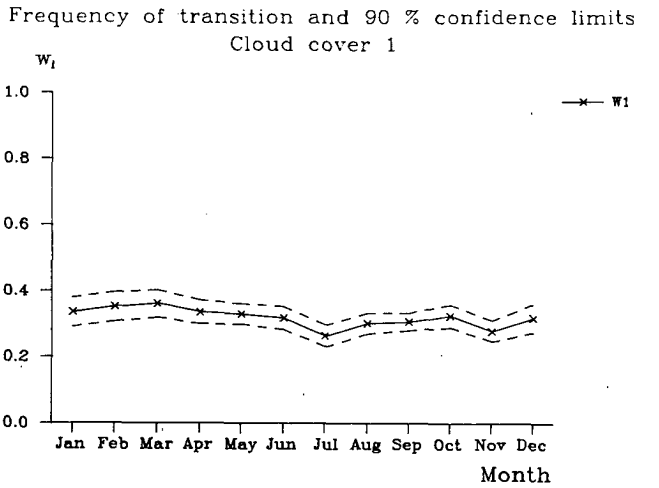


Fig. 12. 90% confidence band for $\hat{w}_1$.

Considering the uncertainties, it seems reasonable to smooth the annual variation for certain estimates.

6. NUMERICAL METHODS

We consider the problems of the optimization of the likelihood function and the calculation of the exponential of the Q-matrix.

For unconstrained optimization the IMSL-routine ZXMIN (IMSL (1980)) was used. The routine is based on the David-Fletcher-Powell variant of the quasi Newton method, as described by Luenberger (1973). To account for the bounded parameter space

$$q_i \in]0,\infty[$$

$$w_i \in [0,1]$$

quadratic penalty functions were introduced, as suggested by Luenberger.

The routine ZXMIN does only find a local maximum, but several facts contribute to our belief in that the local maximum corresponds to the global maximum. First of all we observed the expected connection between the simple estimates and the ML-estimates. Further it is noted that the same maximum was found independent of quite different starting values for the iterative search. Singer and Spilerman (1975, 1976) have discussed the uniqueness of the solution to $P = \exp(Q)$. It may be verified that if Q is a tridiagonal matrix, the solution is unique. In the case where Q has the structure given in section 5.2.1 there might be more than one solution when all w_i are zero. In the present situation the data indicate that, e.g., w_9 is close to 1 and hence that the solution to $P = \exp(Q)$ is unique. Finally, during the search the routine never found areas where the likelihood function was convex.

The iterative search for the maximum implies repeated evaluation of $\exp(Q)$. Moler and van Loan (1978) have given a general discussion of the

The iterative search for the maximum implies repeated evaluation of exp(Q). Moler and van Loan (1978) have given a general discussion of the problems associated with the computation of the exponential of a matrix. For the parameter values in the present study it turned out, that the best numerical stability was obtained, when the exponential was determined by means of an ordinary series expansion using 12 terms.

7. SUMMARY AND FUTURE WORK

Fifteen years of hourly observations of cloud cover are analysed in order to obtain models of cloud cover variations suitable for numerical simulations. A very clear annual variation of cloud cover is found. The obtained diurnal variation of the mean is smaller than the annual variation.

The cloud cover is measured by the integers 0 through 9. The discrete observations and an investigation of the hourly correlations make it attractive to consider Markov processes. As a first approximation homogeneous Markov processes are considered, which implies that the diurnal variation is neglected. In order to describe the annual variation a model is set up for every month.

Models in both discrete and continuous time are considered. Generally, in continuous time only jumps to neighbouring cloud cover categories are possible. Hence a model based on a continuous time Markov process requires less parameters than the corresponding discrete time model does. The estimation of the parameters of the different models is described. The matrix of transition rates belonging to the embedded Markov process in continuous time is estimated taking into account the physical conditions for transitions. The ML-estimate is found by numerical methods using an iterative search for the maximum of the likelihood function. Confidence bands are obtained as an indication of the possible estimation error under the model. Considering the bands it seems reasonable to smooth the annual variation for certain estimates.

A drawback of the homogeneous model considered is that the diurnal variation is neglected. The diurnal variation could be described by introducing inhomogeneous Markov processes. Thus the matrix of transition rates and the transition probabilities are depending on the time of the diurnal cycle. In this situation Kolmogorov's differential equation becomes

$$\frac{\partial P(t)}{\partial u} = P(t,u)Q(u)$$

where the structure of $Q(u)$ is as in the homogeneous case. A simple solution is obtained by assuming $Q(u)$ being constant in periods of length T. It is natural to choose T as the sampling period (T = 1 hour). Thus the number of required parameters in our case will be (20 * 24) for each month. The number of parameters in continuous time is again less than the number of parameters in discrete time, which is (90 * 24) for each month. An additional reduction of the numbers of required parameters is obtained by modelling the diurnal variation of $Q(u)$ by simple functions, for example by some terms of a Fourier expansion. Furthermore some of the parameters could be fixed after testing the influence of the diurnal variation. The details and results from the inhomogeneous case will hopefully be available soon.

8. REFERENCES

Anderson, T. W. and Goodman, L., 1957: Statistical Inference about Markov Chains. Ann. Math. Stat., 28, 89-109.

Box, G.E.P. and Jenkins, G. M., 1976: Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco.

Cox, D. R. and Miller, H. D., 1980: The Theory of Stochastic processes. Champman and Hall Ltd., New York.

IMSL Library, 1980, Houston, Texas.

Luenberger, D. G., 1973: Introduction to Linear and Nonlinear Programming. Addison-Wesley, London.

Moler, C. and van Loan, C.: Nineteen Dubious Ways to Compute the Exponential of a Matrix. Siam Review, 20, 801-836.

Singer, B. and Spilerman, S., 1975: Identifying Structural Parameters of Social Processes Using Fragmentary data. Bull. Int. Statist. Inst., 46, 681-697.

Singer, B. and Spilerman, S., 1976: Some Methodological Issues in the Analysis of Longitudinal Surveys. Ann. Econ. Soc. Meas., 5, 447-474.

World Meteorological Organization 1972, 1974: Manual on Codes. Vol. 1 and 2. WMO, Post Box No. 5, CH-1211, Geneva 20, Switzerland.

STATISTICAL APPROACH TO SOIL MOISTURE CONTENT BASED ON A MARKOV CHAIN MODEL FOR PRECIPITATION

Tibor Farago

Central Meteorological Institute
Budapest, Hungary

1. INTRODUCTION

The surface water balance has widely been investigated by climatologists as an essential constituent of the atmospheric water cycle. Its climatological characteristics, variability and trends are important environmental factors in agricultural and hydrological planning and management /WMO, 1979/. Certain average ratios within this balance are reliable indicators for climatic classification /Thornwaite, 1948/. Various atmospheric circulation and climate models comprise the boundary conditions of the surface water balance, as well /Sellers, 1976; Marchuk, 1979; Boer and McFarlane, 1979/. Additional problems come from the human activities resulting in widespread change of vegetation and land use in general that, in turn, modifies the surface water and heat balance of large areas /Budyko, 1977/.

Precipitation is the most intensively studied factor of the water balance. As one possible approach, Gabriel and Neumann/1962/ proposed a Markov chain model to describe the process of rainfall occurrences. This technique has proved advantageous in the time series analysis of other meteorological processes, too /Gorbacheva et al., 1977/. Katz /1974, 1977/ has derived a more advanced expression of the distribution statistics in a recurrence form. It has been showed this method is a straightforward tool for estimating the number of wet days, the length of dry and wet spells, maximum amounts of precipitation and the starting date of rainy periods /Wagner, 1968; Stern, 1980a, 1980b, 1981/.

The investigation of other elements of the surface water budget, especially that of the soil moisture content being eventually of direct concern in agriculture, is a more complicated problem. Climatonomical models of the surface water balance count on the soil moisture due its feedback effects on the evaporivity and the surface albedo affecting in turn convective instability, rain shower activity and radiation absorption, respectively /Lettau and Baradas, 1973/. Similar relations, but on an other time scale are taken into consideration in the global circulation modelling. The short-term effects of the soil humidity on the heat and moisture fluxes in the surface boundary layer are usually parameterized /Marchuk, 1979; GARP, 1972/.

Based on a Markov chain model for the precipitation /Katz, 1977/, it is now examined the possibility to deduce an extended probability model of the soil moisture.

As another more general aspect of the problem treated below is the derivation of a simple stochastic version of a "deterministic" equation describing a hydrometeorological process. An extended stochastic approach has widely been used in the frame of the statistical and moreover in the stochastic dynamic /Hasselmann, 1977/ modelling. Probability moments can easily been expressed for simplified differential equations /Epstein, 1969/. However, the direct analytical deduction of the time-dependent probability distributions of terms originally consisted in ordinary dynamical /differential and approximative difference/ equations is usually possible only at the cost of serious simplifications. For the surface water balance in this study, it means that only one of its constituents, the daily precipitation will be entered in its "own" chain-dependent stochastic form, while all other components will be taken into consideration as dependent /or parameterized/ variables. This simplification allows us to produce approximate time-dependent probability distributions of the soil moisture.

2. SIMPLIFIED APPROACH TO THE SURFACE WATER BALANCE

The water balance for the land surface is usually given by the equation

$$R = \frac{\partial T}{\partial t} + A + I \qquad /1/$$

/GARP, 1972; Lettau and Baradas, 1973; Marchuk, 1979/, where R denotes the precipitation, I, T and A stand for the runoff, the soil moisture and the actual evaporation, respectively. All components are assigned to a unit area. The runoff term includes both the surface and the underground runoff. The evaporation term, in the general case, should be replaced by the evapotranspiration. In most studies, the upper one metre of the underlying soil is involved in the measurements and computations of the productive soil moisture content T, for it is, in general, the actually active ground layer in plant production and in other processes. For certain periods /and areas/ a term related to the snow melting rate should be added /Boer and McFarlane, 1979/, but we are primarily concerned with periods with snow-free land surface.

In particular, soil moisture investigations are important in areas, where the actual evaporation is essentially less than the potential evaporation. Below a critical value $\hat{T}$, the soil humidity becomes fundamental in the determination of the actual evaporation. This dependence is close to linear in such a way that $A = P\,T/\hat{T}$, where P denotes the potential evaporation. To some extent, the runoff can also be expressed as a linear function of the soil moisture /Budyko, 1971/.

Consequently, under the conditions of insufficient humidification /GARP, 1972/ eq./1/ can be reduced to the form

$$R = \frac{\partial T}{\partial t} + \varepsilon T \,, \qquad /2/$$

where ε is a dimensionless coefficient of proportionality. /For the sake of simplicity, it is taken to be constant, nevertheless all further reasoning can simply be extended for a time-dependent coefficient $\varepsilon = \varepsilon(t)$./

Finally, for further computations we introduce the simple difference approximation to eq./2/

$$T_n = (1-\varepsilon)T_{n-1} + R_n \,, \qquad /3/$$

where n denotes the number of the day, R_n is the amount of precipitation on the nth day and T_n is the soil moisture content on the same day /or more formally, at the end of that day/.

3. MARKOV CHAIN MODEL FOR THE PRECIPITATION

After the real series of daily occurrence of the precipitation have successfully been approached by homogeneous Markov chains /Gabriel and Neumann, 1962; Wagner, 1968/, this model was thoroughly extended for the daily amounts of the precipitation /Katz, 1977/. The fitness of this model as compared with the hypothesis of the stochastic independence of the daily precipitation amounts and with the unnecessity of involving more than first order chain-dependence, have been investigated and proved /Feyerheim and Dean Bark, 1965, 1967; Katz, 1977/. It was showed that for periods essentially shorter than one month, it is reasonable to introduce this interdiurnal dependence in the computation of the various statistics, in particular, of the total amount of the precipitation in n days.

Under certain conditions, Katz/1977/ has derived a recurrence formula for the conditional probability distribution of this amount

$$H_n(x|i) = P_r\{\sum_{k=1}^{n} R_k \le x \,|\, J_0 = i\} = P_{i0}H_{n-1}(x|0) + P_{i1}F_i(x) * H_{n-1}(x|1) \,, \qquad /4/$$

$$H_0(x|i) \equiv 1 \,, \qquad x \ge 0 \,, \quad i = 0,1 \,,$$

where J_n is the indicator function of R_n, i.e.

$$J_n = \begin{cases} 1 & \text{if precipitation occurs on } n \text{ th day} \\ 0 & \text{otherwise} \end{cases}$$

and P_{ij} are the transition probabilities for the Markov chain $\{J_n,\ n = 0,1,\ldots\}$. The asterisk in eq./4/ denotes the convolution operator. F_i, $i = 0,1$ correspond to the conditional distribution functions of any stochastic variable R_n

$$F_i(x) = P_r\{R_n \le x \,|\, J_{n-1} = i,\ J_n = 1\} \,,$$

since R_n, $n = 1,2,\ldots$ are assumed to be identically distributed given J_{n-1} and $J_n = 1$.

4. PROBABILITY DISTRIBUTION OF SOIL MOISTURE CONTENT

The first order Markov property of the precipitation sequence R_n can be expressed in an appropriate way

$$\Pr\{R_n \le x \mid T_0, J_0, R_1, J_1, \ldots, R_{n-1}, J_{n-1}=i\} = \Pr\{R_n \le x \mid T_0, J_{n-1}=i\}, \quad /5/$$

$i=0,1$. This condition is just a slightly generalized form of a similar assumption introduced by Katz/1977/. Given T_0 eq./5/ states that the amount of the precipitation on a given day depends only on the occurrence of precipitation on the previous day. The application of this condition immediately yields

$$\Pr\{S'_n \le x \mid T_0, J_1=i\} = \Pr\{S_n \le x \mid T_0, J_0=i\}, \quad /6/$$

where

$$T_n = (1-\varepsilon)^n T_0 + S_n, \quad T_{n+1} = (1-\varepsilon)^{n+1} T_0 + (1-\varepsilon)^n R_1 + S'_n \quad /7/$$
$$S_n = \sum_{k=1}^{n} (1-\varepsilon)^{n-k} R_k, \quad S'_n = \sum_{k=2}^{n+1} (1-\varepsilon)^{n+1-k} R_k,$$

that is, both expressions $T_n - (1-\varepsilon)^n T_0$ and $T_{n+1} - (1-\varepsilon)^{n+1} T_0 - (1-\varepsilon)^n R_1$ are equally distributed.

Let Φ_n denote the conditional probability distribution of soil moisture content on the nth day given T_0

$$\Phi_n(x) = \Pr\{T_n \le x \mid T_0\} \quad \text{for} \quad x \ge (1-\varepsilon)^n T_0.$$

The lower boundary is reached when there is no precipitation in n days, i.e. $J_1 = J_2 = \ldots = J_n = 0$. By conditioning on J_0 we obtain

$$\Phi_n(x) = \lambda_0 \Phi_n(x \mid 0) + \lambda_1 \Phi_n(x \mid 1),$$
$$\Phi_n(x \mid i) = \Pr\{T_n \le x \mid T_0, J_0=i\}, \quad \lambda_i = \Pr\{J_0=i\}.$$

A further decomposition with respect to J_1 results in

$$\Phi_n(x \mid i) = P_{i0} \Pr\{T_n \le x \mid T_0, J_0=i, J_1=0\} + P_{i1} \Pr\{T_n \le x \mid T_0, J_0=i, J_1=1\} \quad /8/$$

for $x \ge (1-\varepsilon)^n T_0$. The substitution of /7/ and the use of /6/ in /8/ gives

$$\Phi_n(x \mid i) = P_{i0} \Phi_{n-1}(x + \varepsilon\langle 1-\varepsilon\rangle^{n-1} T_0 \mid 0) + P_{i1} F_i\left(\frac{x}{(1-\varepsilon)^{n-1}}\right) * \Phi_{n-1}(x + \varepsilon\langle 1-\varepsilon\rangle^{n-1} T_0 \mid 1) \quad /9/$$

for $x \ge (1-\varepsilon)^n T_0$. In this deduction we have taken the advantage of the conditional independence of R_1 and R_k given J_1, that is a direct consequence of eq./5/. Especially, the probability of the dry spell through n days is

$$\Phi_n(\langle 1-\varepsilon\rangle^n T_0 \mid i) = \Pr\{R_1 = R_2 = \ldots = R_n = 0 \mid T_0, J_0=i\} = P_{i0} P_{00}^{n-1}. \quad /10/$$

+ Probabilities λ_i, P_{ij} are assumed to be independent of T_0.

The initial phase of the recurrence relation should be expressed separately, so that

$$\Phi_1(x \mid i) = P_{i0} + P_{i1} F_i(x - \langle 1-\varepsilon\rangle T_0), \quad x \ge (1-\varepsilon) T_0$$
$$\Phi_1(\langle 1-\varepsilon\rangle T_0 \mid i) = P_{i0}. \quad /11/$$

5. RECURRENCE FORMULA FOR EXPECTED VALUES

One of the simplest ways to characterize the generalized stochastic process T_n is a proper estimation of its conditional expected values $E(T_n \mid T_0)$. Like the derivation of formulae /9/-/11/ for the entire distribution, a similar recurrence equation can be given for the expected values.

The Markov property in this case results in the relation like eq./6/

$$E(S'_n \mid T_0, J_1=j) = E(S_n \mid T_0, J_0=j). \quad /12/$$

In order to produce the desired recurrence formula we express the expected values of soil moisture content for the indices n and $n+1$ as follows

$$E(T_n \mid T_0, J_0=i) = (1-\varepsilon)^n T_0 + E(S_n \mid T_0, J_0=i),$$
$$E(T_{n+1} \mid T_0, J_0=i) = P_{i0}\left[E(S'_n \mid T_0, J_1=0) + (1-\varepsilon)^{n+1} T_0\right] + P_{i1}\left[E(S'_n \mid T_0, J_1=1) + (1-\varepsilon)^n E(R_1 \mid T_0, J_0=i, J_1=1) + (1-\varepsilon)^{n+1} T_0\right]. \quad /13/$$

Therefore, from /12/ and /13/

$$E(T_{n+1} \mid T_0, J_0=i) = P_{i0} E(T_n \mid T_0, J_0=0) + P_{i1} E(T_n \mid T_0, J_0=1) - (1-\varepsilon)^n \left[\varepsilon T_0 - P_{i1} E(R_1 \mid T_0, J_0=i, J_1=1)\right], \quad /14/$$

where the last expression can be rewritten in another form, too

$$E(R_1 \mid T_0, J_0=i) = P_{i1} E(R_1 \mid T_0, J_0=i, J_1=1). \quad /15/$$

To initialize eq./14/ we set

$$E(T_1 \mid T_0) = \lambda_0 E(T_1 \mid T_0, J_0=0) + \lambda_1 E(T_1 \mid T_0, J_0=1),$$
$$E(T_1 \mid T_0, J_0=i) = (1-\varepsilon) T_0 + E(R_1 \mid T_0, J_0=i). \quad /16/$$

6. NUMERICAL EXAMPLE

There are certain limitations in the further analytical solution of the recurrence formulae /9/-/11/. In their study, Todorovic and Woolhiser /1975/ have approached the daily precipitation amounts by means of exponential distribution. It is well-known that the sum of independent

exponentially distributed /with the same parameter/ variables obeys gamma-distribution. Even if our variables were stochastically independent with identical exponential probability density function $\beta^{-1}\exp(-x/\beta)$, the sum $\sum(1-\varepsilon)^k R_k$, $k=1,2,\dots,n$ would have a complicated probability behaviour according to the density function of the generalized Erlang's distribution

$$g_n(x)=(-1)^{n-1}\frac{1}{\beta_1\beta_2\dots\beta_n}\cdot$$

$$\cdot\sum_{k=1}^{n}\left[\exp(-x/\beta_k)/\prod_{i\neq k}(1/\beta_k-1/\beta_i)\right],\quad x>0$$

/Rényi, 1966/ with parameters $\beta_k=\beta(1-\varepsilon)^{n-k}$. Further investigations showed the gamma-distribution fits much better to the daily precipitation data. This distribution is frequently selected as the theoretical distribution of monthly rainfall sum, too /Panofsky and Brier, 1958; Bagrov, 1965; Szigyártó, 1972; Gray, 1976/. However, the Katz's assumption /1977/ cannot help to avoid the difficulties in our case, by taking F_0 and F_1 to be gamma-distribution functions with a common scale parameter β. Although the population of such distributions is closed with respect to the convolution operator, i.e. the sum of these independent variables has a similar type distribution with the same scale parameter, just again the coefficients $(1-\varepsilon)^{n-k}$ distort this value. Consequently, we are to solve the recurrence equations /9/-/11/ by finite difference approximation.

Note that the model under discussion reduces to Katz's/1977/ precipitation model by taking $T_0=0$ and $\varepsilon=0$.

The daily precipitation amounts for May at Budapest /Institute for Atmospheric Physics/ have been used for the numerical illustration. The sequence of the consequtive amounts is characterized by a homogeneous Markov chain with transition probabilities $P_{00}=0.72$, $P_{10}=0.40$, and the parameters of the conditional gamma-distributions F_0 and F_1, i.e.

$$F_i(x)=\int_0^x f_i(t)\,d(t)\quad,\qquad /17/$$

$$f_i(t)=\frac{1}{\beta_i^{\alpha_i}\Gamma(\alpha_i)}\,t^{\alpha_i-1}\exp(-t/\beta_i)\,.$$

The latter parameters can be estimated by the maximum likelihood method /Panofsky and Brier, 1958/

$$\alpha_i=(1+\sqrt{1+4/3\,A_i})/4A_i$$

$$\beta_i=r_i/\alpha_i\qquad i=0,1,$$

where r_i denotes the conditional sample mean and

$$A_i=\ln r_i-\frac{1}{N}\sum_{j=1}^{N_i}\ln R_{ij}\,.$$

The subsequence $\{R_{ij}, j=1,2,\dots,N_i\}$ consists of all non-zero / $R_{ij}\geq 0.01$ cm/ elements of the sample series $\{R_j\}$ with wet preceding day if $i=1$ and with dry preceding day if $i=0$, respectively. With this approach, we have

$$\alpha_0=0.677\,,\qquad \beta_0=0.525\ \text{cm}$$

$$\alpha_1=0.769\,,\qquad \beta_1=0.628\ \text{cm}.$$

The density and distribution functions are showed in Fig.1.

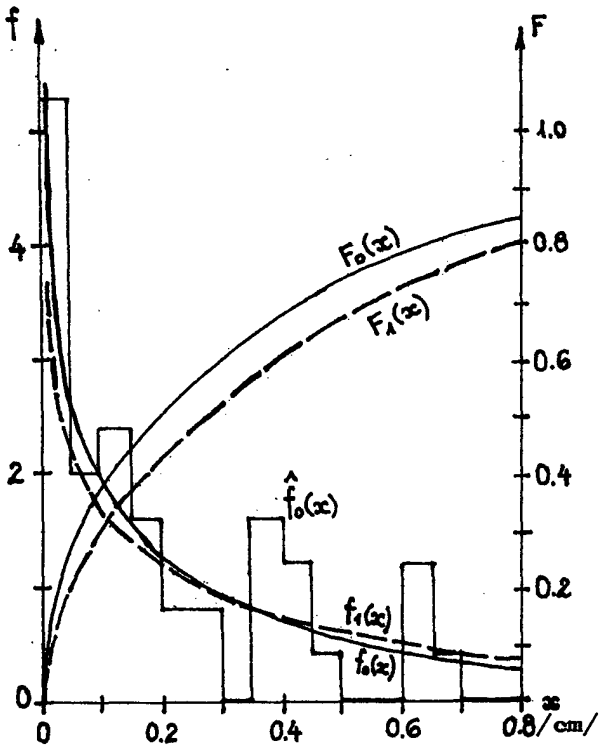


Fig.1. Conditional density and distribution functions for the Budapest daily precipitation amounts given dry preceding day and wet preceding day /dotted curves/, respectively. For the former, the sample histogram is also showed

The a priori probabilities π_i of the events $J_0=i$ have been chosen equal to the asymptotic stationary probabilities

$$\pi_i = \lim_{n\to\infty} P_r\{J_n = i\} = \begin{cases} P_{10}/(P_{01}+P_{10}) , & i=0 \\ P_{01}/(P_{01}+P_{10}) , & i=1 . \end{cases} /18/$$

In this example, $\pi_0 = 0.620$, $\pi_1 = 0.380$. For the initial value of the productive soil moisture, $T_0 = 8$ cm has been taken. It corresponds to the usual requirements /GARP, 1972/ of insufficient humidification with a critical productive soil

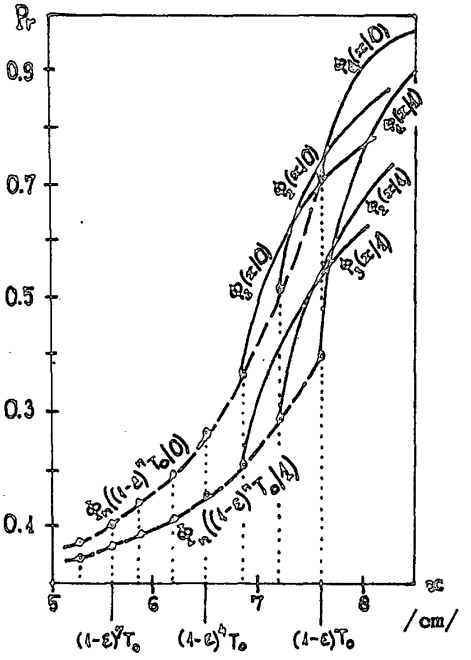


Fig. 2. Conditional probability distributions of soil moisture $\Phi_n(x|i)$, $i=0,1$ for the days $n = 1, 2, 3$ given $T_0 = 8$ cm and $\varepsilon = 0.05$. Dotted lines are plotted against the minima of variables T_n on the consecutive days, i.e. those show the probability rate for periods with no precipitation

moisture value of 10–20 cm in the upper 1 m layer of the soil. An empirical value of 0.05 has been selected for ε. Fig. 2 illustrates the computed conditional distributions $\Phi_n(x|i)$ for $n=1,2,3$. The curves of the probabilities for dry spells bound from below the plotted distributions. These values show the probability rate at which the soil dries out and those can easily be estimated from eq. /10/.

The recurrence formulae /14/–/15/ for the expected /mean/ values can essentially be simplified using the real data as follows

$$E(R_1|T_0, J_0=i) = P_{i1}\alpha_i\beta_i = \begin{cases} 0.10 , & i=0 \\ 0.29 , & i=1 \end{cases}$$

$$E(T_{n+1}|T_0, J_0=i) = P_{i0}E(T_n|T_0, J_0=0) + P_{i1}E(T_n|T_0, J_0=1) - C_{i,n}$$

where $E(T_1|T_0, J_0=0) = 7.70$, $E(T_1|T_0, J_0=1) = 7.89$, $C_{0,n} = 0.30\cdot 0.95^n$, $C_{1,n} = 0.11\cdot 0.95^n$. Long run calculations show that $E(T_n|T_0)$ tends to the value of $\tau = 3.60$ that is just the asymptotic solution to the equation

$$E(T_{n+1}|T_0) = (1-\varepsilon)E(T_n|T_0) + E(R_n|T_0)$$

i.e., with the appropriate substitutions /due to the choice $\lambda_i = \pi_i$/

$$\tau = (1-\varepsilon)\tau + \pi_0 E(R_1|T_0, J_0=0) + \pi_1 E(R_1|T_0, J_0=1).$$

For various values of T_0 and ε, the march of the expected values is showed in Fig. 3. The changes of the soil moisture content for the periods with no precipitation also indicated. Note, that setting a low value of ε, the soil moisture content will increase on average up to the corresponding theoretical limit τ or the field capacity if the latter is the smaller. /Moreover, it has been assumed in this study that T_n does not exceed, in general, a threshold $\hat{T}$ ./ In the special case under consideration with $\lambda_i = \pi_i$ $i = 0, 1$ the expected values can be treated more simply by noticing $E(R_n|T_0) = E(R_1|T_0)$, $n = 2, 3, \ldots$. However, either the conditional values given J_0 or any of these expected values with other a priori probabilities λ_i can be computed by means of the more sophisticated equations /14/–/16/.

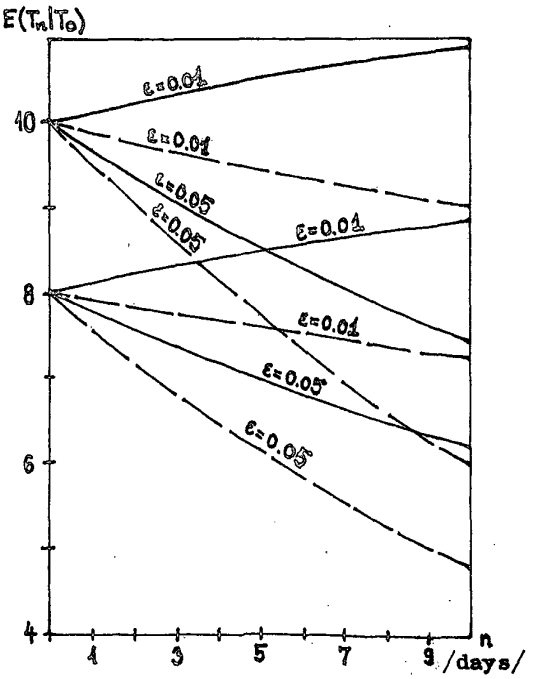


Fig. 3. Expected values of soil moisture given T_0 for various values of ε. Besides, the minima of T_n are showed for the same values of T_0 and ε /dotted curves/

7. CONCLUSIONS

The simplified surface water balance equation has been generalized so that the probability distribution of its soil moisture component be expressed in a recurrence form. A similar equation has been deduced for the first moments, as well. The consideration of the stochastic dependence of the consecutive daily precipitation amounts is usually unavoidable for periods essentially shorter than one month. Gamma distributions are used in fitting to the stochastic variables of daily precipitation amounts. The model could have been solved only in an approximate numerical way due its complexity. Nevertheless, either the distribution function, or in a simpler case, the mean value of the soil moisture can effectively be estimated given all parameters of the Markov chain of daily precipitation, an initial value of the soil moisture and the coefficient of parameterization of other terms in the water balance equation. Such a parameterization is acceptable under the conditions of insufficient humidification.

It is clear that further improvements are necessary for a better validation of this model. Especially, there is a need to retain a more realistic version of the water balance equation for the generalization.

8. REFERENCES

Bagrov, N.A., 1965: On distribution of monthly precipitation amounts. CIP, No.139, Moscow, 3-20 /in Russian/

Boer, G.J., N.A. McFarlane, 1979: The AES atmospheric general circulation model. GARP P.S. No.22, 409-460.

Budyko, M.I., 1971: Climate and life. Gidrometizdat, Leningrad /in Russian/

Budyko, M.I., 1977: Global ecology. Misl, Moscow /in Russian/

Epstein, E.S., 1969: Stochastic dynamic prediction. Tellus, 21, 739-759.

Feyerherm, A.M., L. Dean Bark, 1965: Statistical methods for persistent precipitation patterns. J. Appl. Meteor., 4, 320-328.

Feyerherm, A.M., L. Dean Bark, 1967: Goodness of fit of a Markov chain model for sequences of wet and dry days. J. Appl. Meteor., 6, 770-773.

Gabriel, K.R., J. Neumann, 1962: A Markov chain model for daily rainfall occurrence at Tel Aviv. Quart. J. Roy. Meteor. Soc., 88, 90-95.

GARP. 1972: Parameterization of sub-grid scale processes. P.S. No.8, WMO

Gorbacheva, N.A. et al., 1977: Statistical analysis of march of types of monthly mean temperature anomaly fields. VNIIGMI No.53, Gidrometizdat, Obninsk, No.53, 13-21 /in Russian/

Hasselmann, K., 1977: The dynamical coupling between the atmosphere and the ocean. Rep. Marine Sci. Affairs, No.11, WMO No.472, 31-44.

Katz, R.W., 1974: Computing probabilities associated with the Markov chain model for precipitation. J. Appl. Meteor., 13, 953-954.

Katz, R.W., 1977: Precipitation as a chain-dependent process. J. Appl. Meteor., 16, 671-676.

Lettau, H.H., M.W. Baradas, 1973: Evapotranspiration climatonomy II. Mon. Wea. Rev., 101, 636-649.

Marchuk, G.I. et al., 1979: Numerical simulation of the global circulation of the atmosphere. GARP P.S. No.22., 318-370.

Panofsky, H.A., G.W. Brier, 1958: Some applications of statistics to meteorology. Univ. Press, Pennsylvania.

Rényi A., 1966: Probability theory. Tankönyvkiadó, Budapest /in Hungarian/

Sellers, W.P., 1976: A two-dimensional global climatic model. Mon. Wea. Rev., 104, 233-248.

Stern, R.D., 1980a: Analysis of daily rainfall at Samaru. Archiv für Meteor., Ser. B, 28, 123-136.

Stern, R.D., 1980b: The calculation of probability distributions for models of daily precipitation. Archiv für Meteor., Ser. B, 28, 137-148.

Stern, R.D., 1981: Computing a probability distribution for the start of the rains from a Markov chain model for precipitation. J. Appl. Meteor., 21, 420-423.

Szigyártó Z., 1972: Depicting the variability of precipitation amounts for monthly and longer periods. VITUKI-TKE No.36, Budapest /in Hungarian/

Thornwaite, C.W., 1948: An approach toward a rational classification of climate. Geographical Rev., XXXVIII, 55-94.

Todorovic, P.D., A. Woolhiser, 1975: A stochastic model of n-day precipitation. J. Appl. Meteor., 14, 17-24.

Wagner M., 1968: Some statistical characteristics of days without and with small amount of precipitation. Időjárás, Budapest, 72, 30-33. /in Hungarian/

WMO, 1979: World Climate Conference. Ext. sum. of papers. Geneva

Gray, B.M., 1976: Medium term fluctuations of rainfall in Southeastern England. Quart. J. Roy. Met. Soc., 102, 627-638.

APPLICATION OF A REGULAR MARKOV CHAIN TO COMPUTATIONS OF RUNS OF HISTORICAL FLOOD AND DROUGHT YEARS

C. S. YAO

Department of Meteorology
Nanjing University, China

1. INTRODUCTION

In the theory of a Markov chain, it has first been treated as a series of transitions between states without after-effect, that is to say the probability law of the future development of the process depends only on the present state and not on how the process arrived in this present state. This simple Markov model introduced by different authers in studying sequences of weather events is only the two-state Markov chain. Gabrial and Neumann(1957) first suggested a simple Markov chain model to study the Tel Aviv daily rainfall occurrences. Gabrial (1959) in terms of the simple Markov model derived the distribution of the numbers of wet days in a sequence of dependent observations. However, a question whether the sequences of wet and dry days may be described by the simple Markov chain model or not was controversial in the past times and this question concerning Tel Aviv daily rainfall data is also debatable in the recent times. A brief recommendation of this problem has been given by Yao(1983).

Now, we shall deal with the problem about how to apply the theory of regular Markov chains to computations of runs of the historical flood and drought records. The methods here suggested can also be applied to the historical records of coldness and warmth. Various definitions of runs were used in the literature. As a classical term, a run is defined as a sequence of observations of the same kind preceded and succeeded by one or more observations of a different kind (cf. Feller, 1950; Yevjevich, 1972). However, the theory of runs for the Markov chain is relatively undeveloped, except for some cases of the simple Markov chain with two transitional states (Yevjevich, 1972). Since the simple Markov chain model introduced by different authors in studying runs of precipitation events is only the two-state Markov chain, it might be unsatisfactory to describe the sequences of precipitation events by such a model because of persistence existing in these sequences of data. Billard and Meshkani (1980) subclassified the wet days into (m-1) classes designating the degree of wetness as measured by the amounts of rainfall actually received. In the case of this Markov chain with m states, they ascertained the transition probabilities of moving from one weather classification type to another in any given number of days and the long run probabilities of certain weather types.

For the purpose of improving the Markov property of minimal feedback in studying the probabilities of changes of wet and dry days, Yao(1966) treated stochastic changes of weather as trials and took the effect of any past trials into consideration in the states themselves, while the states were classified. By this technique, a regular Markov chain with n-step transition may be formed without considering the problem of how to determine the order of the Markov chain. This kind of Markov chains under consideration of the after-effect may satisfactorily be used to describe the random changes of precipitation events. According to the above idea, various states are then similarly classified by run-lengths of 0, 1, 2, ... years of flood(drought) conditions in the historical records, therefore the theory of runs in historical flood and drought records can be calculated by means of the regular Markov chain.

2. REGULAR MARKOV CHAIN

Let S_j (j=0, 1, 2, ..., m) represent state with the index j. Let $p_{jk}^{(n)}$ represent the n-step transition probability, that is, the probability of finding the process to be in the state S_k at the current time, given that it has been in the state S_j in the other time, with n time intervals from S_j to S_k. The n-step transition probabilities will be arranged in a n-step transition matrix

$$P^n = \begin{pmatrix} p_{00}^{(n)} & p_{01}^{(n)} & \cdots & p_{0m}^{(n)} \\ p_{10}^{(n)} & p_{11}^{(n)} & \cdots & p_{1m}^{(n)} \\ \vdots & \vdots & \cdots & \vdots \\ p_{m0}^{(n)} & p_{m1}^{(n)} & \cdots & p_{mm}^{(n)} \end{pmatrix}, \qquad (1)$$

where the first subscript stands for row, and the second for column. We know that

$$p_{jk}^{(n)} = \sum_i p_{ji}\, p_{ik}^{(n-1)} ; \qquad (2)$$

$P^1=P$. The Matrix P together with the initial distribution completely defines the regular Markov chain with states S_o, S_1, S_2, ... S_m.

A regular Markov chain is one which has a single ergodic set, that is, it is possible to go from every state to every other one. If P is a regular transition matrix, then the power P^n tends as $n\to\infty$ to a probability matrix A, each row of which is the same probability vecter $\alpha=(a_o, a_1, \ldots, a_m)$, where $a_k>0$, $k=0, 1, 2, \ldots, m$. In other words, the matrix P gives the following property

$$\lim_{n\to\infty} p_{jk}^{(n)} = a_k , \qquad (3)$$

where j and k denote the row subscript and the column subscript respectively for any state of the chain, and a_k will be independent of j. The matrix $A=\{a_k\}$ and the vector α are called the limiting matrix and the limiting vector respectively.

It can be show that

$$a_k = \sum_{j\in S} a_j p_{jk} , \quad k \in S , \qquad (4)$$

where S represents the state space of the regular Markov chain. Since the chain (1) is a finite Markov chain, $\{a_o, a_1, a_2, \ldots, a_m\}$ is a probability distribution, namely

$$\sum_{k=o}^{m} a_k = 1 \qquad (5)$$

Let t_{jk} be a function whose value is the number of steps before entering S_k for the first time from S_j. For the regular Markov chain, the mean first passage time from S_j to S_k is $m_{jk} = E(t_{jk})$, if $j \neq k$. It is called the mean recurrence time of the state S_k, if $j=k$. It can be proved(Kemeny and Snell, 1960) that the mean first passage matrix with entries m_{jk} is given by

$$M = (I - Z + EZ_{dg}) D, \qquad (6)$$

where I is a unit matrix, E a square matrix with all entries 1, D a diagonal matrix with diagonal elements $d_{kk}=\frac{1}{a_k}$, and

$$Z = (I - P + A)^{-1} \qquad (7)$$

is called the fundamental matrix for the regular Markov chain determined by P.

Let $V=\{\sigma_{jk}^2\}$ be the variance of the first passage time. Then, it can be proved(Kemeny and Snell, 1960) that

$$V = M(2Z_{dg} D - I)+2(ZM - E(ZM)_{dg}) - M_{sq} , \qquad (8)$$

where the Matrix M_{sq} is formed from M by squaring each entry.

It is known that the mean and variance are the two most important parameters to define the frequency distribution. Therefore, the mean and variance of the first passage time from S_j to S_k can enable us to study the behaviour of the regular Markov chain.

Besides, we can compute the probability distribution of the occupation time. Let S_k be a recurrent state with the finite mean recurrence time m_{kk} and finite variance σ_{kk}^2. Let $N_k(n)$ be occupation times of the state S_k, namely the number of times that the chain is in state S_k in the first n transitions. It may be shown(Parzen, 1962) that as n tends to infinity, the occupation times $N_k(n)$ have a distribution approaching the normal distribution with mean

$$E[N_k(n)] = \frac{n}{m_{kk}} \qquad (9)$$

and variance

$$Var\ N_k(n) = \frac{n\sigma_{kk}^2}{(m_{kk})^3} \qquad (10)$$

Thus, for any real number X,

$$\lim Pr\left[\frac{N_k(n) - \frac{n}{m_{kk}}}{\sqrt{n\sigma_{kk}^2/(m_{kk})^3}} \leq X\right] = \Phi(X), \qquad (11)$$

where $\Phi(X)$ is the standard normal distribution function. This is the Central Limit Theorem in the theory of Markov chains.

The probability that the variable X is smaller than or equal to the p-fractile X_p is equal to the cumulative probability $\Phi(X_p)$. The fractiles, u_p, of the standardized distribution of occupation times may be given by

$$Pr(u \leq u_p) = \Phi(u_p), \qquad (12)$$

where

$$u = \frac{N_k(n) - \frac{n}{m_{kk}}}{\sqrt{n\sigma_{kk}^2/m_{kk}^3}} \qquad (13)$$

and the fractiles of occupation times, $N_k(n)$, are equal to

$$\frac{n}{m_{kk}} + u_p\sqrt{n\sigma_{kk}^2/m_{kk}^3} \qquad (14)$$

for large samples.

3. CALCULATION METHODS OF RUNS OF FLOOD/ DROUGHT YEARS FOR REGULAR MARKOV CHAINS.

Let state S_o designate the non-flood year, state S_1 the run of a flood of one-year length, state S_2 run of floods of two-year length, and in general, state S_ν run of floods of ν-year length. Let the transition from S_o to S_1 and that from S_o to S_o have transition probabilities p_o and q_o respectively, the transition from S_1 to S_2 and that from S_1 to S_o have the transition probabilities p_1 and q_1 respectively, and in general the transition from S_ν to $S_{\nu+1}$ and that from S_ν to S_o have transition probabilities p_ν and q_ν respectively. Now, the transition matrix for the regular Markov chain known as the regular transition matrix for flood years is expressed in the form:

$$P = \begin{pmatrix} q_o & p_o & o & o & o & \cdots & o \\ q_1 & o & p_1 & o & o & \cdots & o \\ q_2 & o & o & p_2 & o & \cdots & o \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ q_\nu & o & o & o & o & \cdots & p_\nu \\ 1 & o & o & o & o & \cdots & o \end{pmatrix} \qquad (15)$$

Let F denote the flood year, and N the non-flood year. Let f_ν be the conditional probability of the non-flood year occurring

for the first time at the νth trial after an initial trial (non-flood year), i.e.,

$f_1=\frac{Pr(NN)}{P(N)}$, $f_2=\frac{Pr(NFN)}{P(N)}$, $f_3=\frac{Pr(NFFN)}{P(N)}$, . . . ,

$f_1+f_2+\cdots=1$. It can thus be seen that

$$p_0=Pr(F|N)=\frac{Pr(N)-Pr(NN)}{P(N)}=1-f_1\ ,$$

$$p_1=Pr(F|NF)=\frac{Pr(NF)-Pr(NFN)}{Pr(N)-Pr(NN)}=\frac{1-(f_1+f_2)}{1-f_1}$$

$$p_2=Pr(F|NFF)=\frac{Pr(NFF)-Pr(NFFN)}{Pr(N)-Pr(NN)-Pr(NFN)}=\frac{1-(f_1+f_2+f_3)}{1-(f_1+f_2)}$$

. . . . . .

In general,

$$p_k=\frac{1-(f_1+f_2+\cdots+f_{k+1})}{1-(f_1+f_2+\cdots+f_k)} \qquad (16)$$

and

$$q_k=1-p_k=\frac{f_k+1}{1-(f_1+f_2+\cdots+f_k)} \qquad (17)$$

where $k=0, 1, 2, \ldots$; $f_0=0$. Though the same relations have been found from the consideration of recurrent events(Feller, 1957), yet for Markov chain model the obove results are convenient in practical calculations of runs of flood years from the historical records. Similar calculations of runs of drought years by the Markov chain model may be performed, if the non-drought year is assigned to be the initial trial.

Fom (4), (5) and (15), we get the following set of equations

$$\begin{aligned} a_0+a_1+a_2+\cdots+a_\nu+a_{\nu+1}&=1\\ p_0a_0-a_1&=0\\ p_1a_1-a_2&=0\\ &\cdots\cdots\\ p_\nu a_\nu-a_{\nu+1}&=0 \end{aligned} \qquad (18)$$

The solutions of these equations are given by

$$\begin{aligned} a_0&=\frac{1}{1+p_0+p_0p_1+p_0p_1p_2+\cdots p_0p_1p_2\cdots p_\nu}\\ a_1&=p_0a_0\\ a_2&=p_1a_1\\ &\cdots\cdots\\ a_{\nu+1}&=p_\nu a_\nu \end{aligned} \qquad (19)$$

Hence, the required limit matrix A may be arranged and runs of flood(drought) years may be computed by means of (6)-(19).

In order to give computation example, we would recite 500-year historical graded records of floods and droughts in Beijing(1470-1974) which have been given out in another paper (Yao, 1982). These historical records of floods and droughts were graded into five different gradations, 1, 2, 3, 4 and 5 representing flood, subflood, normal, subdrought and drought respectively. On account of applying the theory of the Markov chain model to the analysis of runs of flood/drought years, the 5-grade series should be converted into a dichotomous series{0, 1}, in which "1" denotes both flood (drought) and subflood (subdrought), now called by flood (drought) year, and "0" the otherwise cases, now called by non-flood (non-drought) year.

4. STATISTICAL CHARACTERISTICS OF RUNS OF FLOOD/DROUGHT YEARS STUDIED BY A REGULAR MARKOV CHAIN MODEL FOR HISTORICAL RECORDS OF FLOOD AND DROUGHT IN BEIJING

Using the historical flood/drought records of the latest 505 years in Beijing (1470-1974) in the dichotomous variate form, the regular transition matrices for flood years and drought years respectively were obtained as follows:

Flood years: $P_=$

$$\begin{pmatrix} 0.74 & 0.26 & 0 & 0 & 0 & 0 & 0 & 0\\ 0.61 & 0 & 0.39 & 0 & 0 & 0 & 0 & 0\\ 0.64 & 0 & 0 & 0.36 & 0 & 0 & 0 & 0\\ 0.69 & 0 & 0 & 0 & 0.31 & 0 & 0 & 0\\ 0.50 & 0 & 0 & 0 & 0 & 0.50 & 0 & 0\\ 0.50 & 0 & 0 & 0 & 0 & 0 & 0.50 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1\\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Drought years: $P_=$

$$\begin{pmatrix} 0.67 & 0.33 & 0 & 0 & 0 & 0 & 0 & 0\\ 0.62 & 0 & 0.38 & 0 & 0 & 0 & 0 & 0\\ 0.71 & 0 & 0 & 0.29 & 0 & 0 & 0 & 0\\ 0.42 & 0 & 0 & 0 & 0.58 & 0 & 0 & 0\\ 0.29 & 0 & 0 & 0 & 0 & 0.71 & 0 & 0\\ 0.20 & 0 & 0 & 0 & 0 & 0 & 0.80 & 0\\ 0.25 & 0 & 0 & 0 & 0 & 0 & 0 & 0.75\\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The transition probability that $S_0\to S_1$ for drought year is larger than that for flood year. However, the probabilities that $S_1\to S_2$ and $S_2\to S_3$ for flood years are larger than that for drought years, though probabilities of that $S_3\to S_4$, $S_4\to S_5$, and $S_5\to S_6$ are larger for drought years than that for flood years. Recognizing that long run-lengths for flood/drought years are of the rare cases, Yao(1982) arrived at the conclusion that the flood-dominant climate was slightly prevaling in Beijing in the lastest 500 years.

The two-step transition matrices for flood/drought years computed by matrix multiplication were as follws:

Flood years: $P^2=$

0.71	0.19	0.10	0	0	0	0	0
0.70	0.16	0	0.14	0	0	0	0
0.72	0.17	0	0	0.11	0	0	0
0.67	0.18	0	0	0	0.15	0	0
0.62	0.13	0	0	0	0	0.25	0
0.37	0.13	0	0	0	0	0	0.50
1	0	0	0	0	0	0	0
0.74	0.26	0	0	0	0	0	0

Drought years: $P^2=$

0.65	0.22	0.13	0	0	0	0	0
0.68	0.21	0	0.11	0	0	0	0
0.59	0.24	0	0	0.17	0	0	0
0.44	0.14	0	0	0	0.42	0	0
0.33	0.10	0	0	0	0	0.57	0
0.33	0.067	0	0	0	0	0	0.60
0.92	0.083	0	0	0	0	0	0
0.67	0.33	0	0	0	0	0	0

It seems that these transitional probabilities in different steps are useful for the climatic diagnosis and probability prediction for flood/drought years. The similar transitional probabilities in any steps for flood/drought years can also be computed by (2). It has but been shown by matrix multiplieation that P^5 should almost approach the limiting matrix A. Referring to the results of these calculations, it is sufficient to call attension to the fact that we might consider any present state of different run-lengths of flood(drought) years of little influence on the predictions for any states of runs of flood (drought)years for five years later. This fact obtained agrees well with the investigation of a three state Markov chain with states R, N, and S for rain, nice, and snow respectively in the Land of Oz(Kemeny and Snell, 1960). The probabilities of transitions in 5 steps from any state to other states of runs longer than 4 years, conditions being considered as impossible events, are zeros, except that $S_0 \to S_0$, $S_1 \to S_1$ and $S_3 \to S_3$.

The mean first passage matrices for flood and drought years respectively in Beijing have been computed from (6) as follows:

Flood years: $M=$

1.4	3.8	12.4	37.0	123.5	249.0	500.0	501.0
1.6	5.5	8.5	33.2	119.7	245.2	496.2	497.2
1.6	5.4	13.9	24.6	111.1	236.6	487.6	488.6
1.6	5.5	14.0	38.6	86.5	212.0	463.0	464.0
2.0	5.8	14.4	39.0	125.5	125.5	376.5	377.5
2.0	5.8	14.4	39.0	125.5	251.0	251.0	252.0
2.0	5.8	14.4	39.0	125.5	251.0	502.0	1.00
1	4.8	14.4	38.0	124.5	250.0	501.0	502.0

Drought years: $M=$

1.6	3.0	10.5	39.4	69.3	98.4	124.3	167.0
1.7	4.7	7.5	36.4	66.3	95.4	121.3	164.0
1.8	4.8	12.3	28.9	58.8	87.9	113.7	156.5
2.6	5.6	13.1	42.0	29.9	59.0	84.8	127.6
2.7	5.7	13.3	42.1	72.0	29.1	55.0	97.7
2.4	5.4	12.9	41.8	71.7	100.8	25.9	68.6
1.8	4.8	12.3	41.2	71.0	100.2	126.0	42.8
1	4.0	11.5	40.4	70.3	99.4	125.3	168.0

It is of interest to see, for example, in Beijing that the mean number of years between S_0 and S_1, between S_0 and S_2, between S_0 and S_3, between S_0 and S_4, between S_0 and S_5, between S_0 and S_6 and between S_0 and S_7 for flood years were 3.8, 12.4, 37.0, 123.5, 249.0, 500.0 and 501.0 years respectively, while the corrponging values for drought years were 3.0, 10.5, 39.4, 69.3, 98.4, 124.3 and 167.0 years. The first passage times from S_7 to S_0 for both flood year and drought year were 1 year, because S_7 should turn into S_0 for any conditions in the historical records in Beijing.

The mean recurrence times of states S_1, S_2, S_3, S_4, S_5, S_6 and S_7 for flood years were 5.5, 13.9, 38.6, 125.5, 251.0, 502.0, and 502.0 years respectively, while the corresponding values for drought years were 4.7, 12.3, 42.0, 72.0, 100.8, 126.0 and 168.0 years respectively. The observed values of mean recurrence times of states S_1, S_2, S_3 were 5.5, 13.7, 40.1 and 4.7, 13.1, 41.3 for flood years and for drought years respectively, with which the theoretical values are in good agreement. However, there exist large errors for the observed values of run-lengths more than 4 owing to the rareness of observed long runs for both flood years and drought years.

On the whole, the mean numbers of years between any state of runs and another were smaller for drought years than for flood years, except that between any states of runs and the run-length of 3 years. Seeing that the long run-lengths of drought years are of the rare cases, the most cyclic periods are more frequent for flood years than for drought years, and thus a flood-dominant climate in Beijing had prevailed in the last 500 years, though it inclined to a slight domination effect(Yao, 1982).

The variance matrices computed from (8) for flood and drought years in Beijing have been given respectively as follows:

Flood years: $V =$

$$\begin{pmatrix} 0.8 & 10.9 & 120.8 & 1209 & 14489 & 59979 & 244958 & 244958 \\ 1.0 & 11.9 & 109.9 & 1198 & 14478 & 59968 & 244947 & 244947 \\ 1.1 & 12.0 & 121.9 & 1088 & 14368 & 59858 & 244837 & 244838 \\ 1.3 & 2.2 & 122.1 & 1211 & 13280 & 58770 & 243749 & 243749 \\ 1.5 & 2.4 & 122.3 & 1211 & 14491 & 45490 & 230469 & 230749 \\ 1.0 & 11.9 & 121.8 & 1210 & 14490 & 59980 & 184979 & 184979 \\ 0.0 & 10.9 & 120.8 & 1209 & 14489 & 59979 & 244958 & 0.00 \\ 0 & 10.9 & 120.8 & 1209 & 14489 & 59979 & 244958 & 244958 \end{pmatrix}$$

Drought years: $V =$

$$\begin{pmatrix} 1.1 & 6.0 & 84.7 & 1388 & 4380 & 8899 & 14211 & 25920 \\ 1.6 & 7.6 & 78.7 & 1382 & 4374 & 8893 & 14205 & 25914 \\ 2.2 & 8.2 & 86.9 & 1304 & 4295 & 8814 & 14126 & 25836 \\ 2.7 & 8.7 & 87.4 & 1391 & 2992 & 7510 & 12823 & 24532 \\ 1.6 & 7.6 & 86.3 & 1390 & 4382 & 4518 & 98307 & 21540 \\ 0.6 & 6.6 & 85.3 & 1389 & 4381 & 8899 & 53122 & 17022 \\ 0.2 & 6.2 & 84.9 & 1388 & 4380 & 8899 & 14211 & 11709 \\ 0 & 6.0 & 84.7 & 1388 & 4380 & 8899 & 14211 & 25920 \end{pmatrix}$$

The following results which would be useful to analyse and predicte flood/drought years at least for the purpose of reference were arrived at from these variance matrices: 1) Variabilities of the first time from any state to another for flood/drought years are quickly enlarged as long as the run-lengths of these years are lengthened. 2) Variabilities of the first passage time of runs of flood/drought years are nearly independent of the starting states.

From (14), the fractiles of occupation times, $N_k(n)$, of runs of flood/drought years for historical records in Beijing are presented as follows:

Fractile of occupation times	Run of flood years						
	S_1	S_2	S_3	S_4	S_5	S_6	S_7
5 % fractile	82	28	8	1	0	0	0
25 % fractile	88	33	11	3	1	0	0
50 % fractile	92	36	13	4	2	1	1
75 % fractile	96	39	15	5	3	2	2
95 % fractile	102	44	18	7	4	3	3
Fractile of occupation times	Run of drought years						
	S_1	S_2	S_3	S_4	S_5	S_6	S_7
5 % fractile	98	33	7	3	2	1	0
25 % fractile	104	38	10	5	4	3	2
50 % fractile	108	41	12	7	5	4	3
75 % fractile	112	44	14	9	6	5	4
95 % fractile	118	49	17	11	8	7	6

These fractiles characterize the theoretical distribution of occupation times of the run S_k. The 50% fractile called the median corresponds to the most probable occupation time of runs, if the distribution of occupation times of runs approaches the normal distribution for large samples. The point of present interest is that the fractiles of occupation times of runs in Beijing were in general smaller for flood years than for drought years except for fractiles of occupation times of S_3. The observed values of occupation times of S_1, S_2, S_3, S_4, S_5, S_6, S_7 for flood years and drought years were 92, 35, 12, 3, 1, 1 and 105, 38, 11, 5, 3, 2 respectively. Now, a matter for our consideration is that for flood years the observed values of occupation times of runs approximate to the calculated 50% fractiles of occupation times while for drought years the observed values approximate only to the 25% fractiles of occupation times. In connection with these facts it should be noted that the distribution of occupation times of runs of drought years requires a sufficient large sample in order that (14) may be satisfied. Another point to which attention may be drawn is that the fractiles calculated for runs of long run-lengths, such as S_5, S_6 and S_7, were not so exact as ones calculated for runs of shorter run-lengths, for the distribution of occupation times of the long, rare run-lengths in Beijing can not approach the normal distribution especially so far as the drought years are concerned.

REFERENCES

1. Billard, L. and M. R. Meshkani, 1980: Modelling weather data as a Markov chain. Statistical Climatology, Proc. First Int. Conf. on Statistical Climatology, 1979, Hachioji, Japan, S. Ikeda et al. Elsevier, 149-163.

2. Feller, W., 1968: An Introduction to Probability Theory and Its Applications. Vol. 1, 3rd ed., Wiley, 461pp.

3. Gabriel, K. R., 1959: The distribution of the number of successes in a sequence of dependent trials. Biometrika, 96, 959-960.

4. Gabriel, K. R. and J. Neumann, 1962: A Markov chain model for daily rainfall occurrence at Tel Aviv. Quart. J. Roy. Meteor. Soc., 88, 90-95.

5. Kemeny J. G. and J. L. Snell, 1960: Finite Markov Chains. Van Nostrand, 210pp.

6. Parzen E., 1962: Stochastic Processes. Holden-Day, 324pp.

7. Yao, C. S., 1966: The probabilities of stochastic changes of wet and dry days. Acta Meteor. Sinica, 36, 249-260.

8. ——— , 1982: A statistical approach to historical records of flood and drought. J. Appl. Meteor., 21, 588-594.

9. ——— , 1983: Fitting a linear autoregressive model for long-range forecasting. Mon. Weather Rev., April.

10. Yevjevich, V. 1972: Stochastic Processes in Hydrology. Water Resources Publications, Fort Collins, 276pp.

A BOX-JENKINS ANALYSIS OF SERIES OF WINTER-STRENGTH IN THE DANUBE REGION

Konrad Cehak

Institute for Meteorology and Geophysics of the University of Vienna
and
Central Office for Meteorology and Geodynamics in Vienna
A 1190 Vienna, Austria

While a large number of scientistis, who in former times had been engaged in research in numerical weather forecasts, have taken up the modeling of climate in the sense of extensions of their former models of the general circulation another kind of models has been established during the latest two decades, the time series models. One of the sources of this stream of statistical investigations was the series of lectures given by C.L. Godske (1966) at the WMO Seminar on the Application of Statistics in Meteorology and Climatology held in Paris in 1962. The results of the statistical models are not only of interest for their own sake but also for the application in the numerical models as data basis as well as in giving the background against which the output of the numerical climate models has to be verified.

The time series models widely used are models in the frequency domain. This could be caused by the fact that the first time series models (Rice, 1954) had been applied to communications problems and the engineers are used to think in terms of frequencies. A lot of standard computer programs were written and included into the software packages of each computer firm, which especially after the invention of the fast Fourier analysis made the way open for the use of the variance spectrum analysis in Climatology. Most of the applications were made in the field of climatic change in order to detect rhythms in the climatological time series. The advent of the maximum entropy method for the calculation of variance spectra was again a step forward in the application of the analysis of variance spectra. This method helps efficiently to find unknown periodicities in data series and permits also to make considerations about the significance of the spectral peaks hinting at the rhythms looked for.

However, the frequency domain models do not always give the same clear insight into the underlying statistical processes as the time domain models. The theory of these models in quite advanced and it is used in the practical work in the form of the so-called Box-Jenkins analysis (Anderson, 1976), which has found further development by various authors recently. The Box-Jenkins approach to modelling is not fully objective but, nevertheless, it yields very good informations on the statistical process behind the observations. The method shows even the way for making forecasts, which of course not always may meet the final purpose of the forecaster, so one must be careful especially in this part of the work.

The Box-Jenkins Analysis is applied here to five climatological time series, namely the series of a parameter for the strength of the winter season at five stations in the Danube Region of Bavaria, Austria and Slovakia. They have in common that the length of the observations amount to at least 200 years and the climatologists in charge of these series have done a good job in making them homogeneous as far as this has been necessary. The parameter used to describe the winter strength is the sum of negative daily mean temperatures during the course of the three winter months December, January and February each year. It can be shown that this parameter is closely correlated to other similar characteristic values for the winter strength and it is easily obtained from the data. Table 1 lists the simple statistical characteristics of the five series.

The visual inspection of the five series could give the impression that they are not quite stationary, however, the degree of this instationarity might be rather low. Therefore, the Box-Jenkins Analysis was started by the first guess that the effective statistical process might be an integrated autoregressive, moving average process ARIMA(p,d,q). The lengths p and q of the autoregressive and moving average parts have to be determined as well as the degree d of the unstationa-

Fig. 1. Hohenpeißenberg. Negative Temperature Sums.

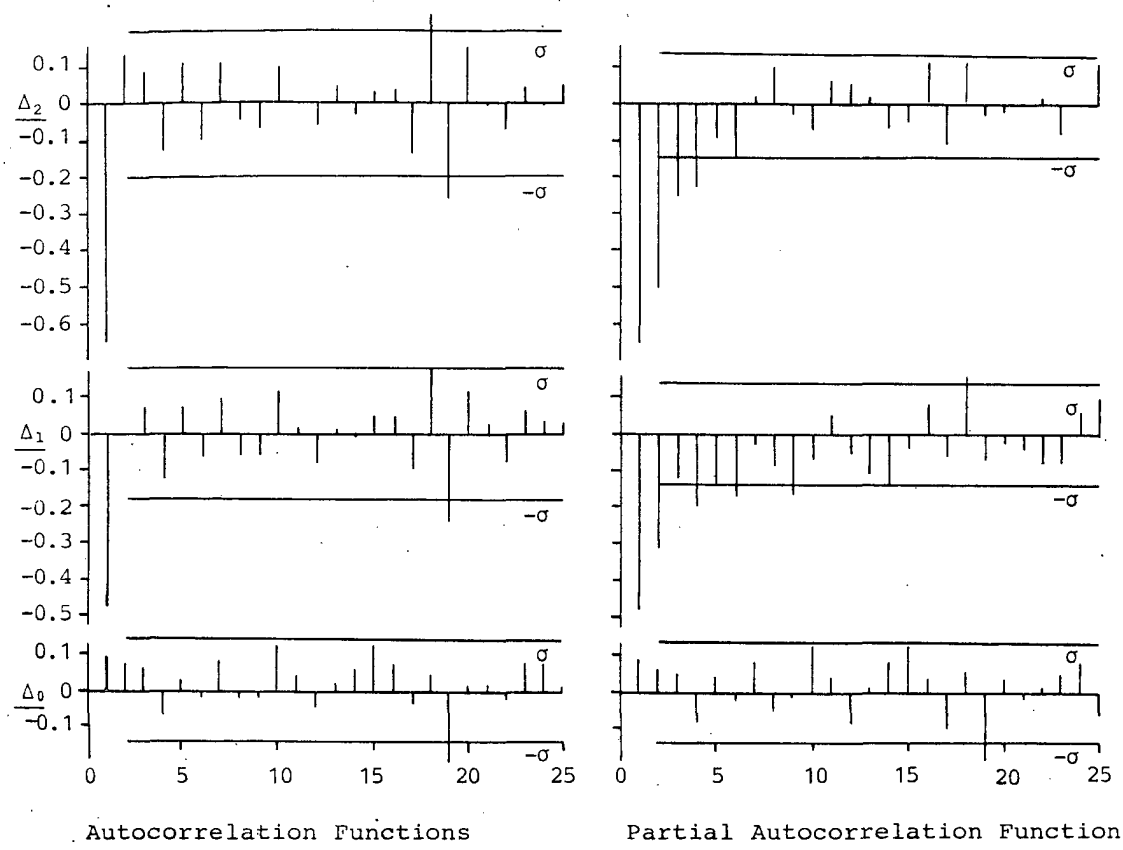


Table 1: Mean value $\bar{x}$, standard deviation s, coefficients of skewness g_1 and kurtosis g_2, minimum, maximum and first order autocorrelation coefficient r_1 of the series of negative temperature sums at the five stations in the Danube Region.

Station	Observations	$\bar{x}$	s	g_1	g_2	r_1	min	max
Hohenpeißenberg	1781/2-1977/8	3593.6	1286.6	0.5	-0.3	0.09	1264	7232
München	1781/2-1977/8	2777.9	1356.1	0.9	0.7	0.08	374	7649
Kremsmünster	1765/6-1978/9	2763.3	1402.9	1.0	1.5	0.15	295	8937
Wien	1775/6-1978/9	1923.0	1189.5	1.0	0.9	0.06	72	6626
Bratislava	1775/6-1977/8	1958.8	1212.8	0.9	0.7	0.005	55	6758

rity. Therefore, autocorrelation functions up to the lag-maximum of 25 were calculated for the data series itself and for the first and second differences. The resulting three autocorrelation functions are represented in figure 1.

The first step in the analysis of the ARIMA process is the determination of d, the number of differences necessary. By using differences in stead of the original values shifts of the level or of the slope of the process could be removed. The resulting series should look like a stationary series. In figure 1 already the original autocorrelation function shows the features of a stationary process, the series of the differences are more correlated between themselves than the undifferenced series. Hence, differencing seems not to be necessary, the degree of unstationarity is set d = 0, and instead of the ARIMA process one gets an ARMA(p,q) process.

The estimation of the length p of the autoregressive part of the model is based on Bartlett's formula (Bartlett, 1946) for the variance of the autocorrelation coefficients for an index k > K:

$$\mathrm{var}(r_k) = \frac{1}{N} \sum_{i=-K}^{K} \rho_i^2 \ .$$

In this case the autocorrelation coefficients of higher than first order seem to be so small that K is set to 1 and, therefore,

$$\mathrm{var}(r_k) = \frac{2}{N} \rho_1^2 \quad \text{for} \quad k > 1 \ .$$

Straight lines in the distance $\pm\sigma$ are drawn in the figure and it is to be seen that only the magnitude of the autocorrelation coefficient for the lag 19 is larger than σ. Thus, there is no contradiction against the supposition that the order of the autoregressive part is $p = 1$.

The partial autocorrelation function of an autoregressive process of first order is

$$\phi_{11} = \rho_1$$
$$\phi_{kk} = 0 \quad k > 1$$

This seems to be fulfilled in our case, too.

The estimation of the length q of the moving average part makes use of the duality between the behaviour of the autocorrelation function and that of the partial autocorrelation function of the moving average and the autoregressive process. The result is that the moving average part of the process might be of the same order 1 like the autoregression or even of order 0.

The estimations of the lengths p and q suggested the use of an ARMA(1,1) model. The admissible region for the autocorrelation coefficients for such a model are given by

$$2r_1^2 - |r_1| < r_2 < |r_1| \ .$$

This inequality is fullfilled in each of the five stations. Next, the coefficients for the model representation

$$z_i = \phi z_{i-1} + a_1 + \theta a_{i-1}$$

are calculated from

$$\phi = r_2 : r_1$$

$$\theta = (b \pm \sqrt{b^2 - 4}) : 2$$

$$b = (1-2r_2+\phi^2) : (r_1-\phi) \ .$$

These coefficients, together with the input values of the first two autocorrelation coefficients are contained in table 2. The "shocks" a_i are normally distributed with mean value 0 and the standard deviation given in table 2, they are part of a white noise random process. The variance of a_i is given by

$$\sigma_a^2 = c_o(1-\phi^2) : (1-2\phi\theta+\theta^2)$$

c_o being the variance of the observations. Finally, table 2 contains even the standard deviations of the estimates ϕ and θ, derived from

$$\sigma_\phi^2 = (1-\phi)^2 (1+\phi\theta)^2 : N(\phi+\theta)^2$$

$$\sigma_\theta^2 = (1-\theta)^2 (1+\phi\theta)^2 : N(\phi+\theta)^2$$

N being the length of the series. If we regard only those coefficients to be significant, which are in magnitude larger than twice their standard deviation, we see that only ϕ for Hohenpeißenberg and Kremsmünster are significant, that of Bratislava is nearly so.

This last result suggests that one could consider the series of Hohenpeißenberg and Kremsmünster as autoregressive series of first order (AR(1)-processes), the other ones could be taken as white noise processes. Taking into consideration the small magnitude of r_1 at Hohenpeißenberg and Kremsmünster and the other

Table 2: Characteristic numbers for the ARMA(1,1) fit for the five series. The stations are represented by their initials.

		H	M	K	W	B
Autocorrelation-coefficients	r_1	0.09	0.08	0.15	0.06	0.05
	r_2	0.07	-0.01	0.07	-0.00	0.03
Coefficients	$\hat{\phi}$	0.78	-0.13	0.47	-0.08	0.60
	$\hat{\theta}$	-0.80	0.21	-0.33	0.14	-0.55
Standard deviations of						
	$\hat{\phi}$	0.29	0.97	0.23	1.26	0.38
	$\hat{\theta}$	2.39	0.68	0.57	1.00	1.47
shocks	a_i	1283	1348	1382	1173	1210

equations valid for ϕ and its standard deviation in the case of an AR(1), namely

$$\phi = r_1$$
$$\sigma^2_\phi = c_o(1-r_1^2)$$

which again gives no significance for the estimations of ϕ, one is led to accept the white noise model even for those two stations.

This result is not very encouraging, if one had wanted to use the time series for forecasting the climate for certain future years. With a white noise as background model the forecasting errors would be rather high, even if one had the good luck that no climatic change would occur in form of a change of level or trend of the series. On the other hand the result shows the constant character of the climate at least for the latest 200 years. If there would be no change in this character one could at least simulate the future behavious of the winterstrength in this Pre-Alpine region and derive from this basic material for economic forecasts, as the winterstrength is of utmost importance in the use of energy in which form ever one has to procure it.

REFERENCES

Anderson, O.D., 1976: Time Series Analysis and Forecasting. Butterworth, London and Boston.

Bartlett, M.S., 1946: On the theoretical specification of sampling properties of autocorrelated time series. J.R. Statist. Soc., B 8, 27-41.

Godske, C.L., 1966: Methods of Statistics and some Applications to Climatology. WMO Technical Note No. 71: Statistical Analysis and Prognosis in Meteorology, 9-86.

Rice, S.O., 1954: Mathematical analysis of random noise. In: Wax, E., Ed.: Selected Papers on Noise and Stochastic Processes. Dover, New York, 133-294.

EVIDENCE FOR A 40,000 YEAR SCALING REGIME IN CLIMATOLOGICAL TEMPERATURES

S. Lovejoy,
D. Schertzer

Météorologie Nationale,
EERM/CRMD,
2 Ave. Rapp,
Paris, 75007

1. INTRODUCTION:

Geochemically infered paleotemperatures have conclusively established the existence of climatological temperature fluctuations at all observed time scales. This is not surprising since any system of equations governing climate evolution is certainly quite non-linear; as Lorenz (1964) first demonstrated, such non-linearity is likely to give rise to an erratically varying climate. Instrumental records and climatological proxy records of temperature dataare used in this paper to show that climatological temperature fluctuations from at least 5 to about 40,000 years are scaling, with parameter $H \sim 0.4$. This means that the energy of fluctuations at frequency f is of the form $f^{-(1+2H)} = f^{-\beta} \sim f^{-1.8}$ which is also very close to the $f^{-5/3}$ spectrum predicted for the temperature in isotropic, homogeneous turbulence in a fluid.

A function T(t) is said to be scaling with parameter H if the fluctuations ΔT follow the scaling relationship:

$$\Delta T(\lambda \Delta t) \stackrel{d}{=} \lambda^H \Delta T(\Delta t) \qquad (1)$$

for arbitrary Δt and factor λ. Equality is $\stackrel{d}{=}$ is understood in the sense of probablity distributions.

Mandelbrot (1975, 1982) has shown that scaling is widespread in nature, and scaling objects are central to the theory of fractals. Of course, the idea of scaling is present in the work of Richardson, and Kolmogorov but its extent and signifigance was underestimated. In particular, Mandelbrot and Wallis (1969) find a number of long geophysical time series, including varves and sediments, as well as climate related series such as tree rings, river discharges and precipitation to be scaling.

Although no connection between scaling and the equations is made here, some progress has recently been made in the important case of the Navier-Stokes equations for fluid flow (see Chorin (1981)).

2. DATA ANALYSIS:

An empirical test of the scaling hypothesis and a determination of H and requires the examination of temperatures or temperature spectra. Spectra have already been extensively investigated as part of the search for the effects of periodic forcing. These spectra do indeed show a broadly $f^{-1.8}$ form (e.g. Hays et Al (1976)), a fact confirmed by the present authors for the data described below. As argued below, it is signifigant that this is less steep than the f^{-2} (H=1/2) "Browninan motion" spectrum proposed theoretically by Hasselman (1976), Battacharya et Al (1982) and Letreut and Ghil (1982). Most time series analysed spectrally so far, covered only a decade or so in frequency hence the signifigance of the overall $f^{-1.8}$ envelope has not been widely appreciated. In the following, we analyse temperature and paleotemperature series in the time domain rather than in the frequency domain. Direct analysis has the advantage that the result can be directly interpreted in terms of a physically meaningful temperature difference.

In order to establish equation (1) for climatological temperatures, we empirically determine the standard deviation of the temperature change denoted by ΔT as a function of the time interval Δt. This standard deviation, denoted $S(\Delta t)$, is the square root of the delta variance or structure function and is defined as follows:

$$S(\Delta t) = \langle (\Delta T(\Delta t))^2 \rangle^{1/2}$$

H can be determined from the slope of a log S vs log Δt graph.

Fig. 1 shows the result of such an analysis when Manley's (1974) central England data covering the period 1659-1974, and Budyko's (1969) northern hemisphere data covering the period 1880-1969 have been used to determine the ΔT behaviour at small Δt. The Manley data shows that S is essentially independent of Δt over the $\sim$300 year period that it covers, indicating that central England temperatures are roughly stationary over this period, with fluctuations $\sim .88$K (a fact confirmed by spectral analysis e.g. Mason (1976), and Kutzbach and Bryson (1974)). The Budyko northern hemisphere temperatures are however quite different, following the law $S(\Delta t) \sim .077 \Delta t^H$ K with $H \sim 0.4$ (c.f. the straight line in fig. 1). Unfortunately, Budyko passed his data through a 5 year moving average, so that they cannot be used to study fluctuations over shorter time scales. The behaviour of $S(\Delta t)$ for $\Delta t < 5$ years is an

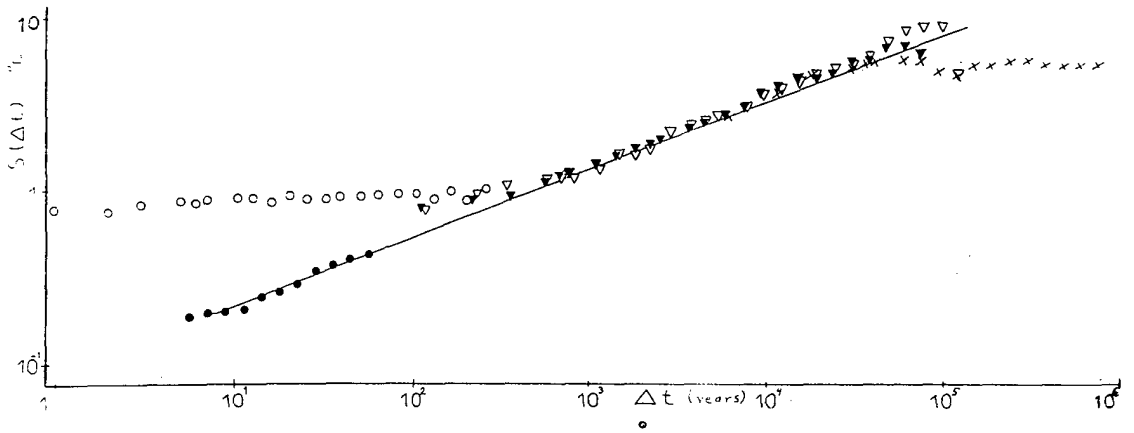


Fig. 1: A plot of $S(\Delta t)$ against Δt for the following temperature data (see text for calibration details): black circles, Budyko northern hemisphere averages (Budyko (1969)), white circles, Manley (1974) central England temperatures, black traingles, antartic ice core data, and white triangles, artic core data (both from Johnsen et Al (1972)), X's are ocean core data (Shackleton and Opdyke (1973)). The straight line is the function $S(\Delta t) = 0.077\ \Delta t^{0.4}$.

important object of future research. One interesting possiblity is the global $S(\Delta t)$ function decreases more or less contiuously until it joins the scaling regime which is known to hold locally- at least on time scales from seconds to days (e.g. Monin and Yaglom (1975). If the line shown in fig. 1 is extrapolated back to Δt=1 year, we find $S(1) \sim 0.077$K, which is comparable to other estimates of the characteristic yearly temperature fluctuation (e.g. Fraederich (1978) uses the value 0.16K- the reason for the discrepancy between the two values is probably due to the difficulty of accurately defing the hemispheric temperature variability over a year).

The difference of a factor $\sim$10 between the yearly variation in Central England and that of the northern hemisphere reflects the fact that regional variations are largely suppressed when averages are taken over an entire hemisphere. Indeed, hemispheric variations cannot be expected to dominate over local ones until $0.077\Delta t^{0.4} > 0.88$ or until $\Delta t > 450$ years. We therefore expect local temperatures to be approximately stationary (flat $S(\Delta t)$ function) until $\sim$450 years after which they should follow the formula $S(\Delta t) = 0.077\Delta t^{0.4}$. A systematic study of the different relationship between temperature fluctuations over different spatial and temporal scales would undoubtedly clarify the nature of this apparent transition at 450 years (see Mitchell (1976) for some discussion of the relationship of temporal and spatial variability). This transition, that we propose to call a "local/global bifurcation", may apply to other meterological/ climatological parameters.

These ideas may be tested by using O^{18}/O^{16} ratios taken from Artic and Antartic ice cores as proxy for paleotemperatures (see Dupplessy (1978) for a discussion). Ice cores have a deficit of O^{18} because the lighter H_2O^{16} evaporates more readily from the oceans than the heavier H_2O^{18}. This deficit is highly correlated with the temperature at the time of evaporation and deposition. If this is true, then we expect O^{18}/O^{16} ratios to provide an estimate of local temperatures since the local environment over the ice cap is expected to strongly influence these processes. Fig. 1 shows $S(\Delta t)$ for O^{18}/O^{16} ratios from both Artic and Antarctic cores. Both analyses on high resolution (every $\sim$100 years for $2X10^4$ years) as well as low resolution (every $\sim$700 years for the last $1.2X10^5$ and $9X10^4$ years for artic and antarctic respectively) data. The data was taken from Johnsen et Al (1972). The value of $S(\Delta t)$ at $\Delta t \sim 300$ years (Manley data) has been used as a calibration point for these proxy records. When this is done, $S(\Delta t)$ for the paleotemperatures follows very closely the predicted behaviourn i.e. it falls on the straight line in fig. 1.

This behaviour must eventurally break down- if only because the temperature of the earth is bounded. In the ice data, evidence for this change in behaviour comes from the peak in $S(\Delta t)$ which occurs unfortunately near the end of these data sets (at $\Delta t \sim 8X10^4$ years). In order to investigate $S(\Delta t)$ at these scales, we turn to the analysis of data from deep-sea cores.

The interpretation of deep-sea cores is not as straightforward as that of ice cores. For example, O^{18}/O^{16} ratios may be obtained from the calcium carbonate deposited by fossil plankton. In this case, the ratios depend not only on the temperature of the surface layer where the plankton once lived, but also on the O^{18}/O^{16} ratio in their surrounding water, which in turn depends on the volume of water in the ice caps. It is now generally agreed that the latter is the dominant effect (see Dupplessy (1978)), implying that the ocean core ratios are best interpreted in terms of the volume

of water in the ice caps. However, on a time scale of 10^4 to 10^6 years, fluctuations in the size of the ice caps are surely correlated with the mean global temperature. This justifies using these cores as temperature proxies in the region 10^4 -10^6 years.

In fig. 1 we have therefore plotted data analysed from Shackleton and Opdyke (1973). The $S(\Delta t)$ curve obtained has a scaling behaviour with $H \sim 0.4$ for $\Delta t < 4X10^4$ years, after which it is fairly constant. The curve was calibrated by aligning the $H \sim 0.4$ part to the ice core data according to the dating specified in Shackleton and Opdyke (1973). This curve fits the ice data fairly well except for the region near the peak in the ice data. If this peak is real, then it is an indication of the existence of an oscillatory behaviour of T on a scale of $2X4X10^4 = 8X10^4$ years (the factor 2 is necessary because $4X10^4$ corresonds to a half-period).

The above analysis suggests the following simple interpretation of temperature change in the last $9X10^5$ years. The yearly temperature change of 0.077K can be extrapolated to $\sim 4X10^4$ years by assuming it is scaling, with exponent $H \sim 0.4$. This yields a net temperature change of $\sim 0.077(4X10^4)^{0.4} \sim 5.3$K over this period. At about this time scale the process starts to have a stationary variance (which is compatible with oscillatory behaviour). Thus, over scales longer than $\sim 10^5$ years, the temperature varies by $\pm(5.3/2) \sim \pm 2.7$K with a characteristic time of $2X4X10^4 \sim 8X10^4$ years (it typically takes $4X10^4$ years for the temperature to vary from one extreme to another, or for half of a complete "oscillation"). The exact nature and scope of this behaviour is outside the scope of this paper. The above interpretation is in accord with the independent Emiliani and Shackleton (1974) analysis showing temperature variations of ±3K with about 6-7 strong oscillations in the last $4\text{-}6X10^5$ years, depending on the dating used.

3. THE ACCURACY OF THE SCALING EXPONENT H:

It is important to establish the accuracy of our estimate of H since the amplitude of ΔT for $\Delta t > 40{,}000$ years is very sensitive to its exact value. For example, H=.5 yields a temperature change of ~ 10.8K ($=\pm 5.4$K) which is probably too large to be compatible with other estimates of the magnitude of ΔT, and H=1/3 yields $\Delta T \sim 3.4$K ($=\pm 1.7$K), probably too small.

The value of H is also important theoretically. For instance, the value of H=1/3 is the temperture scaling for isotropic homogeneous turbulence (e.g. Obukhov (1949) and Corrsin (1951)). (It is interesting to note that the experimental tests of Kolmogorov scaling may in fact support $H \sim 0.4$ rather than H=1/3. For example Tatarskii (1956) obtained H 0.41 and Deschamps et Al (1978) H 0.40± .1). The value H=1/2 is predicted for Brown noise and certain types of deterministic aperiodic behaviour (e.g. Hasselman (1976) and Letreut and Ghil (1983) respectively).

In order to estimate the accuracy of H, a robust measure of long-run dependence known as R/S analysis was used (see Mandelbrot (1972)). This statistic is robust, because it is insensitive to large fluctuations around the long-term trend. It has the disadvantage that it is biased for short time series. When it was applied to the Artic and Antarctic ice core series in the range $4{,}000 < \Delta t < 40{,}000$ years, this method yielded H=0.420 and 0.404 respectively (with correlation coefficients between log R/S and log Δt of .995, .985). We therefore estimate $H \sim 0.41 \pm .02$. However, a formula more precise than $S(\Delta t) = 0.077 \Delta t^{0.4}$ is not warranted because of the uncertainties in the value of S(1).

4. INTERPRETATION:

If climatological temperature fluctuations are to be explained by the action of invariant physical laws, as expressed by the equations of the climate, then it is important to establish the stationarity of certain statistics of the fluctuations. The variation of $S(\Delta t)$ in fig. 1 indicates that the temperature is non-stationary on scales $\Delta t < 40{,}000$ years. However, it is possible that the temperature increments themselves are stationary. We therefore define $S'(\Delta t)$ as the square root of the delta variance

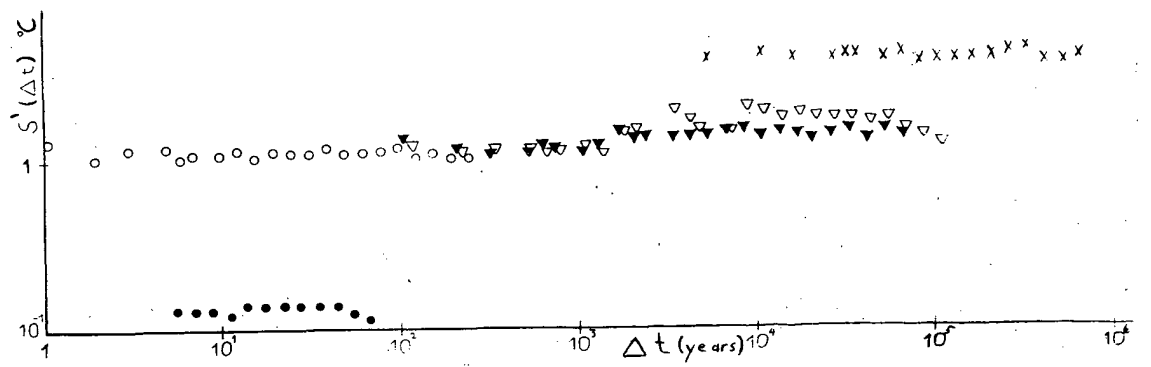


Fig. 2: Same as fig. 1 except $S'(\Delta t)$ is determined from the time series of temperature and paleotemperature increments as explained in the text. The paleotemperatures are calibrated as in fig. 1.

of the increments in T and our paleotemperatures ($S'(\Delta t)$ is calculated in the same way as $S(\Delta t)$ but for the time series composed of the first differences in T rather than T itself.). Using the same calibration as before, we obtain the fairly constant $S'(\Delta t)$ functions shown in fig.2. Note that the magnitude of the $S'(\Delta t)$ function is in general different for the different kinds of data. However the ice data have a very similar $S'(\Delta t)$ value to the Manley central England data, a fact which supports their use as proxies for local temperatures.

If we accept the stationarity of the increments, then it is possible to imagine many different mechanisms to account for the empirical $S(\Delta t)$ function. For example, turbulence theory might lead us to interpret the change in slope at $\Delta t \sim 40{,}000$ years to mark the boundary between a regime dominated by energy injection (characterised by the magnitude $\pm 2.7K$), and a regime dominated by the cascade of energy to smaller and smaller scales via non-linear interactions (for $\Delta t < 40{,}000$ years). Another interpretation not necessarily in contradiction to the first, is that in the scaling regime, the temperature changes by the accumulationof random fluctuations. If these fluctuations were independent, then the temperature would vary in the same way as the coordinate of a Brownian particle: we would obtain H=.5 as stated before. the fact that H<.5 indicates that fluctuations have a tendency to cancel, even over very long time periods. Physically, this could result form the exist .ce of negative feedback between the various parts of the climate system operating at all time scales. At $\Delta t \sim 40{,}000$ years, the magnitude of the fluctuations may be such that new physical processes become important and prevent the temperature from straying outside a fairly well defined range of $\pm 2.7K$.

5. CONCLUSIONS:

We may summarise our results as follows:

a) The climatological temperature averaged over a hemisphere is less than the magnitude of local variations ($\sim 0.88K$) for $\Delta t < 450$ years, after which it dominates the local fluctuations.

b) The hemispheric temperature fluctuations have a magnitude of 0.077 K/year and vary according to the scaling law $S(\Delta t) \sim 0.077 \Delta t^H$ with $H \sim 0.4$ for $5 < \Delta t < 4X10^4$ years.

c) For time scales longer than $4X10^4$ years, the temperature variations are stationary on a characteristic time scale of $2X4X10^4$ $8X10^4$ years, with amplitude $\pm 2.7K$.

d) The increments in temperature are stationary. This allows temperature fluctuations to be interpreted in terms of a non-linear cascade or as a series of feedback loops operating at all scales up to 40,000 years.

If the above analysis is correct, then it may be possible to calibrate paleotemperatures by fitting the empirical $S(\Delta t)$ curve to that shown in fig. 1. This method would work if the required calibration in either temperature or age is approximately linear.

Finally, it is interesting to speculate on the behaviour of the hemispheric $S(\Delta t)$ for $\Delta t < 5$ years. It is possible that it continues unabated down to scales of the order of seconds where a very similar scaling of the local temperatures is known to hold. This "local/global bifurcation" is an important subject for future research.

6. ACKNOWLEDGEMENTS:

We would like to thank J.C. Duplessy, P. Pestiaux, P. Ladoy and J.P. Lafore for helful discussions and comments. We especially thank B. Mandelbrot for providing advice and constructive criticism throughout the project. M. Ghil is also thanked for helpful comments on earlier drafts of this paper. S. Lovejoy would also like to acknowledge the National Science and Engineering REsearch Council (Canada) for financial support.

7. REFERENCES:

Bhattacharya, K., M. Ghil, I.L. Vulis, 1982: Internal variability of an energy-balance model with delayed albedo effects. J. Atmos. Sci., 39, 2362-2367.

Budyko, M.I. 1969: The effect of solar radiation variations on the climate of the earth. Tellus 21, 611-619.

Chorin, A.J. 1981: Estimates of intermittency, spectra, and blow-up in developped turbulence. Comm. Pure Appl. Math., 34, 853-866.

Corrsin, S. 1951: On the spectrum of isotropic temperature fluctuations in an isotropic turbulence. J. Appl. Phys., 22, 469-473.

Deschamps, P.Y., R. Frouin, L. Wald 1981: Satellite determination of the meso-scale variability of the sea surface temperature. J. Phys. Ocean., 11, 864-870.

Dupplessy, J.C. 1978: Climatic Change, J. Gribbin ed. p. 46-67.

Emiliani, C., N.J. Shackleton 1974: The Bruhnes epoch: Isotopic paleotemperatures and geochronology. Science, 183, 511-514.

Fraedrich, K. 1978: Structural and stochastic analysis of a zero-dimensional climate system. Q. J. Roy. Met. Soc. 104, 461-474.

Hasselman, K. 1976: Stochastic climate models Part I; Theory. Tellus, 28, 473-485.

Hays, J.D.,J. Imbrie, N.J. Shackleton 1976: Variations in the earth's orbit: pacemaker of the ice ages. Science, 194, 1121-1132.

Johnsen, S.J., W. Dansgaard, H.B. Clausen, C.C. Langway Jr. 1972: Oxygen isotope profiles through the Antarctic and Greenland ice sheets. Nature, 235, 429-433.

Kutzbach, J.E., R.A. Bryson 1974: Variance spectrum of Holocene climatic fluctuations in the north Atlantic sector. J. Atmos. Sci., 31, 1958-1963.

Le Treut, H., M. Ghil 1983: Orbital forcing, climatic interactions and glaciation cycles. In press J. Geophy. Res.

Lorenz, E.N. 1963: Deterministic non-periodic flow. J. Atmos. Sci.,20, 131 141.

Mandelbrot, B., J.R. Wallis 1969: Some long-run properties of geophysical records. Wat. Resources Res.,5, 321-340.

Mandelbrot, B. 1972: Statistical methodology for non-periodic cycles: from the covariance to R/S analysis. Annals of Econ. Soc. Meas., 1, 259-290.

Mandelbrot, B. 1975: Les objets fractals: formes, hasard et dimension. Paris, Flammarion.

Mandelbrot, B. 1982: The fractal geometry of nature. 468pp, Freeman and co., San Francisco.

Manley, G. 1974: Central England temperatures: monthly means 1659 to 1973. Quart. J. R. Met. Soc., 100, 389-405.

Mason, B.J. 1976: Towards the understanding and prediction of climatic variations. Quart. J. R. Met. Soc., 102, 473-498.

Mitchell, J.M. Jr. 1976: An overview of climatic variability and its causal mechanisms. Quat. Res., 6, 481-493.

Monin, A.S., A.M. Yaglom 1975: Statistical Fluid Mech. Vol. II, MIT Press, Boston.

Obukhov, A. 1949: Structure of the temperature field in a turbulent flow. Izv. Akad. Nauk. SSSR Ser. Geogr. i Geofiz, 13, 58-69.

Shackleton, N.J., N.D. Opdyke, 1973: Quat. Res., 3, 39-43.

Tatarskii,V.I. 1956: Microstructure of the temperature field in the atmospheric surface layer. Izv. Akad. Nauk. SSSR, Ser Geofiz. 6, 689-699.

LEVEL AND INTERANNUAL VARIABILITY OF SEASONAL MEAN TEMPERATURES IN EUROPE

C.J.E. Schuurmans and A.J. Coops

Institute of Meteorology and Oceanography
University of Utrecht
Utrecht, The Netherlands

1. INTRODUCTION

In discussions of climatic change it is often assumed, that changes of temperature level automatically involve changes of interannual variability of temperature. For instance, the much debated Northern Hemisphere cooling in the 1960's was expected to be paralleled by a general increase of climate variability (e.g. more extremes), which on the basis of observations, however, could not convincingly be established (Hare, 1979). Moreover, in studies adressing more directly the relation between temperature level and variability (Van Loon and Williams (1978)) a simple association did not appear to exist. Could these apparently negative results be due to the use of inappropriate statistical methods, or would the assumed relation indeed be absent?

Using some more rigorous methods, we studied this problem for seasonal mean temperatures in Europe. The study was based on a carefully selected set of data for the period 1875-1975. It includes the following 13 stations: Lisbon, Madrid, Milan, Geneva, Frankfurt, Brussels, Valentia (Ireland), Central England (Manley-series), De Bilt (The Netherlands), Warsaw, Copenhagen, Edinburgh and Stockholm.

Apart from trends and the relation referred to above spatial and seasonal coherences of the seasonal mean temperatures were also studied.

2. INTERANNUAL VARIABILITY

Interannual variability may be defined in many different ways. One possible way is to use the standard deviation of a set of yearly values. From the point of view of studying the relation with mean temperature level the use of the standard deviation of yearly temperature values is less appropriate, however. The problem is, that for a given length of the record, say 100 years, one wants to preserve as many degrees of freedom as possible, while on the other hand a meaningful definition of interannual variability for a given year (or number of years) is maintained. Van Loon and Williams (1978), who in fact used standard deviations to characterize interannual variability, tried to solve this problem by applying the method of moving averages. This indeed maintains the representativeness of the computed variability for a given year, but reduces the effective number of degrees of freedom to approximately N/n (N being the length of the record, n the number of years used to compute the standard deviation). In order to avoid this severe reduction of the number of degrees of freedom, as a measure of interannual variability we simply used the temperature difference of two consecutive years. Furthermore, using each temperature value of the original time series only once, we reduced the time dependency in the series still further. So for the series of 100 years, only 50 year-to-year differences were used, and compared with 50 two-year averages. Of course, this procedure may be applied twice, once starting with an even year and once with an odd year. Since there is no a priori preference for one of the two, the best estimate is taken to be the average of the outcomes of the two computations. The additional advantage of this procedure is, that the influence of a possible biennial oscillation on the results is completely eliminated.

Interannual variabilities of seasonal mean temperatures thus obtained are given in Fig. 1. At the lower latitudes and in coastal regions the interannual variability is less than $1^{o}C$; without much annual variation. At the higher latitudes and inland stations interannual variability is higher and in winter about twice as large as in summer, with intermediate values in spring and autumn. In general the patterns are most markedly influenced by the land-sea contrast.

3. METHODS OF ANALYSIS

The presence of trend in our series of temperatures and interannual variabilities of temperatures was investigated using the two sample Wilcoxon test, which is a non-parametric analogue of the well-known t-test (Hájek, 1969). For two(sub)samples (x_i) and (y_j) of size N_x and N_y, respectively, the test statistic W_2 is defined by

$$W_2 = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \text{sign}\,(x_i - y_j).$$

Under the null hypothesis, that trend is absent, it holds, that $EW_2 = 0$ and $\text{var}\, W_2 = N_x N_y (N_x+N_y+1)/3$, while for relatively large values of N_x and N_y W_2

is approximately normally distributed.

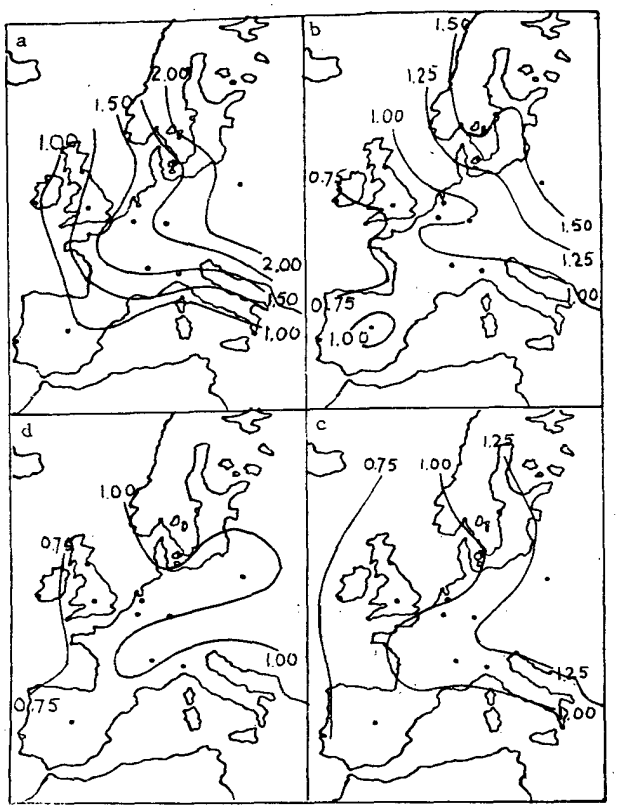


Fig. 1

Fig. 1. Interannual variability of mean winter (a), spring (b), summer (c) and autumn (d) temperatures in °C for the period 1875-1975.

To investigate the relation between temperature level and interannual variability of temperature (or between two series of either temperature or interannual variability) we used the (non-parametric) rank correlation method of Kendall (Kendall, 1962). In this method the relation between the variables x and y is expressed by the rank correlation coefficient τ defined as

$$\tau = \frac{2S}{N(N-1)} ,$$

where $S = \sum_{i<j} \text{sign}(x_i - x_j)\ \text{sign}(y_i - y_j)$

with $1 \leq i, j \leq N$ (N is the total number of variables in each of the two series, while x_i, y_i are the individual observations in these series).

Under the null hypothesis of complete independence between the two variables τ obeys a discrete distribution which is symmetrical with respect to zero, with variance

$$2(2N+5)/9N(N-1).$$

For W_2 as well as for τ a test statistic may be computed which has approximately a standard normal distribution for large (N_x, N_y) and N, respectively. Probabilities of exceeding a certain value of this quantity can easily be obtained from statistical handbooks.

It must be remarked, that the result of Kendall's rank correlation method is sensitive to the presence of trend in the data series. In our research for relations we therefore had to detrend our series or to carefully avoid the drawing of conclusions, if these could be influenced by trends. We have chosen for the latter procedure, mainly on the basis of our finding that at least the series of interannual variability, which we considered, were all free of trend (which means, that at none of the stations and for none of the seasons the null-hypothesis of no trend had to be rejected at the 5% level of statistical significance).

For further details concerning the application and validity of the above methods we may refer to the statistical textbooks.

4. TRENDS OF TEMPERATURE

In view of the remarks at the end of the foregoing section we need only consider here the trend in temperature level, interannual variability of temperature being found to be essentially trend free over the period considered.

First we present trend for the last 50 years of the period (1925-1975) with respect to the foregoing 50 years (1875-1925). See Table 1.

Table 1. Sign of the change of seasonal mean temperatures from 1875-1925 to 1925-1975. Only changes which are significant at the 5% level are indicated.

	Winter	Spring	Summer	Fall
Stockholm		+	+	+
Copenhagen	+	+	+	+
Edinburgh				+
Valentia (Irl.)		+		+
C. England		+	+	+
De Bilt				+
Brussels		+	+	+
Frankfurt		+	+	+
Warsaw		+	+	+
Geneva		+	+	+
Milan		+	+	+
Madrid	+	+	-	+
Lisbon	+	+	+	+

This table confirms the well-known warming in all seasons, which in Europe and also elswhere took place during the first half of this century (Lamb, 1972). This warming has been attributed to a combination of factors: increase of solar activity, lack of volcanic eruptions and an increased CO_2-concentration of the atmosphere. Only recently it seems, that the latter factor is receiving increased emphasis (Kellogg and Bojkov, 1982). Although from our analysis we are not able to discriminate between any of the possible causes mentioned, two of our trend results are certainly not in favour of an assumed CO_2-cause. These additional results are: 1. In winter the warming is smaller (in absolute value) than in any of the other seasons, which is in contradiction with results of GCM-sensitivity

studies (Manabe and Stouffer, 1980). See Fig. 2 for the observed seasonal trends.

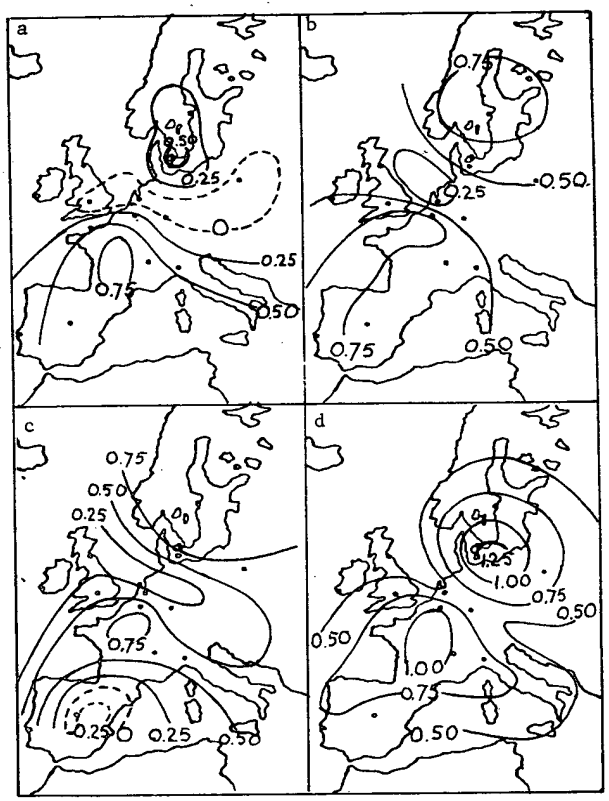


Fig. 2

Fig. 2. Changes of seasonal mean temperature from 1875-1925 to 1925-1975 for winter (a), spring (b), summer (c) and autumn (d).

2. On the basis of the CO_2-warming hypothesis one would expect, that the ongoing and even accelerated increase of CO_2-concentration in the last 50 years would give rise to a continuation of the upward temperature trend. According to our table 2 this is certainly not the case. Rather a significant decrease of temperature occurred at a number of stations. Only wintertemperatures at Madrid and summertemperatures at Lisbon show a significant upward trend from 1925-1950 to 1950-1975. It is not clear, whether this warming could possibly be due to the effect of city growth.

5. LEVEL AND VARIABILITY

As referred to in the introduction, studies on the relation between temperature level and year-to-year variability up till now did not arrive at definite answers. Van Loon and Williams for instance came to the conclusion that "As for the postulated coincidence of cold periods and high variability of climatic elements, our investigation does not provide evidence for such a connection". In a critical comment on Van Loon and Williams' work Duchon and Koscielny (1977) came to the conclusion that this negative result might be due to the lack of sophistication in the statistical analysis used.

Table 2. As table 1, but from 1925-1950 to 1950-1975.

	Winter	Spring	Summer	Fall
Stockholm				
Copenhagen				
Edinburgh				
Valentia (Irl.)	-		-	
C. England			-	
De Bilt			-	
Brussels				
Frankfurt				
Warsaw				
Geneva				
Milan				
Madrid	+			
Lisbon			+	

On the basis of our results we tend to agree with this criticism. Using Kendall's rank correlation method we were able to show, that the supposed negative correlation (cold periods - high variability; warm periods - low variability) is real, but only for winter temperature at middle and high latitudes. See table 3.

Table 3. Kendall rank correlation coefficients between mean temperature and interannual variability. Period: 1875-1975.

$|\tau| \geq 0.19$ is significant at the 5% level (*)

$|\tau| \geq 0.25$ is significant at the 1% level (**)

	Winter	Spring	Summer	Fall
Stockholm	- 0.22 *	- 0.05	0.01	0.04
Copenhagen	- 0.21 *	- 0.07	- 0.07	0.07
Edinburgh	- 0.21 *	0.04	- 0.03	- 0.03
Valentia(Irl.)	- 0.26 **	0.22*	0.11	- 0.05
C. England	- 0.23 *	0.04	- 0.02	- 0.13
De Bilt	- 0.27 **	- 0.06	- 0.12	- 0.17
Brussels	- 0.21 *	- 0.06	- 0.02	- 0.08
Frankfurt	- 0.29 **	- 0.05	0.03	- 0.07
Warsaw	- 0.30 **	- 0.00	- 0.11	- 0.10
Geneva	- 0.16	- 0.08	0.15	- 0.02
Milan	- 0.17	0.04	0.18	0.03
Madrid	0.08	0.07	- 0.10	0.01
Lisbon	0.01	0.03	- 0.11	0.11

For none of the other seasons a statistically significant relation between temperature level and variability could be established. The positive relation in spring at Valentia (Ireland) is not supported by other stations. We neglect it here.

In winter the negative correlation is due to a skewness of the temperature distribution in the sense that the median is greater than the mean (mild winters are more numerous than cold ones), which in North-Western Europe is caused by frequently occuring westerly type winters with a definite temperature maximum determined by ocean temperatures and less frequent continental type winters with rather unlimitedly low temperatures. In view of this probable explanation the negative correlation found here must not be given a global validity.

6. SPATIAL AND SEASONAL COHERENCE

Atmospheric variables defining the state of the atmosphere like pressure and temperature are strongly correlated in space as well as in time. This is also the case for time averages of such variables, like for instance daily or monthly means. In this section we make a comparison between the spatial and temporal correlation characteristics of temperature level at the one hand and temperature variability at the other. As before we make use of the seasonal averages of temperature and their variability from year to year. For the data available we computed the Kendall rank correlation coefficient for: (1) all pairs of stations and for each season; (2) all pairs of seasons and for each station. Again the period considered was 1875-1975.

From the wealth of output data a number of well-known facts could be re-established. For instance, that seasonal temperature level is significantly correlated over distances comparable to the scale of Western Europe. New or poorly known might, however, be the following conclusions:
(a) The spatial coherence of temperature variability is much lower than that of temperature level. As shown in Fig. 3 a statistically significant (5% level) correlation is limited to an area of radius 1000 km.

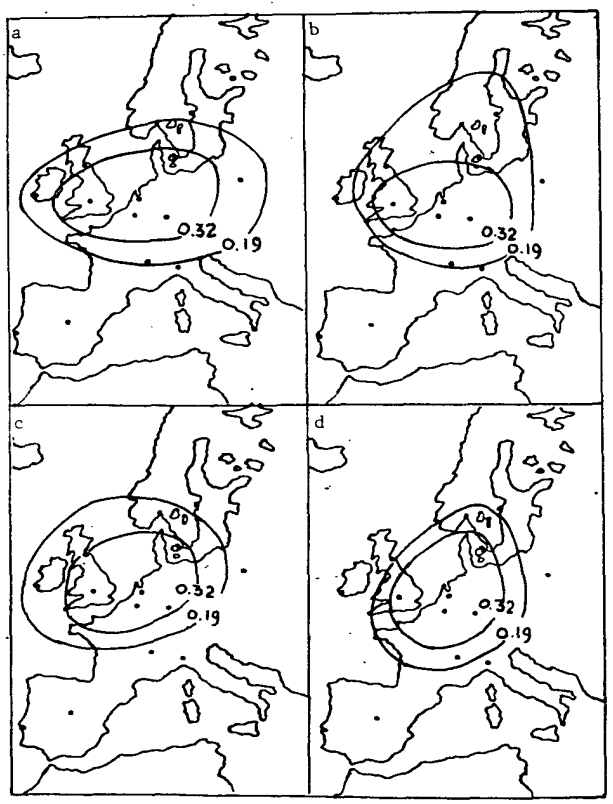


Fig. 3

Fig. 3. Kendall rank correlation coefficients between interannual variability of seasonal mean temperatures at De Bilt (The Netherlands) and that at other stations in Europe, for winter (a), spring (b), summer (c) and autumn (d).

(b) The spatial coherence of temperature as well as that of temperature variability increase, with latitude. For variability this is shown geographically in Fig. 4.

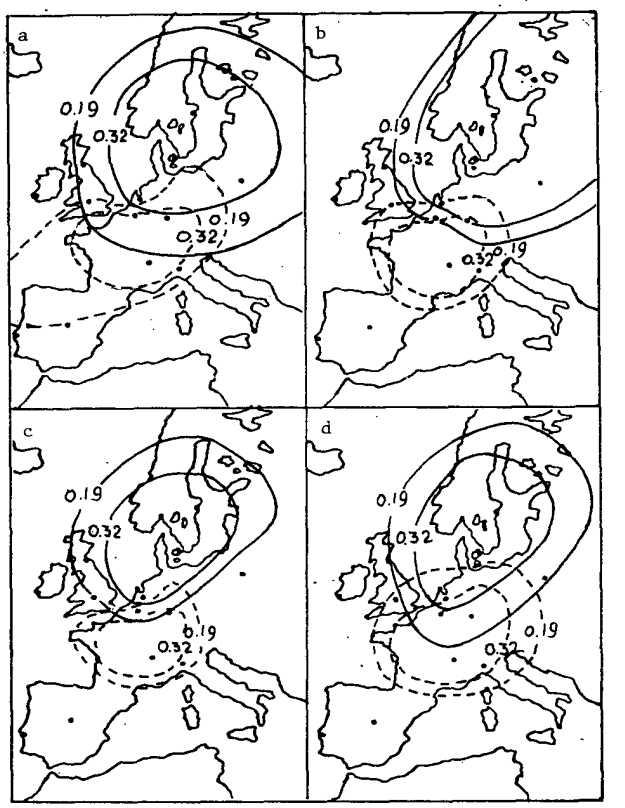


Fig. 4

Fig. 4. The same as Fig. 3, but for the reference stations Geneva (dashed lines) and Copenhagen (full lines).

A full explanation of the observed facts expressed in conclusions (a) and (b) is not attempted here. Rather, we have investigated the possible influence of the trend in our data on the above conclusions. Since in the last 50 years of our data record (1925-1975) seasonal temperature level was found to be trend free at practically all stations, we compared our results for the whole period with those for the last half of the period. Although some influence of the trend could be easily detected, the conclusions (a) and (b) above appeared to be firmly established.

Our results for the temporal, season-to-season correlations of temperature level were found to be very strongly affected by the trend. It is to be expected, that some of the results on seasonal persistence of temperature, published in the past, were in fact also unreliable, due to this effect.

Confirming ourselves to the trend free part of the record, we came to the conclusion that a statistically significant season-to-season correlation of temperature is a rare phenomenon. As shown by Table 4, the only apparently real phenomenon is the winter-to-spring correlation of temperature at stations around the Baltic Sea (which

most probably has to do with the occurrence or non-occurrence of sea ice and snow in that area).

Table 4. Season-to-season coherence of seasonal mean temperature for the period 1925-1975.

$|\tau| \geq 0.28$ is significant at the 5% level (*)

$|\tau| \geq 0.37$ is significant at the 1% level (**)

$|\tau| \geq 0.47$ is significant at the 1‰ level (***)

	Wi-Sp	Sp-Su	Su-Fa	Fa-Wi
Stockholm	0.37 **	0.24	0.01	0.16
Copenhagen	0.51 ***	0.22	0.12	0.06
Edinburgh	0.31 *	0.03	0.05	- 0.13
Valentia(Irl.)	0.24	0.29*	0.12	- 0.10
C.England	0.11	0.16	0.14	- 0.19
De Bilt	0.20	0.18	0.21	- 0.09
Brussels	0.12	0.14	0.06	- 0.06
Frankfurt	0.18	0.25	0.24	- 0.03
Warsaw	0.33 *	0.05	- 0.09	- 0.05
Geneva	0.06	0.26	0.08	- 0.07
Milan	0.07	0.26	0.16	- 0.01
Madrid	0.03	0.28*	0.11	0.13
Lisbon	0.18	0.47***	0.13	0.22

A season-to-season coherence of temperature variability was found to be completely absent at all stations considered.

7. SUMMARY AND CONCLUSIONS

Our statistical study of seasonal mean temperatures in Europe reveals, that the strong upward trend from 1875-1925 to 1925-1975 in temperature level was not paralleled by a trend in interannual variability of seasonal temperatures. In fact temperature level and interannual variability appeared to be uncorrelated, except for the winter season at middle to high latitudes. Here a negative correlation was firmly established, which means, that periods of on average colder winters are more variable from year to year, while warm periods show less variability. The statistical analysis was based on some well-known non-parametric methods, which may guarantee, that the results are not fortuitous.

This perhaps may put to rest (at least for Europe) the debate on the increase of temperature variability in the last decades and the implicitly assumed generally negative correlation between temperature level and variability.

Our analysis furthermore does not support the idea of a strong temperature influence of the increasing atmospheric CO_2-concentration during the foregoing century. Not only the relatively small warming in the winter season as compared to the other seasons, being in contradiction to the theoretically expected effect, but also the absence of an upward trend of temperature during the last 50 years (1925-1975), does make it very difficult to accept the CO_2-effect as a major factor in the determination of the European climate in the last 100 years.

Finally from our analysis we were able to conclude that temperature variability is a much more locally determined quantity than temperature itself. The geographical coherence of this quantity is relatively small, while a persistence from season to season (at some stations clearly present for temperature level) is non-existent.

8. REFERENCES

Duchon, C.E. and A.J. Koscielny, 1979: Comments on on "The association between mean temperature and interannual variability, Mon.Wea.Rev., 107, 496.

Hájek, J., 1969: Non-parametric statistics, Holden-Day, Inc.

Hare, F.K., 1979: Climatic variation and variability. Proc, World Climate Conf., WMO-No. 537, Geneva, 51-87.

Kellogg, W.W. and R.D. Bojkov (ed.), 1982: Report of the JSC/CAS meeting of experts on detection of possible climate change, WCP-29, Geneva, 43 pp.

Kendall, M.G., 1962: Rank Correlation Methods, Griffin & Comp., 3th ed.

Lamb, H.H., 1972: Climate: Present, Past and Future, Methuen, 613 pp.

Manabe, S. and R.J. Stouffer, 1980: Sensitivity of a global climate model to an increase of CO_2 concentration in the atmosphere, J. Geophys. Res., 85, 5529-5554.

Van Loon, H. and J. Williams, 1978: The association between mean temperature and interannual variability, Mon.Wea.Rev., 106, 1012-1017.

THE 500 MB GEOPOTENTIAL HEIGHT AS A FIRST-ORDER MARKOV PROCESS

Juhani Rinne and Annakaisa Sarkanen

Department of Meteorology
University of Helsinki
Hallituskatu 11-13
SF-00100,Helsinki 10,Finland

1. INTRODUCTION

In his study of large-scale motions Willson (1975) mentions that the spectra of the 500 mb geopotential height cannot be approximated by the spectrum of the first-order Markov process. Willson computed spectra for many wave numbers and latitudes north of 20°N in the Northern Hemisphere. The spectra are flattish or peaked at low frequencies, but they fall rapidly above a cutoff frequency. Conversely, Gambo (1979) found a good fit with the Markov spectrum at latitude 50°N. However, Gambo restricted his study to a narrow frequency interval where the spectrum is similar to that given by Willson. Thus there is no disagreement. Lau (1981) has presented a case where the spectrum of an empirical orthogonal function (EOF) is given by a red noise spectrum. Rogers (1979,p.131) mentions the red noise form of EOF spectra. Paegle and Haslam (1982) have fitted red noise spectra in the case of some local EOF's.

In the present study the 500 mb geopotential height is described with the first-order linear Markov process. In order to do this the height fields were represented with empirical orthogonal functions.

2. THE FIRST-ORDER MARKOV PROCESS

A variable z (here the height of the 500 mb surface) can be represented with the empirical orthogonal functions by

$$z(i,t)=zm(i)+\sum_{\nu=1}^{N} C_\nu(t)\,g_\nu(i) \qquad (1)$$
$$+\text{Residual}(t,i,N),$$

where zm is a long-term mean field, C_ν is the time-dependent coefficient and g_ν is the space function of the νth component. Variables i and t refer to the gridpoint and time, respectively, and N is the truncation point of the EOF series. The functions g_ν are orthogonal and are normalized by applying an area weighting (Rinne et al.,1981).

When functions g_ν are known the time-dependent coefficients can be determined from

$$C_\nu(t)=\sum_i g_\nu(i)\left\{z(i,t)-zm(i)\right\}dA(i), \qquad (2)$$

where dA(i) is the area element. In the dependent sample the coefficients are orthogonal. The functions and coefficients are arranged according to their variance reduction.

A spectral forecast model of the 500 mb height can be given by

$$\frac{\partial C_\nu}{\partial t} = \alpha_\nu + \sum_{\mu}^{N} \beta_{\mu\nu} C_\mu + \sum_{\varkappa}^{N}\sum_{\eta}^{N} \gamma^{\nu}_{\varkappa\eta} C_\varkappa C_\eta \qquad (3)$$

Here α, β and γ are model coefficients and C_νs are the EOF coefficients (for the model, see Rinne and Karhila, 1975). Eq.3 can be simplified to

$$\frac{\partial C_\nu}{\partial t} = 2\pi\lambda_\nu C_\nu(t) + \varepsilon_\nu, \qquad (4)$$

where ε_ν stands for a sum of some 15000 terms when N=175.

If ε_ν has a white noise spectrum, the power spectrum of C_ν is given by

$$S_\nu(f) = A_\nu/(f^2+\lambda_\nu^2). \qquad (5)$$

A_ν is directly related to the random forcing, i.e. A_ν is related to the variance of ε_ν. The red noise spectrum given by Eq.5 corresponds to the first-order Markov chain (Eq.4) (cf. Hasselmann,1976).

3. THE DATA AND METHOD

The main data consist of 14672 FNWC (Fleet Numerical Weather Center) analyses of the 500 mb height for the 25-year period 1946-70. The mean field (zm) and the functions (g_ν) are computed with 11876 analyses. The data are given in the FNWC grid, a total of 1404 points north of 20°N. The coefficients are also computed for 1966-79 using the very same functions and NMC (National Meteorological Center) analyses. The NMC data consist of 10017 analyses.

The yearly total variance of the 11876 analyses mentioned is 25175 m^2. The leading components up to N=1,10,100 explain 57.7, 79,0 and 98.7 per cent of the total variance, respectively.

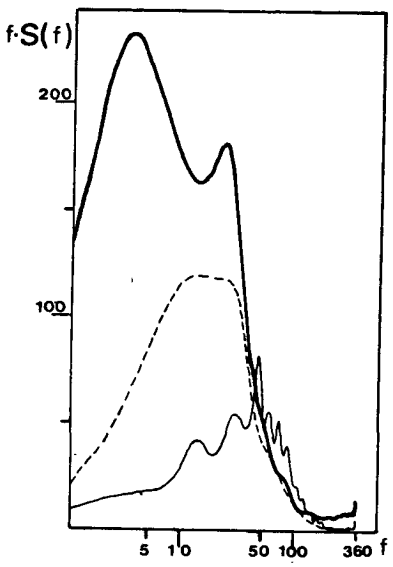


FIG.1 The observed spectra of the components ν =2 (thick line), 10 (dashed line) and 20 (thin line) for 1966-79. Units: f cycles per year, S(f) m^2/cpy.

The time series of coefficients $C_\nu(t)$ are divided into three parts that partly overlap: 1955-62, 1962-70, 1966-79. A comparison of the NMC - FNWC analyses for the overlapping period showed disagreement and so they have been investigated separately. Data values are given twice a day, at 00 and 12 GMT. The missing coefficients are interpolated linearly.

Excepting the first component no prewhitening or detrending was applied. In the case of the first component it was necessary to remove the dominant annual and semiannual cycles. The residual variation is weak and will not be studied here.

The spectra of the coefficients ν =2-30,35,40,...,175 are computed with the maximum entropy method, applying the Akaike final prediction error criterion (Ulrych and Bishop,1975). The critical point of the method is the determination of the parameter that gives the length of the so called prediction error filter. In

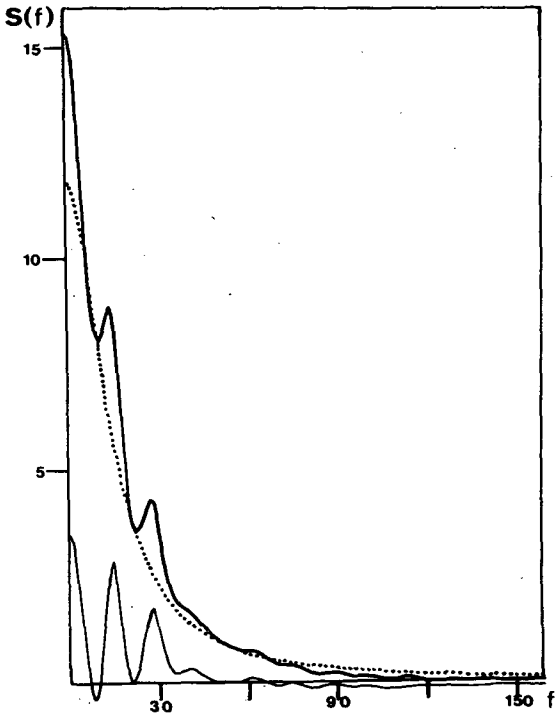


FIG.2 Illustration of the fitting procedure. Given are the true spectrum of the 12th component (thick line), the estimation of the background spectrum (dotted line), the difference of the former spectra (thin line). Sample 1962-70. Note that the coordinate system is different from the one in Fig.1. Units: f cycles per year, S(f) m^2/cpy.

the present case we are dealing with background spectra which are less sensitive to variations in the parameters. Therefore the method will be discussed in more detail in another paper presenting the peaks. Some examples of typical spectra are shown in Fig.1.

4. THE BACKGROUND SPECTRA

A typical spectrum of a coefficient C_ν includes a background spectrum on which relatively weak peaks are superimposed. The observed background spectra fit rather well with Eq.5 in the frequency interval 5-200 cpy (cycles per year). The fitting is performed subjectively. Because $S_\nu(f)$ represents the background spectrum, its values should not exceed the observed ones. In the subjective fitting, minor deviations from this constraint were allowed, especially at higher frequencies (Fig.2).

The values of the parameters λ and A are shown in Figs.3 and 4. The parameters seem to follow a general law that can be given by

$$\lambda_\nu = a\nu^b$$
$$A_\nu = \begin{cases} c\cdot 20^{-d}, & 1<\nu\leq 20 \\ c\cdot \nu^{-d}, & 20\leq\nu . \end{cases} \qquad (6)$$

The approximations are only valid for $2<\nu<175$, because components $\nu>175$ were

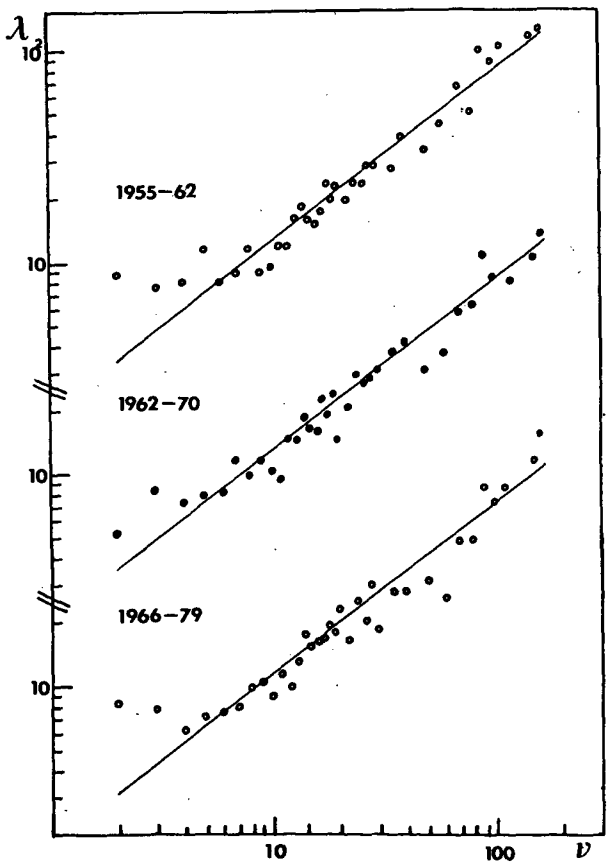


FIG.3 Parameter λ vs. index ν in three samples. Units are in cycles per year. Free hand regression lines are shown.

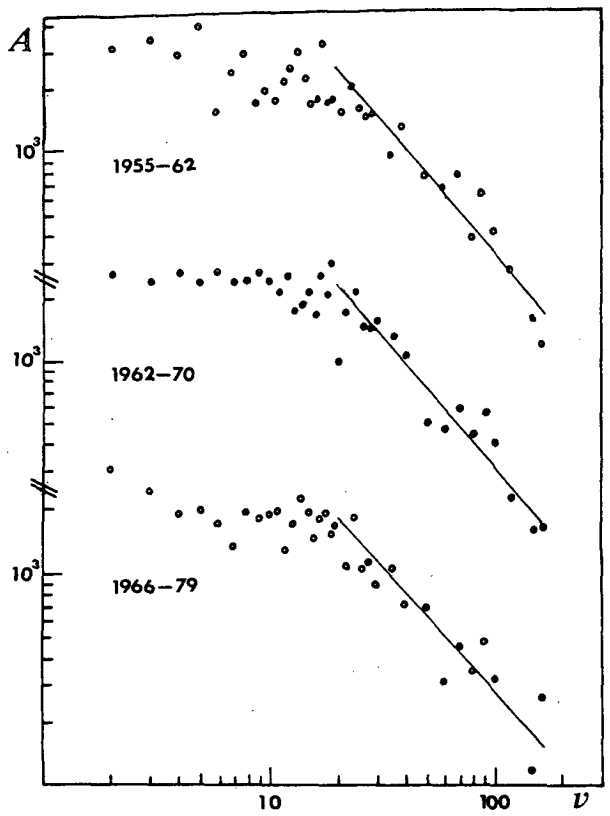


FIG.4 Parameter A vs. index ν in three samples. Units are in m^2cpy. Free hand regression lines are shown.

not studied. The approximation of λ_ν is less satisfactory in the cases ν = 2, 3 and 4. The first component mainly describes the yearly oscillation and therefore it is excluded from Eq.7. The values of coefficients a,b,c and d as estimated with a free hand regression are presented below

sample	a	b	c	d
	(cpy)		(m^2)	
1955-62	2.0	0.79	118000	1.3
1962-70	2.0	0.81	100000	1.25
1966-79	1.8	0.81	66000	1.2

The values vary from sample to sample. The large variation in the forcing noise coefficient c, in particular, is not expected. However, when f~0 the main factor in Eq.5 is A_ν/λ_ν^2, which is related to c/a^2. The values of this parameter in different samples are closer to each other than the values of c. Further, when f is large, the main factor in Eq.5 is A_ν or $c\nu^{-d}$. In this parameter the variations in c and d partly cancel each other out, as seen from experimenting with the values of different samples. Thus the variation in the parameter values does not imply a similar variation in $S_\nu(f)$. Some variation is obviously due to the fitting process and to the differences between the samples. These, in turn, are due to observation and analysis methods and to the climatological variation. The effect of the last factor on the parameters is obviously overwhelmed by the others and cannot be be determined.

5. SOME APPLICATIONS

Integrating Eq.5 with respect to f over the interval 0-∞ one finds for the variance of C

$$\overline{C_\nu^2(t)} = (\pi/2)(A_\nu/\lambda_\nu) \qquad (7)$$

or, substituting from Eq.6,

$$\overline{C_\nu^2(t)} = (\pi/2)(c/a)\nu^{b-d}(20/\nu)^d, \quad 1<\nu\leq 20 \qquad (8)$$
$$= (\pi/2)(c/a)\nu^{b-d}, \quad 175>\nu\geq 20$$

The observed and derived (Eq.8) values of variances in sample 1966-79 are compared in Fig.5. The latter ones are smaller, because the peak mass and the extra mass at low frequencies (f<5cpy) are not present in the background spectra.

Because $z-zm=\sum_\nu C_\nu(t)g_\nu(i)$, the Fourier transform of z-zm is $F(z-zm)=\sum_\nu F(C_\nu)g_\nu(i)$ in grid point i. The power spectrum is then

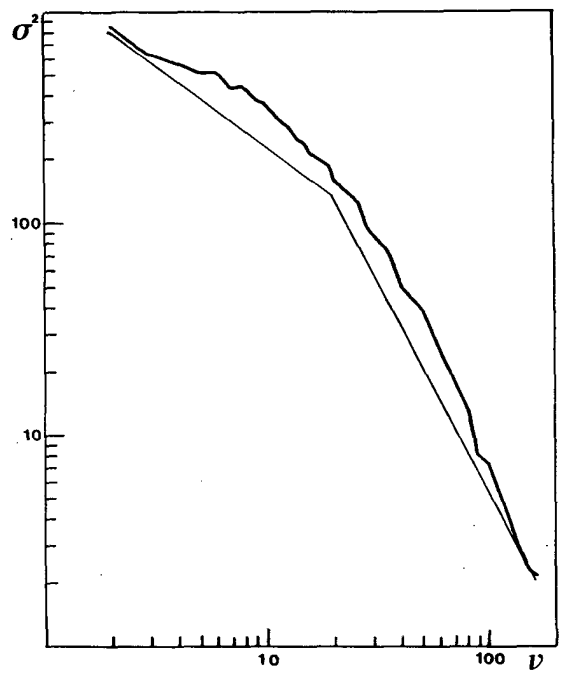


FIG.5 The observed and theoretical variances of the components ν . Units: m^2 . The theoretical values are derived from Eqs.6 and 8.

$$F(z-zm)F^*(z-zm)=\sum_{\nu}\sum_{\mu} F(C_\nu)F^*(C_\mu)g_\nu(i)g_\mu(i).$$

When this is integrated over the area north of 20°N, the products with $\nu \neq \mu$ on the right-hand side vanish owing to the orthogonality of the functions g. The left-hand side gives the average power spectrum north of 20°N. The spectrum is thus

$$S_{\bar{z}}(f) = \sum_{\nu} S_\nu(f).$$

Hence the average spectrum of the 500 mb geopotential height is

$$S_{\bar{z}}(f) = \sum_{\nu} \frac{A_\nu}{f^2 + \lambda_\nu^2}, \qquad (9)$$

where A_ν and λ_ν are estimated or taken from Eq.8. The result is an underestimation because the peaks and the extra mass at f<5cpy are excluded.

Spectrum (5) corresponds to the autocorrelation function $e^{-\gamma|\tau|}$ where τ is the time lag and $\gamma = 2\pi\lambda$ (Jenkins and Watts, 1968). Following Leith (1973), we take

$$T_o = \int_{-\infty}^{+\infty} e^{-\gamma|\tau|}\, d\tau$$

for the characteristic time between effectively independent sample values. Thus $T_o = 2/(2\pi\lambda)$. The characteristic times of the components ν=2,3 and 4 are 13-23 days when we use the values λ_ν=5-9 cpy from Fig.4.

6. SUMMARY

The index of the empirical orthogonal functions gives approximations via Eqs.5 and 6 of the following:

- The red-noise or background spectrum of the components in the frequency interval 5-200 cycles per year.
- The variance of the component
- The characteristic time of the component

The component number seems to be as essential to EOF's as the wave numbers are to spherical harmonics.

The spectrum of the 500 mb geopotential height can be computed as the sum of the spectrum of EOF's.

The characteristic times indicate that the components with a low index are predictable over a few weeks.

The conclusions are valid only at the accuracy of the fitting process.

ACKNOWLEDGEMENTS: This work has been financially supported by the Academy of Finland.

7. REFERENCES

Gambo, K.,1979: A note on the spectral density of vorticity of ultra-long waves in middle latitudes at 500-mb level in the winter season. J.Meteor.Soc. Japan, 57, 373-385.

Hasselmann, K.,1976: Stochastic climate models. Part I. Theory. Tellus, 28, 473-485.

Jenkins, G.M. and D.G. Watts,1968: Spectral analysis and its applications. Holden-Day, 524 pp.

Lau, N.C.,1981: A diagnostic study of recurrent meteorological anomalies appearing in a 15-year simulation with a GFDL general circulation model. Mon. Wea. Rev., 109, 2287-2311.

Leith, C.E.,1973: The standard error of time-averaged estimates of climatic means. J.Appl.Meteor., 12, 1066-1069.

Paegle, J.N. and R.B. Haslam,1982: Empirical orthogonal function estimates of local predictability. J.Appl. Meteor., 21, 117-126.

Rinne, J. and V. Karhila,1975: A spectral barotropic model in horizontal empirical orthogonal functions. Quart.J.Roy. Meteor. Soc., 101, 365-382.

Rinne, J.,V.Karhila and S.Järvenoja,1981: The EOF's of the 500 mb height in the extratropics of the Northern Hemisphere. Report No.17, Dept. of meteorology, University of Helsinki, Finland, 16 pp. + Appendix.

Rogers, J.C.,1979: The North Pacific Oscillation and eigenvectors of of Northern Hemisphere atmospheric circulation during winter. NCAR Cooperative thesis No. 56, 177 pp.

Ulrych, T.J. and T.N. Bishop,1975: Maximum entropy spectral analysis and autoregressive decomposition. Rev. Geophysics and Space Physics, 13,183-200.

Willson, M.A.G.,1975: A wavenumber-frequency analysis of large-scale tropospheric motions in the extratropical Northern Hemisphere. J.Atmos.Sci., 32, 478-488.

WHY DO WE GET SPURIOUS CORRELATIONS IN CLIMATOLOGY: SOME EXPLANATIONS

Ian T. Jolliffe

Mathematical Institute,
University of Kent at Canterbury,
CT2 7NF, U.K.

1. INTRODUCTION

The appearance of large correlations, both positive and negative, between climatological variables at various time lags, is a frequent occurrence. All too often, however, such correlations disappear, or even reverse their signs, when a different set of data is collected on the same variables. Bell (1976, 1977) gives several instances of this type of behaviour; for example, the correlation between sunspot numbers and the intensity of the north-east monsoon in January in the Far East varies from 0.6 early in this century, to - 0.5 in the 1940s, based on fifteen-year periods of data.

The aim of the present paper is to put forward various explanations for such 'nonsense' correlations Two explanations which are straightforward and well known are dealt with briefly in Section 2 while Sections 3 and 4 cover somewhat more subtle explanations at greater length. Section 3 deals with the result that the distribution of the correlation coefficient, R, is far from normal (Gaussian), when the sample size, n, is small. In fact, the distribution of R is U-shaped for very small sample sizes, with values of R close to ± 1 more likely than values near zero, even when the two variables of interest are independent. Although this property helps to explain at least some of the substantially non-zero correlations which are observed for small sample sizes, it turns out that use of a simple normal approximation to the distribution of R will produce too few, rather than too many, significant correlations.

The U-shape of the distribution of R can persist for much larger sample sizes if the variables of interest have a high degree of autocorrelation, as is common in climatological series. Furthermore, the variance of R is inflated in the presence of autocorrelation, and this also contributes to spuriously 'significant' values of R, if independence is assumed. The behaviour of R when the variables are autocorrelated is discussed further in Section 4. In particular, the case of two quasi-periodic series is examined in some detail. It is argued that such series can often be modelled satisfactorily by second-order autoregressive processes, and that such processes may have correlations between them which switch from large positive to large negative values.

The dangers of spurious correlation are as great for economic data as they are in climatology, and some work has been reported in the economic literature on how such correlations can arise. Some of this work, which is related to the material of Section 4, is reviewed briefly in Section 5.

2. SOME STRAIGHTFORWARD EXPLANATIONS FOR SPURIOUS CORRELATION

The causes of spurious correlations which are discussed in this section are very well-known, and perhaps obvious, but are included for completeness.

The first problem is that in many climatological studies a large number of variables are examined. If an even larger number of sample correlations between the variables are then calculated, and compared with an appropriate critical level of the distribution of the sample correlation coefficient, then some will be 'significant' purely by chance. For example, if 1000 correlations are calculated independently, then even if the test is done at a fairly stringent 1% level, we would still expect to get about 10 falsely significant results. This fact is widely appreciated and allowed for when 1000 correlations are actually calculated. However, it is easy to erroneously ignore the problem if only a few correlations are calculated, but these correlations have been deliberately selected by looking at a large amount of data first and calculating correlations only for the relationships that look 'interesting'.

A second straightforward cause of spurious correlation, which is more common in economics than in climatology, is that the two variables measured over time both have a steady upward, or downward, trend. The trend is determined by extraneous factors, and although it may cause a large correlation between the variables, it does not imply any direct relationship. A similar

argument holds in the case where both variables display, instead of trend, regular cyclic behaviour with similar period and phase for each variable throughout the data set. Some aspects of cyclic behaviour will be discussed further in Section 4.

3. THE DISTRIBUTION OF THE SAMPLE CORRELATION COEFFICIENT, R.

A third fairly straightforward reason for spurious correlation might be that an inappropriate test of significance is done, due to the use of an incorrect approximation to the distribution of the sample correlation coefficient R.

Suppose that X and Y are independent normally distributed random variables, and that a random sample of n pairs of observations is taken on X, Y, namely $(X_1,Y_1),(X_2,Y_2)\ldots(X_n,Y_n)$. Then the correlation coefficient, R, computed from $(X_1,Y_1),(X_2,Y_2)\ldots(X_n,Y_n)$ has a distribution with probability density function

$$f(r) = \frac{1}{B\{\frac{1}{2},\frac{1}{2}(n-2)\}}(1-r^2)^{\frac{1}{2}(n-4)} \quad 0\le r\le 1 \quad (3.1)$$

The mean and variance of this distribution are zero and $(n-1)^{-1}$ respectively.

To test whether or not an observed correlation coefficient is significantly different from zero, the distribution given by (3.1) may be used directly or, equivalently, the statistic

$$t = \{(n-2)\ R^2/(1-R^2)\}^{\frac{1}{2}}, \quad (3.2)$$

which has a t-distribution with (n-2) degrees of freedom when X and Y are uncorrelated, can be computed. There are also two approximate results which are available for testing whether X and Y are uncorrelated, namely that

$$Z = \tfrac{1}{2}\log\left(\frac{1+R}{1-R}\right) \underset{\sim}{\text{approx}} N(0,(n-3)^{-1}) \quad (3.3)$$

and that $R \underset{\sim}{\text{approx.}} N(0,(n-1)^{-1})$, (3.4)

when X and Y are uncorrelated.

The last approximation is the simplest, and it is therefore tempting to use it. However, it is not such a good approximation as (3.3), and, of course, the exact results based on (3.1) and (3.2) are even better. Table 1 compares the approximation given by (3.4) with the exact distribution of (3.1), at the 5% and 1% critical levels for $|R|$.

It can be seen that the simplest approximation (3.4) is reasonable, for n as small as 30 at the 5% level, although the approximation is worse at the 1% level. It also demonstrates that very large values of R are necessary for $n \le 10$ in order to declare a significant result; a value of $|R| = 0.9$ looks impressive, but is not particularly unlikely if n=5. Furthermore, the approximation is conservative, in that it will favour acceptance of no correlation between

Table 1. 5% and 1% critical levels for $|R|$, for the exact distribution (3.1), and for the approximation (3.4).

Sample size, n	5% Exact	5% Approximate	1% Exact	1% Approximate
3	0.997	>1	0.9999	>1
4	0.950	>1	0.990	>1
5	0.878	0.980	0.954	>1
10	0.632	0.653	0.765	0.859
20	0.444	0.450	0.561	0.591
30	0.361	0.364	0.463	0.479
42	0.304	0.306	0.393	0.402
62	0.250	0.251	0.325	0.330

X and Y more frequently than an exact test. Use of the approximation will therefore not cause spurious correlations, provided that two assumptions hold:

(i) X and Y are normally distributed;

(ii) $(X_1,Y_1), (X_2,Y_2), \ldots, (X_n,Y_n)$ form a random sample, i.e. (X_i,Y_i), (X_j,Y_j) are independent for $i \ne j$.

However, if either of these assumptions is relaxed, the validity of each of (3.1) - (3.4) becomes questionable. The effect of non-normality will depend on what particular type of non-normal distributions is possessed by X and Y and will not be discussed further here. Correlation between different values of X and/or Y will now be considered further and will be shown to be a major potential cuase of spurious correlation.

It is common for successive observations on a climatological variable to be correlated; and if such autocorrelation is present in X and Y, then the distribution of the correlation coefficient R can be very different from that given by (3.1).

One of the first authors to recognise this problem was Yule (1926) in his presidential address to Royal Statistical Society. By means of simulation he showed that certain types of autocorrelation lead to markedly U-shaped distributions for R. For random samples, (3.1) is U-shaped for $n < 4$, flat for n=4, and unimodal for $n > 4$, but the undesirable U-shaped behaviour persists for larger values of n when autocorrelation is present in X and Y. Yule's (1926) results were (understandably) for a small sample size, n=10, but if strong autocorrelation is present the U-shape may persist for much larger sample sizes, which means that the true (U-shaped) distribution looks nothing like the normal approximation, nor, indeed, like the exact (unimodal) distribution. Even if n is large enough, or the autocorrelation weak enough, for the true distribution of R to be unimodal, the use of the normal approximation (3.4) will lead to spuriously significant values of R. This will occur because the variance used in the

approximation may be a gross underestimation of the true value. The variance of the correlation between two autocorrelated but independent data sets is approximately (Bartlett, 1955, p.289, who actually uses n, rather than n-1, though the difference will be small for large n)

$$(n-1)^{-1} \sum_{j=-\infty}^{\infty} \rho_X(j)\, \rho_Y(j) , \qquad (3.5)$$

where $\rho_X(j)$, $\rho_Y(j)$, $j = 0, \pm 1, \pm 2, \ldots$ are autocorrelations for X and Y. If X and Y have a finite number of non-zero autocorrelations, say $\rho_X(1), \rho_X(2), \ldots, \rho_X(q_1)$, $\rho_Y(1), \rho_Y(2), \ldots, \rho_Y(q_2)$, then (3.5) becomes

$$(n-1)^{-1} \left[1 + 2 \sum_{j=1}^{q} \rho_X(j)\, \rho_Y(j) \right],$$

where $q = \mathrm{Min}(q_1, q_2)$. Depending on the size and sign of these non-zero autocorrelations, $2 \sum_{j=1}^{q} \rho_X(j)\, \rho_Y(j)$ could be fairly large and positive, or close to zero. It is unlikely to be negative, since autocorrelations between successive observations are usually positive, unless the observations are taken a long time apart and there is some cyclic behaviour.

An interesting special case of (3.5) occurs when the variables X and Y are both first-order autoregressive [AR(1)] processes, so that

$$\begin{aligned} X_t &= \phi_X X_{t-1} + \varepsilon_t \\ Y_t &= \phi_Y Y_{t-1} + \eta_t \end{aligned} \qquad (3.6)$$

where X_t, Y_t denote the values of X and Y at time t, $\{\varepsilon_t\}, \{\eta_t\}$ are purely random (white noise) series, and ϕ_X, ϕ_Y are constants which equal $\rho_X(1)$, $\rho_Y(1)$ respectively. For such a model, $\rho_X(j) = \phi_X^j$,

$$\rho_Y(j) = \phi_Y^j \quad \text{for } j = 0, 1, 2, \ldots,$$

so that (3.5) becomes

$$(n-1)^{-1} \frac{(1 + \phi_X\phi_Y)}{(1 - \phi_X\phi_Y)} . \qquad (3.7)$$

Table 2 gives values of the quantity

$$f = \frac{(1 + \phi_X\phi_Y)}{(1 - \phi_X\phi_Y)} ,$$

which measures the factor by which the variance of R is inflated when autocorrelation is introduced, for various values of ϕ_X, ϕ_Y. The values show that for $\phi_X = \phi_Y = 0.7$, a degree of autocorrelation which is not uncommon when observations are taken fairly close together, the variance is increased by a factor of about 3; when $\phi_X = \phi_Y = 0.9$ there is nearly a ten-fold increase. Clearly, in such cases it is very misleading to approximate the variance of R by $(n-1)^{-1}$ and may result in many spuriously significant correlations.

Table 2. Values of $f = (1 + \phi_X\phi_Y)/(1 - \phi_X\phi_Y)$ for various values of ϕ_X, ϕ_Y.

		ϕ_X				
		0.1	0.3	0.5	0.7	0.9
ϕ_Y	0.1	1.02	1.06	1.11	1.15	1.20
	0.3		1.20	1.35	1.53	1.74
	0.5			1.67	2.08	2.63
	0.7				2.92	4.41
	0.9					9.53

Additional information regarding the shape of the distribution of R is also available from values such as those given in Table 2. The distribution is either U-shaped, unimodal or uniform; when it is unimodal, its variance is less than that of the uniform distribution, and when it is U-shaped its variance exceeds that of the uniform distribution. But the uniform distribution on the interval (-1,1) has variance 1/3, so that the distribution is U-shaped if Var R > 1/3. But Var R = f/(n-1), so Var R > 1/3 if f > (n-1)/3, so the distribution is U-shaped for $n \leq 20$ when $f > 6$, as occurs when $\phi_X = \phi_Y = 0.9$. For really strong autocorrelation, the distribution may be U-shaped for much larger n; for example, when $\phi_X = \phi_Y = 0.97$, the distribution is U-shaped for n as large as 100.

4. QUASI-PERIODIC BEHAVIOUR OF CLIMATOLOGICAL VARIABLES AND ITS IMPLICATION FOR SPURIOUS CORRELATIONS.

Quasi-periodic behaviour in climatological series is a widespread phenomenon, and series which exhibit such behaviour clearly have non-zero autocorrelations. If the cyclic behaviour has a fixed period, which is shared by two series, then there will generally be a correlation between the series which may be positive or negative depending on the phase shift between the two series. In practice, apart from annual and diurnal cycles, the behaviour of climatological series is more often quasi-periodic, with both the period and the amplitude of cycles varying over time. Quasi-periodic behaviour in two series can still lead to large correlations between them, provided that their periods are not too dissimilar. Furthermore, the correlations can switch between large positive and large negative values as the two series drift in and out of phase with each other. This switching between positive and negative values which is observed in practice (Bell 1976, 1977) cannot occur with two series possessing a consistent

upward or downward trend, nor with two cyclic series with the same fixed period.

Jolliffe (1983) has argued that quasi-periodic behaviour can often be explained by postulating second-order autoregressive [AR(2)] models for the series. In addition, such models have an intuitively plausible justification for many stationary series - see Jolliffe (1983). Such models have the form

$$(X_t-\mu) = \phi_1(X_{t-1}-\mu) + \phi_2(X_{t-2}-\mu) + \varepsilon_t , \quad (4.1)$$

where ϕ_1, ϕ_2 are constants, μ is the mean of the series, and $\{\varepsilon_t\}$ is a white noise series.

The autocorrelation structure of a series generated by (4.1) is somewhat similar to that of AR(1) models described in the previous section since autocorrelation can persist for many lags, but an advantage of AR(2) models is that they can adequately describe the common phenomenon of quasi-periodicity.

The explanation given above of how quasi-periodic behaviour can lead to large, switching, correlations is intuitive. To investigate its validity, AR(2) models were simulated with parameters chosen so that quasi-periodic behaviour would occur. The results of one such simulation are summarised in Table 3 and Fig. 1. The data in this case was all generated from AR(2) processes (4.1) with $\phi_1 = 1.51$, $\phi_2 = -0.81$; with these values of ϕ_1, ϕ_2, the series displays quasi-periodic behaviour with period 11. Various independent realisations, each of length 50, were generated from such a model, and the correlation calculated for each pair of independent realisations. Table 3 gives the distribution of the values of $|R|$ for 174 pairs of independent series, together with the expected distribution if the simple approximation (3.4) is used.

Table 3. Distribution of $|R|$ for 174 pairs of independent simulated AR(2) series of length 50 with parameters $\phi_1 = 1.51$ and $\phi_2 = -0.81$, together with the expected distribution using approximation (3.4)

Value of $\lvert R\rvert$:	0.0 - 0.1	0.1 - 0.2	0.2 - 0.3	0.3 - 0.4	0.4 - 0.5	0.5 - 0.6	0.6 - 0.7
Observed	37	49	29	24	16	12	7
Expected	90	56	22	5	1	0	0

The approximation (3.4) would be quite good if there were no autocorrelation present in the series, but it is clear that the actual distribution of $|R|$ is far from that given by the approximation, for the present data. For example, 36% of the observed values lie beyond the upper 5% level of the approximate distribution for $|R|$. The type of behaviour seen in Table 3 could well be a contributory factor in the example cited earlier and given by Bell (1977) since one of the series involved was for sunspot numbers, which are notable for their very strong quasi-periodic behaviour.

Another aspect of spurious correlation is that if a predictive relationship is sought between two series, then not only is the correlation calculated between X_t and Y_t when they are measured simultaneously, but correlation is also often calculated when X_t leads or lags Y_t by 1, 2, 3, ... time periods. The extra correlations increase the chance of spurious significance and this is demonstrated particularly well for quasi-periodic series. As the phase of one series is shifted relative to the other, so the correlation may change from large positive values when the series are in phase, to large negative values when they have opposite phase. Figure 1 illustrates this for two AR(2) series with the same parameters as before, but with realisations of length 200.

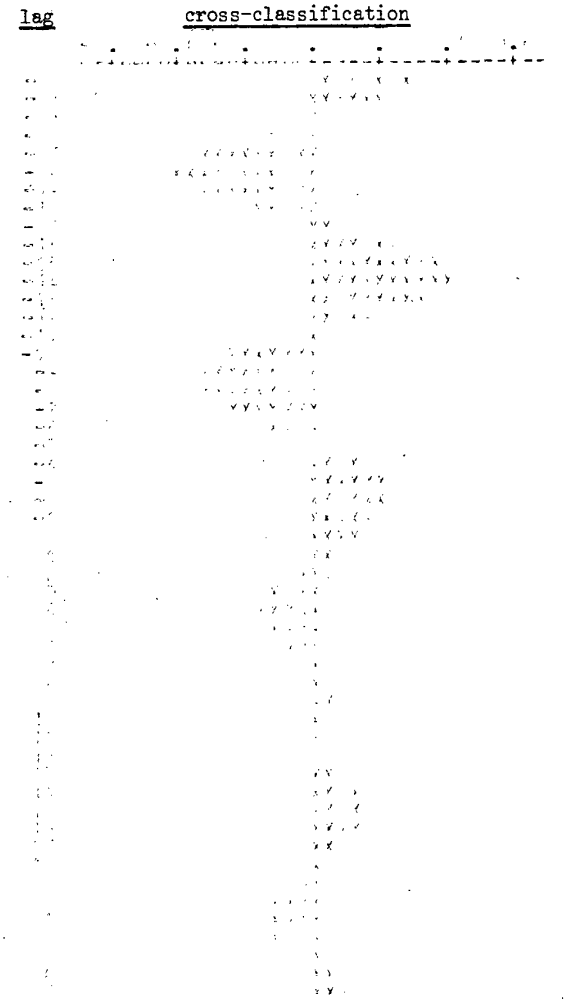


Fig.1(a) Cross-correlations between two independent simulated AR(2) series, each of length 200, and with parameters $\phi_1 = 1.51$, $\phi_2 = -0.81$.

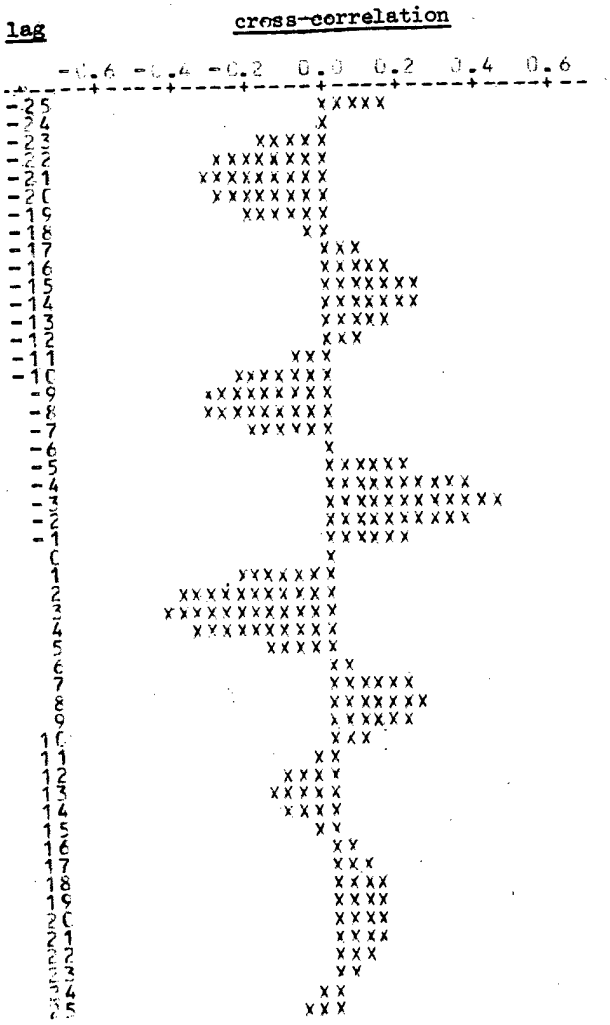


Figure 1(b) As Figure 1(a) but for a different pair of independent AR(2) series.

It is seen that with lags of between ± 25 time periods, the correlation ranges from -0.390 to 0.387 for the pair of series in Figure 1(a), and from -0.442 to 0.422 for those in Figure 1(b). This compares with the 5% significance level of 0.139 given by approximation (3.4). The behaviour of the cross-classification shown in Figure 1 is fairly typical of the simulations carried out so far. For some pairs of independent simulated series, the cross-correlation function is flatter, but there are usually a substantial number of cross-correlations beyond the usual significance levels, in both positive and negative directions.

5. RELATED WORK IN ECONOMICS

In economics the problem of spurious regression is common, as Granger and Newbold (1977) has stressed in Chapter 6.4 of their book. Linear models may be fitted in order to predict one economic variable, Y, from several explanatory variables, and often very good fits, as measured by the squared multiple correlation coefficient, R^2, are reported. In the special case where there is only one explanatory variable, X, the multiple correlation coefficient is simply R, the correlation coefficient between X and Y, so that the problem of spurious regression is equivalent to the problem of spurious correlation between X and Y. Granger and Newbold (1977) argue that many economic regression equations with impressively large values of R^2 are, in fact, spurious, with R^2 being inflated because of an inappropriately specified error structure for the regression equation. Independence between error terms is usually assumed when fitting regression equations, but this is inappropriate when autocorrelation is present in one or more of the variables in the regression equation. Thus, as discussed in Sections 3 and 4, it is autocorrelation that is mainly responsible for spuriously large values of R^2.

Autocorrelation in economic series is most frequently due to a trend (i.e. non-stationarity) in the series, and the implications of trend being ignored when fitting regression equations has been studied using simulation. In particular, Granger and Newbold (1977) report some results where two series were generated according to the model

$$X_t = X_{t-1} + \varepsilon_t + \theta_X \varepsilon_{t-1} \qquad (5.1)$$

$$Y_t = Y_{t-1} + \eta_t + \theta_Y \eta_{t-1} ,$$

where θ_X, θ_Y are constants and $\{\varepsilon_t\}, \{\eta_t\}$ are independent white noise series. These two series are independent, and both follow a IMA(1,1) model. This means that they are non-stationary, but differencing once will induce stationarity, and the first differences then follow a first-order moving-average model. IMA(1,1) models apparently 'provide a very good representation of a wide range of economic time series'. For a set of values of θ_X, θ_Y the percentage of occasions on which a t-test at the 5% level declared a 'significant' relationship between X and Y is recorded, along with other information, in Granger and Newbold's (1977) Table 6.4. X and Y are independent, so that the values should be close to 5%. All the values given exceed 5%, and they are generally much larger, between 40% and 70%, except when θ_X or θ_Y are large and negative. A second study reported by Granger and Newbold (1977, pp.205-206) suggests that it is mostly non-stationarity, rather than moving-average behaviour that is responsible for spurious regression/correlation. This is not surprising because non-stationarity will induce much longer-lasting autocorrelation than a moving-average model, and it is autocorrelation persisting for many lags which seems to be the main cause of spurious regression.

Non-stationarity is not a particularly prevalent problem in climatology, at least on short time scales. However, first- and second-order autoregressive models which often provide a reasonable fit to the major variation in climatological series also have autocorrelation at large time lags, and so display the same spurious correlations.

6. DISCUSSION

There are many reasons for the occurrence of spurious correlations, and most have been known for a considerable amount of time. Nevertheless many are ignored, especially those created by autocorrelation within the variables being correlated. It is hoped that the present paper will encourage more consideration of the possibility that observed large correlations may be due not to any direct relationship, but rather because of the correlation structure within each series. This possibility needs to be particularly borne in mind when dealing with quasi-periodic series.

7. REFERENCES

Bartlett, M.S. (1955) Stochastic Processes. Cambridge: Cambridge University Press.

Bell, G.J. (1976) Seasonal forecasts of Hong Kong summer rainfall. Weather, 31, 208-212.

Bell, G.J. (1977) Changes in sign of the relationship between sunspots and pressure, rainfall and the monsoons. Weather, 32, 26-32.

Granger, C.W.J. and Newbold, P. (1977) Forecasting economic time series. New York: Academic Press.

Jolliffe, I.T. (1983) Quasi-periodic meteorological series and second-order autoregressive processes. J.Climatology. To appear.

Yule, G.U. (1926) Why do we sometimes get nonsense-correlations between time series? - a study in sampling and the nature of time series. J.Roy.Statist.Soc., 89, 1-64.

TREND CORRECTION IN CLIMATOLOGICAL TIME SERIES OBTAINED BY HARD LIMITING PROCEDURE

Masanori. B. Okamoto and Kōsei Iwase

Section of Statistics
Department of Economics
Hiroshima University
Hiroshima City,730,Japan

Department of Applied Mathematics
Faculty of Engineering
Hiroshima University
Higashi-hiroshima City,724,Japan.

1. INTRODUCTION

An sythetic index series of sevral climatic time series are obtained by Okamoto and Iwase(1982). The index series is defined as the total sum of hard limiting time series which is produced by a crossing level suitably chosen. The time series treated there is assumed to stationary. However, many climatic time series have a trend in general. Although the above procedure should be applied to the time series after eliminated the trend, it happens that a time series has trend which cause a deviation of the crossing level from the predetermined level. In this paper, at first we prove two theorems about integral formula concerning with covariance function of clipped series. Next, we give a possible maximum error in covariance function due to this deviation, assuming that trend is linear.

2. COVARIANCE FUNCTION OF CLIPPED SERIES

Two random variable X and Y are assumed to follow bivariate normal population $N(0,0,1,1;\rho)$. Hard limiting procedure of X is defined by

$$C_u(X) = \begin{cases} 1 & x > u \\ 0 & |X| \leqq u \\ -1 & X < -u \end{cases}$$

where u is any positive number and means a crossing level. $C_u(Y)$ is defined similary. Under the condition $|\rho_u|<1$, Okamoto and Iwase(1982) gives the integral form of covariance function, i.e.,

$$E[C_u(X)C_u(Y)] = \frac{1}{\pi}\int_{-\rho}^{\rho} \frac{1}{\sqrt{1-x^2}} \exp[-\frac{u^2}{1+x}]dx .$$

We now prove that this integral expression holds still for the case $|\rho| = 1$. We consider the above formula;

$$\frac{1}{\pi}\int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} \exp[-\frac{u^2}{1+x}]dx$$

$$= \frac{1}{\pi}\int_{0}^{1} \frac{1}{\sqrt{1-x^2}} \exp[-\frac{u^2}{1+x}]\,dx + \frac{1}{\pi}\int_{0}^{1} \frac{1}{\sqrt{1-x^2}} \exp[-\frac{u^2}{1-x}]dx .$$

We denote the first term and the second term of the right hand of the above equation with A and B, respectively. Then we have with $u^2/(1+x) = t^2/2$,

$$A = \frac{2u}{\pi}\int_{u}^{\sqrt{2}} \frac{1}{t\sqrt{t^2-u^2}} \exp[-\frac{1}{2}t^2]\,dt$$

$$= \frac{2u}{\pi}\int_{u}^{\infty} \frac{1}{t\sqrt{t^2-u^2}} \exp[-\frac{1}{2}t^2]\,dt - \frac{2u}{\pi}\int_{\sqrt{2}}^{\infty} \frac{1}{t\sqrt{t^2-u^2}} \exp[-\frac{1}{2}t^2]\,dt$$

$$= A_2 - A_1 .$$

With transformation of variable $t/\sqrt{2} = x$ and $v = u/\sqrt{2}$, we have

$$A_1 = \frac{2u}{\pi}\frac{1}{\sqrt{2}}\int_{\sqrt{2}}^{\infty} \frac{1}{x\sqrt{x^2-u^2}} \exp[-x^2]\,dx .$$

By the use of formula (Gradshteyn and Ryzhik (1965), p.339)

$$\int_{\sqrt{2}u}^{\infty} \frac{1}{\sqrt{x^2-u^2}} e^{-x^2} \frac{dx}{x} = \frac{\pi}{4u}(1-\Phi(u))^2 , \quad u>0$$

where

$$\Phi(x) = \frac{2}{\sqrt{\pi}}\int_{0}^{x} \exp[-t^2]\,dt ,$$

we obtain

$$A_1 = \frac{1}{2}(1-\Phi(\frac{u}{\sqrt{2}}))^2 .$$

Putting $\sqrt{t^2-u^2} = x$, we have

$$A_2 = \frac{2u}{\pi}\exp[-\frac{1}{2}u^2]\int_{0}^{\infty} \frac{1}{x^2+u^2} \exp[-\frac{1}{2}x^2]\,dx .$$

By the use of formula (Gradshteyn and Ryzhik (1965), p.338)

$$\int_{0}^{\infty} \frac{1}{x^2+\beta^2} \exp[-\mu^2x^2]\,dx$$

$$= [1-\Phi(\beta\mu)]\frac{\pi}{2\beta}\exp(\beta^2\mu^2) , \quad \mathrm{Re}(\beta) > 0 ,$$

we have

$$A_2 = \frac{2u}{\pi}\exp[-\frac{1}{2}u^2](1-\Phi(\frac{u}{\sqrt{2}}))\frac{\pi}{2u}\exp[\frac{1}{2}u^2]$$

$$= 1 - \Phi(\frac{u}{\sqrt{2}})$$

Thus we obtain

$$A = A_2 - A_1$$

$$= [1-\Phi(\frac{u}{\sqrt{2}})] - [\frac{1}{2}(1-\Phi(\frac{u}{\sqrt{2}}))^2]$$

$$= \frac{1}{2} - \frac{1}{\pi}(\int_0^u \exp[-\frac{1}{2}x^2]\,dx)^2$$

Similarly we have

$$B = \frac{1}{2}(1-\Phi(\frac{u}{\sqrt{2}}))^2$$

Finally we obtain

$$E[C_u(X)C_u(Y)]\Big|_{|\rho|=1} = \frac{1}{\pi}\int_{-\rho}^{\rho}\frac{1}{\sqrt{1-x^2}}\exp[-\frac{u^2}{1+x}]dx\Big|_{\rho=1}$$

$$= 1 - 2\int_0^x \phi(x)dx = 2\int_x^\infty \phi(t)dt = E[1_u(x)]$$

where $\phi(x)$ is the standard normal densit function and

$$1_u(x) = \begin{cases} 1 & |X| > u \\ 0 & |X| \leq u \end{cases}$$

Noting

$$(C_u(x))^2 = 1_u(x),$$

we have for $\rho = 1$

$$E[C_u(x)C_u(Y)] = E[1_u(x)]$$

and for $\rho=-1$

$$E[C_u(x)C_u(Y)] = -E[1_u(X)]$$

Thus we can prove that the integral formula of covariance function of $C_u(X)$ and $C_u(Y)$ holds for $|\rho| = 1$, i.e.

[Theorem 1] For any ρ in $-1 \leq \rho \leq 1$,

$$E[C_u(X)C_u(Y)] = \frac{1}{\pi}\int_{-\rho}^{\rho}\frac{1}{\sqrt{1-x^2}}\exp[-\frac{u^2}{1+x}]dx.$$

This theorems also includes Kedem's corrollary (Kedem (1980), p.34) as a special case. By the use of Theorem 1, we obtain the following theorem.

[Theorem 2] For any ρ in $-1 \leq \rho \leq 1$ and for any positive δ,

$$|E[\mathrm{sgn}(X-\delta)\mathrm{sgn}(Y-\delta)] - E(\mathrm{sgn}X\ \mathrm{sgn}Y)| \leq 2\Phi(\frac{\delta}{\sqrt{2}})$$

where $\Phi(x) = \frac{2}{\sqrt{\pi}}\int_0^x \exp(-t^2)\,dt$. Equality in the above relation holds only for $\rho = -1$.

Noting $1-\exp(-\frac{\delta^2}{1+x})$ is monotone decreasing and positive for all x in $-1 \leq x \leq 1$, we have

$$|E[\mathrm{sgn}(X-\delta)\mathrm{sgn}(Y-\delta)] - E(\mathrm{sgn}X\ \mathrm{sgn}Y)]|$$

$$= |\frac{2}{\pi}\int_0^\rho \frac{1}{\sqrt{1-x^2}}\exp[-\frac{\delta^2}{1+x}]dx + \frac{2}{\pi}(\int_0^\delta \exp(-\frac{1}{2}x^2)dx)^2 - \frac{2}{\pi}\mathrm{Sin}^{-1}\rho|$$

$$= \frac{2}{\pi}|\int_0^\rho \frac{1}{\sqrt{1-x^2}}(1-\exp[-\frac{\delta^2}{1+x}])dx - (\int_0^\delta \exp[-\frac{1}{2}x^2]dx)^2|$$

$$\leq \frac{2}{\pi}(-1)\int_0^{-1}\frac{1}{\sqrt{1-x^2}}(1-\exp[-\frac{\delta^2}{1+x}])dx + \frac{2}{\pi}(\int_0^\delta \exp[-\frac{1}{2}x^2]dx)^2$$

$$= \frac{2}{\pi}\int_0^1 \frac{1}{\sqrt{1-x^2}}dx - \frac{2}{\pi}\int_0^1 \frac{1}{\sqrt{1-x^2}}\exp[-\frac{\delta^2}{1-x}]dx + \frac{2}{\pi}(\int_0^\delta \exp[-\frac{1}{2}x^2]dx)^2$$

$$= 1 - (1-\Phi(\frac{\delta}{\sqrt{2}}))^2 + (\Phi(\frac{\delta}{\sqrt{2}}))^2 = 2\Phi(\frac{\delta}{\sqrt{2}})$$

Thus we have proved the theorem 2.

3. POSSIBLE MAXIMUM ERROR FOR TREND

When original time series $x_T(t)$ for the duration of T has linear trend, i.e.,

$$x_T(t) = a_o + a_1 t + x(t),$$

sample mean over T is

$$\bar{x}_T = a_o + a_1\bar{t} + \bar{x}$$

Deviation from the mean, $z_T(t)$ can be given by

$$z_T(t) = x_T(t) - \bar{x}_T = a_1(t-\bar{t}) + z(t)$$

where $z(t)=x(t)-\bar{x}$. Thus $\mathrm{sgn}\ z_T(t) = \mathrm{sgn}(z(t)-D)$, where $D=a_1(t-\bar{t})$, $(a_1 \neq 0)$. $\mathrm{sgn}\ z_T(t)=\mathrm{sgn}\ z(t)$ if there are no trend ($a_1=0$, i.e., $D=0$). Thus we must consider the maximum values of $E[\mathrm{sgn}(x-\delta)\mathrm{sgn}(y-\delta)]$, where $X=Z(t)$, $Y=z(t+h)$ and $\delta=D/\sigma$ is normalized deviation of trend. The D's are obtained by regression coefficient a_1 and mean $\bar{t}$ and σ will be replaced with sample standard deviation s. For the case of $D=0.455$ for $a_1=0.0013$, $T-\bar{t}=350$ and $s=1.568$ (climatic series(6) in Okamoto and Iwase's paper(1982)), we have $\delta = 0.29$. Possible maximum error in covariance function of clipped series is $\Phi(\delta/\sqrt{2}) = 2\sqrt{2}\,I(\delta/\sqrt{2}) = 2.828I(.205) = 0.23$ by the use of the standard normal distribution function from 0 to x. We note that $D=0$ for the time point $t=350$.

4. CONCLUSION

When original climatic time series has a trend, possible maximum error due to the trend for covariance function of clipped series, from which we can make a synthetic climatic index series, is given. Theorem 2 here obtained is applicable for not only trend case but for occurence of some bias in original series.

5. REFERENCES

Gradshteyn,I.S. and Ryzhik,I.M., 1965: Tables of integrals,series,and products. Academic Press, New York.

Kedem,b., 1980: Binary time series, Marcel Dekker, New York.

Okamoto,M.B. and Iwase,k., 1982: A ststistical index representing a gross feature of several climatic time series by hard limiting. J. Meteor. Soc. Japan, 60, 726-738.

STOCHASTIC MODELING OF RAINFALL TEMPORAL PATTERNS

Van-Thanh-Van Nguyen

Department of Applied Sciences
University of Quebec in Chicoutimi
Chicoutimi, Québec, Canada

1. INTRODUCTION

Knowledge of the time distribution of rainfall within storm events is one of the most important factors in successful design and planning of many water resources systems. However, as reviewed by Nguyen (1979), most previous research on the time distribution of rainfall usually involved the study of temporal distributions of «rainfall exterior» characteristics such as total event depth, total event duration, time between events,..., but very few studies dealt with the time distribution of rainfall within each individual storm event or the temporal pattern of each event.

In a previous paper (Nguyen, 1982) a general stochastic model was proposed for characterizing the temporal pattern of rainfall. The model was then applied to hourly rainfall process in which the hourly rainfall depths were assumed to be independent and identically exponentially distributed random variables while the sequences of wet hours were represented by a first-order stationary Markov chain. Hence, the model did only introduce the dependence of wet hour occurences into the rainfall process through the first-order Markov chain. However, as observed in many previous investigations, there is a definite persitence inherent in the series of rainfall depths, specially in short-time-interval rainfall series (e.g., Pattison, 1965; Chown and Ramaseshan, 1965; Grace and Eagleson, 1966; Raudkivi and Lawgun, 1974). A better description of the variability of temporal storm patterns, therefore, would be achieved with a model which incorporates this essential observed feature of the rainfall time series.

The present study, a continuation of the previous one, suggests a more general model that can take into account both the persistence in hourly rainfall occurences and the dependence between successive hourly rainfall depths. In this paper we will use the same methodology as that proposed in the earlier cited work for characterizing the temporal storm pattern but will replace the independence assumption of rainfall depths by a more general assumption of correlation between successive rainfall amounts. A storm will be defined as an unbroken sequence of consecutive hourly rainfalls. The approach adopted consists of determining analytically the probability distribution functions for cumulative amounts of correlated hourly rainfall depths at successive time intervals after the storm began. We will demonstrate that by taking into account the correlation structure of consecutive rainfall depths the present model seems to be significantly superior to the previous model which was developed for independent rainfall amounts.

2. THEORETICAL DEVELOPMENT

Consider an interval of time which consists of n hours. Let M_n be defined as the number of consecutive rainy hours starting from the first hour of the n-hour period. According to the definition of a storm stated above, the beginning of a storm is the time at which a non-zero hourly rainfall depth occurs. Hence, in order to consider the time distribution of rainfall within different storm events, all storms are arranged to begin with the first hourly rainfall. Thus, the n-hour period considered is defined as an interval of time starting from the first hour of all rainy hour sequences or from the beginning of all the storms. By this definition, the random variable M_n can assume the values, 1,2,...,n.

Let ε_ν denote the hourly rainfall depth in the ν-th rainy hour. The accumulated amount of rainfall S(n) during M_n consecutive rainy hours of the n-hour period can be defined as

$$S(n) = \sum_{\nu=1}^{M_n} \varepsilon_\nu \qquad (1)$$

The random variable S(n) is the sum of a random number of random variables. Then the distribution function $F_n(x)$ of S(n) can be written as follows (Nguyen and Rousselle, 1981):

$$F_n(x) = P\{S(n) \leq x\} = \sum_{K=1}^{n} P\{X_K \leq x,\ M_n = K\} \qquad (2)$$

in which $P\{\cdot\}$ denotes the probability, and

$$X_K = \sum_{\nu=1}^{K} \varepsilon_\nu \quad \text{for } K = 1,2,\ldots,n.$$

In the following for hourly rainfall process, we assume that:

i) The hourly rainfall depths $\varepsilon_1, \varepsilon_2 \ldots, \varepsilon_n$ are correlated and identically distributed random variables with $P\{\varepsilon_\nu \leq x\} = 1 - e^{-\alpha x}$ for every $\nu = 1,2,\ldots,n$;

ii) $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ are independent of M_n; and

iii) The sequences of wet hours can be represented by a first-order stationary Markov chain.

Under the assumption (ii) we can write:

$$F_n(x) = \sum_{K=1}^{n} P\{X_K \leq x\} . P\{M_n = K\} \qquad (3)$$

Hence, to determine the distribution function $F_n(x)$, it is necessary to compute the probabilities $P\{X_K \leq x\}$ and $P\{M_n = K\}$.

As defined above, the random variable X_K is the sum of K correlated and identically exponentially distributed random variables. Using the method similar to that used by Kotz and Adams (1964) we easily obtain the distribution of the random variable X_K as follows:

$$P\{X_K \leq x\} = (-1)^{K-1} \sum_{i=1}^{K} \left[\prod_{j \neq i} \left(\frac{\lambda_j}{\lambda_i} - 1 \right)^{-1} . \left(1 - e^{-\frac{\alpha}{\lambda_i} x}\right) \right] \qquad \ldots (4)$$

where the λ_i's are the eigenvalues of the correlation matrix of successive hourly rainfall depths, and $\prod_{j \neq i} (\lambda_j/\lambda_i - 1)^{-1}$ is the product of the $(\lambda_j/\lambda_i - 1)^{-1}$'s for i,j = 1,2,...,K and $j \neq i$. Thus, from (4) it can be seen that the probability of the random variable X_K depends on the correlation structure inherent in the hourly rainfall process.

In this paper, to make comparisons with the results in the previous study (Nguyen, 1982) we recall that, by assuming the successive hourly rainfall amounts to be independent and identically exponentially distributed, the distribution function $P\{X_K \leq x\}$ can be obtained as follows:

$$P\{X_K \leq x\} = \frac{\lambda^K}{\Gamma(K)} \int_0^x u^{K-1} e^{-\alpha u} \, du \qquad (5)$$

where $\Gamma(K) = (K-1)!$ for K = 1,2,...,n.

On the basis of assumption (iii), the probability $P\{M_n = K\}$ can be computed by the following relations:

$$P\{M_n = n\} = q_{WW}^{(n-1)} \qquad (6)$$

and

$$P\{M_n = K\} = q_{WW}^{(K-1)} . q_{WD} \qquad (7)$$

for K = 1,2,...,n-1, in which q_{WW} and q_{WD} are respectively the transition probabilities of any present hour being wet (or dry) given that the preceding hour was wet. These transition probabilities will be estimated by using the corresponding observed frequencies from the historical rainfall record.

3. NUMERICAL APPLICATION

Nguyen (1982) tested the descriptive capabilities of the theoretical model, equation (3), in the case of uncorrelated hourly rainfall depths by fitting the distribution of X_K, equation (5), and the Markov chain model, equations (6) and (7), to hourly rainfall data from Dorval Airport station on the Island of Montreal (Canada). A total of 2560 summer rainfall events having durations of from 1 to 27 hours have been used. These events have been selected from a 32-year record (1943 - 1974) according to the definition of a storm specified above. To readily make comparisons with the results of the correlated rainfall case in the present study, results of the uncorrelated rainfall case are shown here again in Figs. 2-4 for n = 2,6 and 10 hours.

In this work we will use the same hourly rainfall data as were used in the previous study (Nguyen, 1982). We continue to assume that the sequence of wet hours be represented by a first-order stationary Markov chain but will replace the simplifying assumption of independent hourly rainfall depths by the assumption of correlation between successive hourly rainfall amounts. Serial correlation analyses were performed on the hourly rainfall series and the results for lag 1 to 20 are shown in Fig. 1. The correlogram shows a gradual decay from lag-one correlation coefficient value and thus, may indicate a highly dependent hourly rainfall series. A more objective method to assess the dependence between successive rainfall depths would be through the use of statistical tests of significance for the serial correlation coefficients. However, despite a very low correlation value of 0.01, the size of the statistical sample was such that even this value was, according to the test proposed by Anderson (1942), still significantly different from zero at the 95 percent significance level. It should be noted that the Anderson's test is only applicable to a circular, normal, stationary time series, this test however is used here as an indication of the existence of a strong persistence in the series of hourly rainfall depths at Dorval Airport.

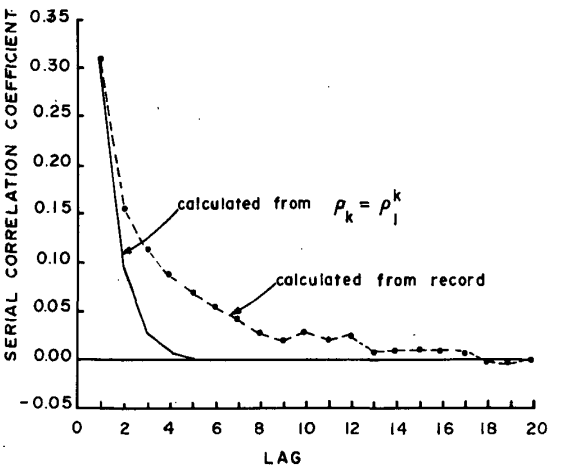


Fig. 1. Correlogram of summer hourly rainfalls at Dorval Airport and the correlogram for a first-order autoregressive process with $\rho_1 = 0.309$.

In the following, the computed values of the sample serial correlation coefficients will be directly used in the estimation of the eigenvalues specified in equation (4). The sampling error of these correlation coefficients could be small because they were estimated from a relatively large data sample. For illustrative purposes, values of the correlation coefficients and the corresponding eigenvalues listed in Table 1 are only for n = 6 hours.

Table 1. Sample correlation coefficient r_K and corresponding eigenvalues λ_i for n = 6 hours.

Lag K	r_K	i	λ_i
0	1.000	1	0.584
1	0.309	2	0.651
2	0.156	3	0.752
3	0.116	4	0.908
4	0.088	5	1.167
5	0.072	6	1.938

From a practical point of view, it would be more useful if the correlation structure of hourly rainfall process could be represented by a simple mathematical function. Previous studies by Chown and Ramaseshan (1965) and Chown and Prasad (1972) suggest that the exponential autocorrelation function $\rho_K = \rho_1^K$, where ρ_K and ρ_1 are respectively the serial correlation coefficients of lag K and lag one, might successfully be used to describe the correlation structure of hourly rainfall. It is known that (e.g., Cox and Miller, 1965) this autocorrelation function charactrizes the first-order autoregressive scheme. Figure 1 shows the comparison between the correlogram obtained from the observed rainfall data and the exponential autocorrelation function (ρ_1= 0.309). It can be readily seen that the first-order structure does not fit very well the summer hourly rainfall data of Dorval Airport. Rainfall process are too complex to allow us to expect that a rainfall time series should conform closely with a simple autoregressive process. Yet, for practical purposes, the usefulness of this first-order scheme need not be rejected one these grounds. Hence, we continue the analysis with this simple autocorrelation function. For purposes of illustration, values of the serial correlation coefficients ρ_K and the corresponding eigenvalues λ_i are given in Table 2 for n = 6 hours. In this paper, the probabilities $P\{X_K \leq x\}$ for K = 1,2,...,n are estimated by using, respectively, the serial correlation coefficients obtained from the observed rainfall data and those computed from the exponential autocorrelation function. The performance of these two models will be compared.

After the probabilities $P\{X_K \leq x\}$ and $P\{M_n = K\}$ were estimated as decribed above, the distribution function $F_n(x)$ for cumulative amount of correlated hourly rainfall depths was computed by using equation (3). The results obtained by using the exponential autocorrelation function were found not to be significantly different from those given by using the sample correlogram. This is fortunate because the simple exponential scheme can safely be used in lieu of the observed correlation structure, the calculation of which is more troublesome.

Table 2. Serial correlation coefficient ρ_K of the first-order autoregressive scheme and corresponding eigenvalues λ_i for n = 6 hours

Lag K	ρ_K	i	λ_i
0	1.000	1	0.550
1	0.309	2	0.623
2	0.096	3	0.764
3	0.030	4	1.000
4	0.009	5	1.344
5	0.003	6	1.719

As mentioned above, the longest unbroken sequence of consecutive hourly rainfalls that has been observed during the 32-year period is 27 hours. The 27-hour period was therefore chosen as the maximum time interval considered in this study. However, to make comparisons with the results of the previous investigation (Nguyen, 1982), observed and theoretical distributions plotted in Figs. 2-4 are only for n = 2,6 and 10 hours. These results show that the model for correlated rainfall depths gives a better fit to the observations than the model for the uncorrelated case.

4. CONCLUSIONS

A stochastic method that can take into account both the persistence in hourly rainfall occurences and the correlation between successive hourly rainfall amounts has been suggested in this study for characterizing the temporal pattern of storm. It can be concluded that the present approach is more general and it may give better results than the method proposed in an earlier work for the case of independent hourly rainfall depths. Furthermore, the use of the suggested stochastic model for constructing analytically distributions of cumulative rainfall amounts at successive time intervals seems preferable to the simulation or fitting approaches that have been used in most previous investigations.

5. ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada, the University of Quebec in Chicoutimi and the Ministry of Education of Quebec (joint project with Ecole Polytechnique de Montréal).

6. REFERENCES

Anderson, R.L., 1942: Distribution of the serial correlation coefficients, Ann. Math. Statist., 13, 1-13.

Chown, V.T., and T. Prasad, 1972: Theory of sto-

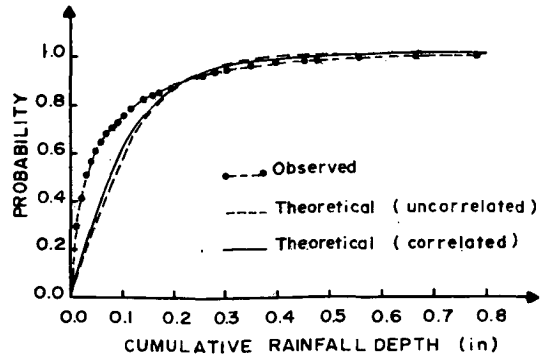


Fig. 2. Observed and theoretical distributions of accumulated rainfalls in the 2-hour period.

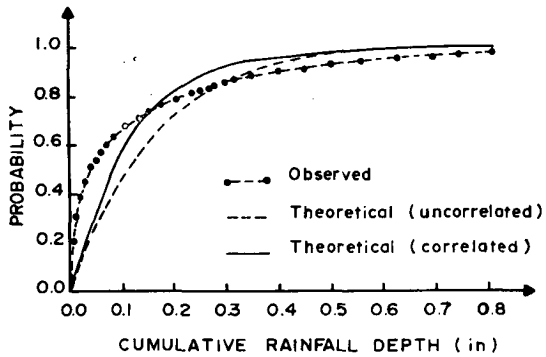


Fig. 3. Observed and theoretical distributions of accumulated rainfalls in the 6-hour period.

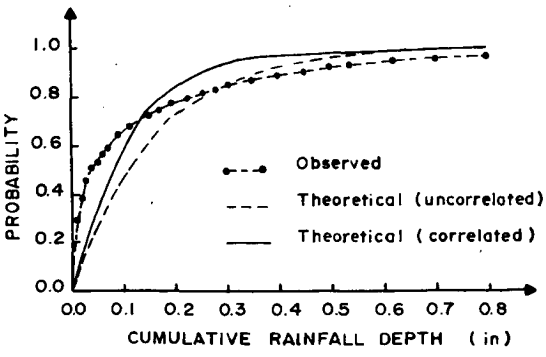


Fig. 4. Observed and theoretical distributions of accumulated rainfalls in the 10-hour period.

chastic modeling of watershed systems, J. Hydrol., 15, 261-284.

Chown, V.T., and R. Ramaseshan, 1965 Sequential generation of rainfall and runoff data, J. Hydraul. Div., Am. Soc. Civ. Eng., 91 (HY4), 205-223.

Cox, D.R., and H.D. Miller, 1965: The theory of stochastic process, Chapman and Hall, London, 398 p.

Grace, G.A., and P.S. Eagleson, 1966: The synthesis of short-time increment rainfall sequences, Report 91, Hydrodyn. Lab., Mass. Inst. of Technol., Cambridge.

Kotz, S., and J.W. Adams, 1964: Distribution of sum of identically distributed exponentially correlated gamma-variables, Ann. Math. Statist., 35, 277-283.

Nguyen, V.T.V., 1979: Caractérisation hydrologique de la pluie pour le contrôle et la planification des systèmes de drainage urbain, Doctoral thesis, Ecole Polytechnique de Montréal, Québec, Canada.

Nguyen, V.T.V., 1982: A stochastic approach to characterization of temporal storm patterns, Proc. of the 1982 Intern. Symp. on Urban Hydrology, Hydraulics and Sediment Control, July 27-29, Lexington, Kentucky, 263-265.

Nguyen, V.T.V., and J. Rousselle, 1981: A stochastic model for the time distribution of hourly rainfall depth, Water Resour. Res., 17(2), 399-409.

Pattison, A., 1965: Synthesis of hourly rainfall data, Water Resour. Res., 1(4), 489-498.

Raudkivi, A.J., and N. Lawgun, 1974: Simulation of rainfall sequences, J. Hydrol., 22, 271-294.

A METHOD OF ANALYZING IMPLIED PERIODS OF CLIMATOLOGICAL DATA USING STEP-WISE REGRESSION

F. Y. MEI, Z. ZHAO and X. G. ZHANG

Academy of Meteorological Science

Beijing, China

1. INTRODUCTION

One of the purposes of time series analysis of climatological data is to abstruct implied periods. To do so, a variety of methods have been employed. Maybe first used was Schuster periodogram based on harmonic analysis (Brooks, 1953), H.A. panofsky (1957) introduced spectra analysis as a powerful tool into climatology. But it has rather low resolution for short data series. This was much improved by maximum entropy spectra analysis (Burg, 1967).

Another approach to detect implied periods is based on variance analysis (Brooks, 1953). Climatological series $X(t)$ is assumed to be composed of several pure periodic series plus a white noise:

$$X(t) = \sum_{i=1}^{M} f_i(t) + E(t) \qquad (1)$$

Where $f_i(t)$ is periodic function with period length l_i, $i=1, 2, \ldots M$. $E(t)$ is white noise.

$f_i(t)$ and l_i are to be estimated based on observations of $X(t)$ $X_1, X_2, \ldots X_n$, where $n \geq 2M$.

The current method named "periodic extrapolation using variance analysis", which will be briefly called as "variance method" hereafter for simplicity, makes a good use of variance analysis to estimate $f_i(t)$ and $E(t)$. Unfortunately when more than one periods are to be found, this method is mathematically incorrect (Details will be discussed in next section).

In this study, a new technique based on multiple step-wise regression is suggested. An example of analyzing implied periods of mean temperature of January 1951 — 1980 in Beijing is given.

2. METHOD

Let

$$X = (X_1, X_2, \ldots X_n)'$$

denote a climatological series.

Construct periodic series

$$Z_1, Z_2, \ldots Z_M$$

$1 \le M \le \left[\frac{n}{2}\right] - 1$ as following:

$$\mathbf{X}_1 = (X_1^1, X_2^1, \ldots X_n^1)' \qquad (2)$$

where

$$X_1^1 = X_3^1 = X_5^1 = \ldots = X_2^1 \times \left[\frac{n+1}{2}\right] - 1$$

$$= \frac{1}{\left[\frac{n+1}{2}\right]} \sum_{i=1}^{\left[\frac{n+1}{2}\right]} X_{2xi-1}$$

$$X_2^1 = X_4^1 = X_6^1 = \ldots = X_2^1 \times \left(\frac{n}{2}\right)$$

$$= \frac{1}{\left[\frac{n}{2}\right]} \sum_{i=1}^{\left[\frac{n}{2}\right]} X_2 \times i$$

$$\mathbf{X}_2 = (X_1^2, X_2^2, \ldots X_n^2)' \qquad (3)$$

where

$$X_1^2 = X_4^2 = X_7^2 = \ldots X_3^2 \times \left[\frac{n+2}{3}\right] - 2$$

$$= \frac{1}{\left[\frac{n+2}{3}\right]} \sum_{i=1}^{\left[\frac{n+2}{3}\right]} X_3 \times i - 2$$

$$X_2^2 = X_5^2 = X_8^2 = \ldots = X_3^2 \times \left[\frac{n+1}{3}\right] - 1$$

$$= \frac{1}{\left[\frac{n+1}{3}\right]} \sum_{i=1}^{\left[\frac{n+1}{3}\right]} X_3 \times i - 1$$

$$X_3^2 = X_6^2 = X_9^2 = \ldots = X_3^2 \times \left[\frac{n}{3}\right]$$

$$= \frac{1}{\left[\frac{n}{3}\right]} \sum_{i=1}^{\left[\frac{n}{3}\right]} X_3 \times i$$

....

$$\mathbf{X}_k = (X_1^k, X_2^k, \ldots X_n^k)' \qquad (4)$$

where

$$X_1^k = X_{k+2}^k = X_{2k+3}^k = \ldots = X_{(k+1)\times\left[\frac{n+k}{k+1}\right] - k}^k$$

$$= \frac{1}{\left[\frac{n+k}{k+1}\right]} \sum_{i=1}^{\left[\frac{n+k}{k+1}\right]} X_{(k+1)} \times i - k$$

$$X_2^k = X_{k+3}^k = X_{2k+4}^k = \ldots = X_{(k+1)\times\left[\frac{n+k-1}{k+1}\right] - (k-1)}^k$$

$$= \frac{1}{\left[\frac{n+k-1}{k+1}\right]} \sum_{i=1}^{\left[\frac{n+k-1}{k+1}\right]} X_{(k+1)xi-(k+1)}$$

....

All the way through $\mathbf{X}_m$. Where $\mathbf{X}'$ denotes the transpose of $\mathbf{X}$, $[A]$ denotes the largest integer not exceeding A.

In order to see if $X_{(t)}$ has an implied period with length 1, according to variance method, an F - test is carried out:

$$F^L = \frac{S^L / (L-1)}{S / (n-L)} \qquad (5)$$

where

$$S^L = \sum_{i=1}^{l} n_i \, (X_i^l - \bar{X})^2$$

$$S = \sum_{i=1}^{n} (X_i - \bar{X})^2$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$N_i = \left[\frac{n+L-i+1}{L+1}\right]$$

When $F^L > F\alpha$ (α is a given significant level), $X_{(t)}$ could be considered to have an implied periodic component with value X_1^L, X_2^L, X_3^L, ... X_n^L.

Usually $F^{L1} = \max_{1 \le L \le m} \{ F^L \}$ is first taken

into account. Once L_1 is defined, a residual series

$$X_1 - X_1^{L1},\ X_2 - X_2^{L1},\ \ldots\ X_n - X_n^{L1} \qquad (6)$$

is constructed. Then the same procedure is repeated for the residual series. Second, third, ... kth periods $L_2, L_3, \ldots L_k$ are found in this process. Suppose K significant periods are detected.

$$\tilde{X}_{(t)} = \sum_{i=1}^{k} X_{(t)}^{Li} \qquad (7)$$

is an approximation of $X_{(t)}$. And

$$\tilde{X}_{n+1} = \sum_{i=1}^{k} X_{n+1}^{Li} \qquad (8)$$

is an extrapolation of $X_{(t)}$.

In the simplest case, k=1 in (1), then

$$X_{(t)} = f_{(t)} + E_{(t)} \qquad (9)$$

Where $f_{(t)}$ is a periodic function with length L, $E_{(t)}$ a white noise. The corresponding series is

$$\begin{aligned} &f_1 + E_1,\ f_2 + E_2,\ \ldots\ f_L + E_L,\\ &f_1 + E_{L+1},\ f_2 + E_{L+2},\ \ldots\ f_L + E_{2L},\\ &f_1 + E_{2L+1},\ f_2 + E_{2L+2},\ \ldots\ f_L + E_{3L},\\ &\ldots. \end{aligned} \qquad (10)$$

It is obvious that the first column $f_1 + E_1$, $f_1 + E_{L+1}$, $f_1 + E_{2L+1}$, ... are distributed as normal distribution $N(f_1, \sigma)$, the second column $f_2 + E_2$, $f_2 + E_{L+2}$, $f_2 + E_{2L+2}$, ... are distributed as normal distribution $N(f_2, \sigma)$, ... Thus, to use variance analysis to make statistical inference that there is an implied period L in climatological series $X_1, X_2, \ldots X_n$ is very artful. But in general case, when $M \geq 2$ in (1), it is incorrect. Still for simplicity, let

$$X_{(t)} = f_{1(t)} + f_{2(t)} + E_{(t)} \qquad (11)$$

Where f_1, f_2 are periodic function with period L, L^* respectively, $L \neq L^*$, $E_{(t)}$ white noise.

Now, series $X_1, X_2, \ldots X_n$ takes form

$$\begin{aligned} &f_{1(1)} + f_{2(1)} + E_{(1)},\ f_{1(2)} + f_{2(2)} + E_{(2)},\ \ldots\\ &f_{1(L)} + f_{2(L)} + E_{(L)},\\ &f_{1(1)} + f_{2(L+1)} + E_{(L+1)},\ f_{1(2)} + f_{2(L+2)} + E_{(L+2)}\\ &f_{1(L)} + f_{2(2L)} + E_{(2L)},\\ &\ldots \end{aligned} \qquad (12)$$

Because $L \neq L^*$, so $f_{1(1)} + f_{2(1)} \neq f_{1(1)} + f_{2(L+1)}$, then first column $f_{1(1)} + f_{2(1)} + E_{(1)}$, $f_{1(1)} + f_{2(L+1)} + E_{(L+1)} \ldots$ are not from same population. In this case, to insist on variance analysis is unreasonable.

To overcome this difficulty, step-wise screen regression technique is adopted instead of variance analysis. A multiple regression model is set:

$$\mathbb{X} = a_0 + \sum_{i=1}^{M} a_i \mathbb{X}_i + \mathbb{E} \qquad (13)$$

Where a_0 is a constant, a_i regression coefficient, $\mathbb{X}_i$ periodic series in (2),(3), (4). $\mathbb{E}$ white noise.

Naturely step-wise screen regression technique (Enslein, 1977) is employed to estimate a_i a_0. When estimation $\hat{a}_i$ is not zero, climatological series $X_{(t)}$ is considered to imply a periodic component $\mathbb{X}_i$ with period i. This method will be called "Regression Method".

3. AN EXAMPLE

The series of mean temperature of January 1951-1980 in Beijing is analyzed. The first three significant periods are 12, 15 and 9 years. The cycles are also convinced by maximum entropy power spectrum analysis (results are not shown). Almost same results could be obtained by variance method, but with comparitively very low significance, see table 1.

Period (year)	Method	F	$F_{0.05}$
12	variance	1.80	2.67
	regression	30.05	
15	variance	2.09	2.60
	regression	22.39	
8	variance	1.87	3.43
9	regression	11.41	3.13

Table 1. Implied periods of mean temperature of January 1951-1980 in Beijing, revealed by variance method and regression method

The data is also fitted by compositing three periodic components picked up by these two method respectively.

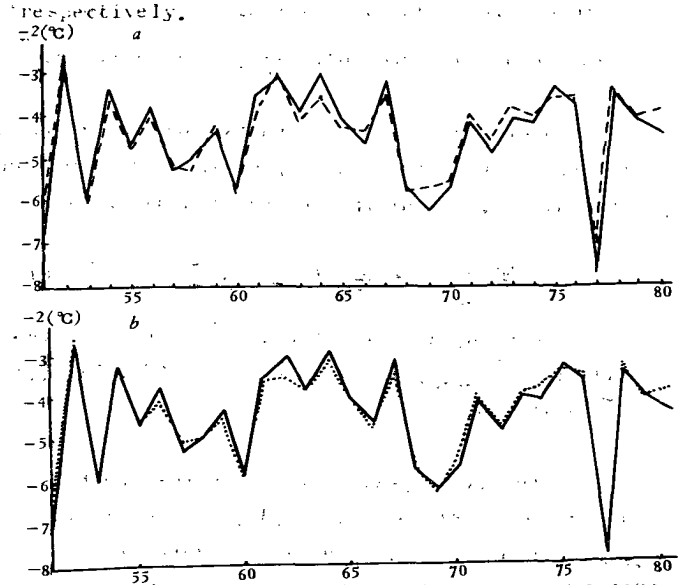


Fig 1. Mean temperature of January 1951-1980 in Beijing (solid line) and a, fitting curve of variance method (dashed line) and b, fitting curve of regression method (dotted line)

It could be seen from Fig. 1 that regression method greatly improves fitting using same number of periodic components.

This technique has also been applied to the Zurich series of annual relative sunspot numbers and precipitation of May to September in Kun Ming.

4. SUMMARY

To improve the resolution of currently used period analysis such as periodogram, spectra analysis, variance analysis, a new technique based on multiple step-wise regression is suggested. The examples investigated give us an impression that this method is better than spectra analysis and variance method in some cases.

This is the authors' first effort to attack the time series problems using multivariate analysis.

5. REFERENCES

Brooks, C. E. P. and Carruthers, N., 1953: Handbook of Statistical Methods in Meteorology. London, Her Majesty's Stationary Office.

Panofsky, H.A., 1957: Spectrum and cross spectrum analysis of hemisphere westerly index, Tellus, 9.

Burg, J.P., 1967: Maximum entropy spectral analysis, Paper presented at the 37th Annual International Meeting, Soc of Explor. Geophys. Oklahama city, okla. Oct.31.

Rao, C.R., 1973: Linear Statistical Inference and Its Application, Wiley, New York.

Anderson, T.W. 1958: An Introduction to Multivariate Statistical Analysis, Wiley, New York.

Enslein, K., Ralstan, A., Wilf, H.S., 1977: STATISTICAL METHODS FOR DIGITAL COMPUTERS, VOLUME III of Mathematical Methods for Digital Computers, John Wiley and Sons, Inc.

SOME APPLICATIONS OF FUZZY SETS TO ANALYSIS OF CLIMATE AND CLIMATE SIMULATION *

Hongxing Cao **

Meteorologisches Institut der Universität Hamburg
D-2000 Hamburg 13, F.R. Germany

1. INTRODUCTION

Climatology development has always been connected with statistics, since climate has been defined as an average state of the atmosphere. Statistics is the most important tool to deal with meteorological observations and to put them into useful information. Of course we may say that there is an uncertainty in climate phenomena. It is necessary to apply statistics to treating this kind of phenomena. If we think deeply about this, it is found that there are two different uncertainties in climate. One is the so called randomness and the other fuzziness. For example, there is no strict limit between the tropical and mesothermal climatic regions. With respect to time, there is also no critical day to separate the four seasons. To deal with this problem, fuzzy sets theory will play a more important role in climatology research. In this paper the application of fuzzy neartude in analysing the climate stage and climate simulation is presented.

2. FUZZY NEARTUDE

What is fuzzy neartude? To introduce this concept, we first need to explain something about an ordinary set and a fuzzy set.

A correction of objects, in which nothing special is assumed about the nature of the individual objects, is called a set, denoted by A, B, . . . X, Y, Z. An individual object x in the collection is called an element or a member of the set A and is said to belong to the set, denoted by

$$x \in A$$

Generally, the characteristic function defined as

$$\mu_A(x) = 1 \quad \text{if } x \in A$$
$$\mu_A(x) = 0 \quad \text{if } x \bar{\in} A \qquad (1)$$

for all x in A, is described a set. It is shown from (1) that the value of $\mu_A(x)$ takes two values 1 or 0 as x does or does not belong to A, denoted as $\{0, 1\}$. This is true for many cases, for example a group of men, a sequence of rational numbers etc. But in numerous cases it cannot be clearly determined whether or not an object belongs to a set, e.g., a collection of new cars, a class of short men. When is a new car 'new'? When is a short man called 'high'? In fact, most of the classes of objects encountered in the real physical world are of this fuzzy, not sharply defined type. So we define a fuzzy set, which is a 'class' with a continuum of grades of membership, denoted as $[0, 1]$. Let X be a set and x be an element of X, then a fuzzy set A of X is a set of ordered pairs (Kaufmann, 1973, Kandel and Lee, 1979)

$$(x \mid \mu_{\tilde{A}}(x)) \quad \forall x \in X,$$

where $\mu_{\tilde{A}}(x)$ is a membership function that has its values in the interval $[0, 1]$. From the above definition of a fuzzy set we observe that the nearer the value of $\mu_{\tilde{A}}(x)$ to unity, the higher the grade of membership of x in $\tilde{A}$, and that the nearer the value of $\mu_{\tilde{A}}(x)$ to zero, the lower the grade of membership of x in $\tilde{A}$. If A is a non-fuzzy set, then $\mu_{\tilde{A}}(x)$ is reduced to a characteristic function, which can only have the two values 0 and 1.

Let level $\lambda \in [0, 1]$; one will call the ordinary subset

$$A_\lambda = \{ x \mid \mu_{\tilde{A}}(x) \geq \lambda \} \qquad (2)$$

the λ-cutoff subset of the fuzzy set A.

Let $A = \{ x_1/0.9, x_2/0.1, x_3/1, x_4/0.5, x_5/0.7, x_6/0.4 \}$, where the figure after the slash shows the membership function of the element x_i (i = 1, 2, . . . 6). The ordinary subset corresponding to $\lambda = 0.5$ and $\lambda = 0.7$ is respectively

$$A_{0.5} = \{ x_1/1, x_2/0, x_3/1, x_4/1, x_5/1, x_6/0 \}$$
$$= \{ x_1, x_3, x_4, x_5 \},$$
$$A_{0.7} = \{ x_1/1, x_2/0, x_3/1, x_4/0, x_5/1, x_6/0 \}$$
$$= \{ x_1, x_3, x_5 \}.$$

A fuzzy neartude is defined as a measure that shows the approaching degree between two fuzzy sets just like a similarity coefficient in multivariate statistical analysis (Wang, 1980). The mathematical concept of a neartude is defined as follows.

The power set of universal set X, denoted by F(X), is a set of fuzzy subsets. Denoting a Cartesian product F(X) x F(X), a mapping T: F(X) x F(X) $\rightarrow$ $[0, 1]$ is called a neartude between fuzzy sets in X, if it satisfies

a1. $T(\tilde{A}, \tilde{A}) = 1$

a2. $T(\tilde{A}, \tilde{B}) = T(\tilde{B}, \tilde{A}) \geq 0$

a3. if there is either $\tilde{A}(x) \geq \tilde{B}(x) \geq \tilde{C}(x)$ or $\tilde{A}(x) \leq \tilde{B}(x) \leq \tilde{C}(x)$, $\forall x \in X$, then $T(\tilde{A}, \tilde{C}) \leq T(\tilde{B}, \tilde{C})$.

* This research is supported by Alexander von Humboldt Foundation.

** Permanent address: Academy of Meteorological Science, Central Meteorological Bureau, 46 Baishiqiaolu, Beijing, China.

Nevertheless, the neartude for fuzzy sets that satisfies a1, a2 and a3 is not unique (Zhao and Chen, 1981)

A distance neartude is defined by

$$T(\underset{\sim}{A}, \underset{\sim}{B}) = 1 - d(\underset{\sim}{A}, \underset{\sim}{B}),$$

where $d(\underset{\sim}{A}, \underset{\sim}{B})$ is normalized Hamming distance, Euclidean distance and Minkowski distance etc.

3. PARTITIONING CLIMATE STAGES WITH FUZZY CLUSTERING

Partition of climate stages is important for both understanding and predicting climate. In recent years, various statistical methods on the basis of historical records to divide climatic change into several stages have been developed. It is more natural to utilize fuzzy sets to do this, since there is only an unclear limit between two climate stages. We can simply consider only two climate stages which alternate between each other: wet periods and drought periods. As an application of fuzzy clustering (Tamura, 1971; Cao and Chen, 1983) with neartude, a tentative result is given as follows.

Let a set $P = \{ p_1, p_2 \ldots, p_n \}$ be n annual precipitations at Shanghai, being preprocessed by 10-year running mean, i.e.

$$p_i = [\, p_{i-4} + \ldots + p_{i-1} + p_i + p_{i+1} + \ldots + p_{i+5} \,]/10 \quad (i = 1, 2, \ldots, n)$$

where p_{i+k} $(k = -4, -3, \ldots 0, \ldots, +5)$ are original records of precipitation. The 10-year running mean being averaged from the annual precipitation observations of 1886-1979, 94 records total, n is equal to 85. The Hamming distance neartude

$$S'_{ij} = 1 - | p_i - p_j | \quad (i, j = 1, 2, \ldots n) \qquad (3)$$

is taken. A n×n matrix S' is calculated and normalized into the interval [0, 1] with the transformation formula

$$S_{ij} = (S'_{ij} - S'_{min})/(S'_{max} - S'_{min}), \qquad (4)$$

where S'_{min} and S'_{max} are respectively the minimum and maximum elements in the matrix S'. The matrix S, which consists of S_{ij} has both reflexivity

$$s_{ii} = 1 \quad (i = 1, 2, \ldots, n),$$

and symmetry $s_{ij} = s_{ji}$ $(i, j = 1, 2, \ldots, n)$. But transitivity $S \circ S \leq S$ is not satisfied, where $S \circ S$ is self-multiplication of S. It is necessary to transform the fuzzy matrix S into a transitive one. Write

$$S^2 \triangleq S \circ S, \ldots, S^z \triangleq S \circ S^{z-1}$$

as 2, . . ., z power of S. Its calculation is similar to matrix multiplication but + and x are substituted by v and ∧. The calculation rule of every element for 2 power of S is

$$S^2_{ik} = \bigvee_{j=1}^{n} (S_{ij} \wedge S_{jk}) \quad (i, k = 1, 2, \ldots, n)$$

in which v and ∧ mean that a v b = max (a, b) and a ∧ b = min (a, b) respectively. A procedure of 2 power is thus repeated, i.e.

$$S \rightarrow S^2 \rightarrow S^4 \rightarrow \ldots S^z \rightarrow S^{2z}.$$

If the matrix S consists of m elements, there must be a $z \leq m$ to make the following equality tenable:

$$S^z = S^{z+1} = \ldots = S^{\infty}$$

Denoting $T = S^z$ which has satisfied transitivity, clustering can be performed in the matrix T. With the λ level decreasing from 1 to 0 and taking

$$Y_{ij} = \begin{cases} 1 & \text{if } t_{ij} \geq \lambda \\ 0 & \text{if } t_{ij} < \lambda \end{cases} \quad (i, j = 1, 2, \ldots, n),$$

many matrices in the sense of ordinary set, which consists of only 1 and 0, are obtained. Investigating the variation of these matrices, the formation process of several climate stages can be understood. The turning point of climate change and different features among the stages can also be found. When we take the certain level λ_0 with empirical judgement, several boundaries between wet and drought period are determined objectively. Even though the 53-67 wet period is nonhomogeneous (Fig. 1), as it is mixed in with especially rainy records like the 60th and 63rd sample, and the boundary between the 53-67 wet period and the 68-85 drought period is vague, it still can be divided into two periods satisfactorily. The result of partitioning participitation stages of Shanghai, i.e. the samples 1-17, 31-52, 68-85 forming the drought periods and the samples 18-30, 53-67 forming the wet periods, is in accordance with experimental judgement and relevant research. During the period of 1886-1979 at Shanghai, the total length of wet periods is shorter than that of drought periods. The variability in the wet periods is small but great in drought periods.

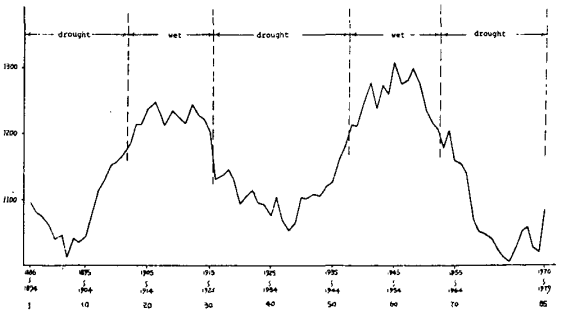


Fig. 1: Annual precipitation (10-year running mean) at Shanghai during 1886-1979 (unit: mm).

4. VERIFICATION OF CLIMATE SIMULATION

Verification of climate simulation with an atmospheric general circulation model (GCM) is made early by simple side-by-side comparisons of individual monthly or seasonal means from a single GCM realization and multiyear averages of the same quantities from observed data (Manabe et al., 1965). This approach tends to be somewhat subjective. Then univariate statistical tests are proposed (Chervin, 1981) which permit a straightforward assessment of the extent to which observed and GCM simulated climates agree or disagree. But making a lot of univariate tests instead of one single multivariate test is not statistically complete. As a global test, a comparative study of model climate with observed by means of multivariate statistics (Storch, 1982) was

performed, which is useful for answering the question whether the simulated states show the tendency to stay within the real climate and revealing some deficiences of the model.

Another approach to this problem is suggested here, which is based on fuzzy neartude as mentioned above. In order to check validity of this approach, a example with the same data in Storch's test is calculated, where only 7 EOFs (Empirical Orthogonal Function), i.e. EOF 16-21 and 30 are taken into account.

The data used for this study are the experimental results produced by the GCM of Hamburg University with 3-layer / 2.8° version. The model generated 500 mb fields cover 90 days. The climate mean field is estimated from daily DWD (Deutscher Wetterdienst) analysis of Januaries 1967-1977. Both the model and the atmosphere 500 mb height fields over the northern hemisphere are expanded into the first 30 EOFs by Rinne and Karhila's (1979) method. Therefore the variable x_i (i = 1, 2, . . . L, L = 30) in multivariate calculation is the EOFs.

A fuzzy neartude defined by

$$T_j(\underset{\sim}{A}, \underset{\sim}{B}) = 1 - \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\underset{\sim}{A}(x_{ij}) - \underset{\sim}{B}(x_{ij}))^2} \qquad (5)$$

(j = 1, 2, . . . M)

is adopted, in which $\underset{\sim}{A}(x_{ij})$ is a membership function of model generated climate and $\underset{\sim}{B}(x_{ij})$ a membership function of observed climate; M is the number of days simulated by GCM. Practically, the exact formulae of membership function under study are

$$\underset{\sim}{A}(x_{ij}) = 1 \quad \text{if } x_{ij} \in [\mu_i - 1.96\sigma_i, \mu_i + 1.96\sigma_i];$$

$$\underset{\sim}{A}(x_{ij}) = \exp(-(x_{ij} - (\mu_i + 1.96\sigma_i))^2/\sqrt{2}\sigma_i) \qquad (6)$$

$$\text{if } x_{ij} > \mu_i + 1.96\sigma_i;$$

$$\underset{\sim}{A}(x_{ij}) = \exp(-x_{ij} - (\mu_i - 1.96\sigma_i))^2/\sqrt{2}\sigma_i)$$

$$\text{if } x_{ij} < (\mu_i - 1.96\sigma_i);$$

(i = 1, 2, . . ., N; j = 1, 2, . . ., M)

where μ_i denotes the ith EOF of the mean field of observed climate. σ_i denotes the ith EOF's standard deviation of the mean field of observed climate; as mentioned above, N = 7, M = 90. For observed climate

$$\underset{\sim}{B}(x_{ij}) = 1 \quad (i = 1, 2, \ldots, N; j = 1, 2, \ldots, M)$$

is taken. This means that the observed climate is considered to ordinary set, because it is taken as a constant within the simulated days.

The membership function has its own meaning. If the simulated value places within the 95% confidence interval $[\mu_i - 1.96\sigma, \mu + 1.96\sigma]$, the membership function is designated as 1. Beyond this interval the membership function falls in a normal curve with the departure of simulated value from the climate mean μ_i.

In order to know to which statistical significance level the λ level corresponds, an empirical approach is used. The Euclidean neartude are calculated daily for the observed 500 mb hemispheric heights of 1967-1982's January with (5) and (6). Therefore, a list of 491 neartudes is composed.

Fixing subjectively a significance level $1 - \alpha$, e.g., $\alpha = 0.01$, there will be λ_α corresponding to α. This means that there should be $491 \times 0.01 = 491 \approx 5$ days whose neartude is smaller than $\lambda_{0.01}$, i.e., if we cut off the fuzzy set of neartudes with $\lambda_{0.01}$, it makes 1% risk exist for a test. From the list we can find the critical neartude, $\lambda_{0.01} = 0.746$ under which there are 5 cases with smaller neartude ($T < 0.746$). Similarly, we obtain $\lambda_{0.05} = 0.907$ (corresponding to $491 \times 0.05 = 24.55 \approx 25$ days) and $\lambda_{0.10} = 0.961$ (corresponding to $491 \times 0.10 = 49.1 \approx 49$ days).

The neartude $T_j(A, B)$ (j = 1, 2, . . ., 90) was calculated daily. A resulting curve and a level $\lambda = 0.907$ corresponding to significance level 95% (dashed line) are shown in Fig. 2. In order to compare, the daily statistic g(x) (Storch, 1982) and X^2 value of 95% fractile (solid line) are also given in Fig. 2. The statistic $g(x_j)$ is

$$g(x_j) = (x_j - \mu)^T S^{-1} (x_j - \mu), \quad (j = 1, 2, \ldots 90)$$

where both x_j and μ are the vector composed of N EOFs for the daily simulated 500 mb height fields and the observed mean field respectively, S^{-1} is the inverse of the covariance matrix corresponding to μ, T denotes transpose. The elements of λ-cutoff set ($\lambda = 0.907$) for neartude fuzzy set are

{1-3, . . ., 5-9, . . ., 12, . . .19, . . ., 22-45,
. . ., 48-50, . . . 52, . . ., 54 . . ., 60, 61, . . .,
63-66, . . ., 70-82, . . ., 84-90},

where the suspension points show the elements that their neartude is less than 0.907.

It is easily seen from Fig. 2 that variation of both curves is quite similar except a couple of days, e.g. 4th, 21th and 59th day. It is interesting that the day which is simulated the worst by the model is the same in both tests, i.e. the 69th day ($T(\underset{\sim}{A}, \underset{\sim}{B}) = 0.672$, $g(x) = 27.9$). The periods that are better simulated by the model are days 22-45 and days 70-90. The periods that are worse simulated by the model are days 13-18, days 55-59 and days 67-69. There are 25 days in total having $T_j(A,B) < \lambda_{0.05}$. This shows that some unrealistic climate states are simulated by Hamburg University GCM.

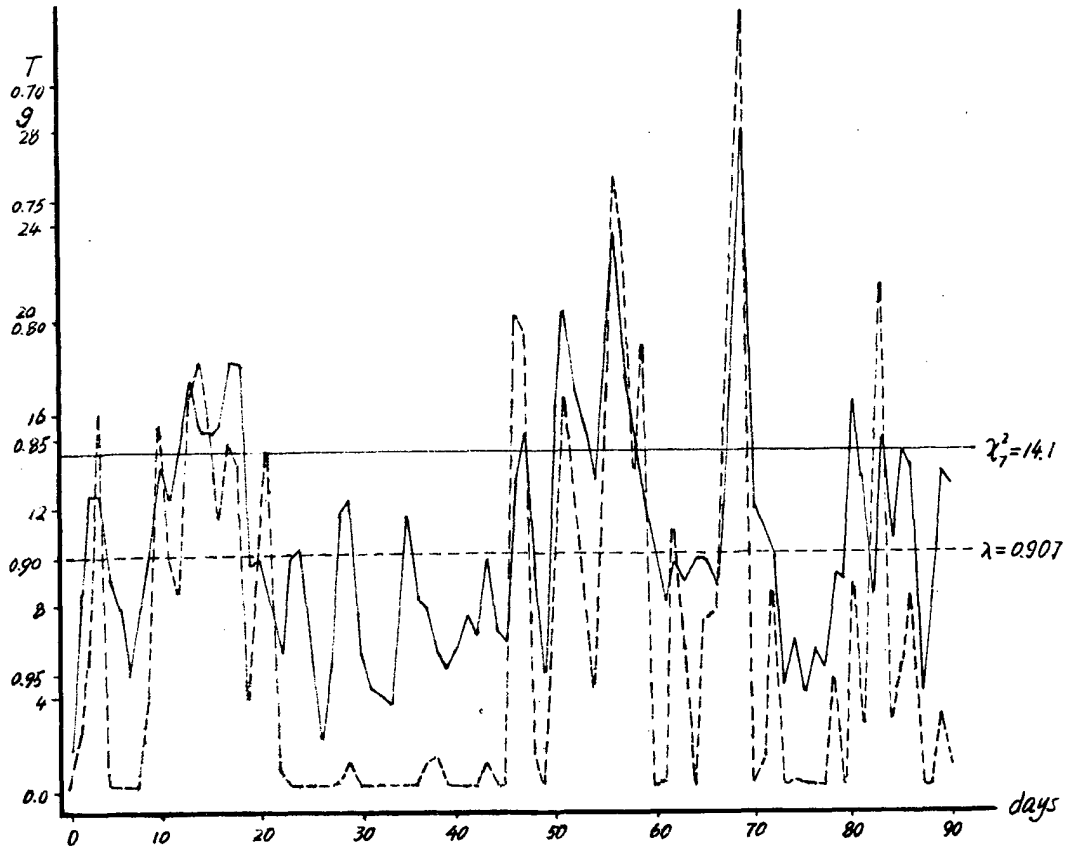


Fig. 2: The Euclidian distance neartude (dashed line) and g statistic (solid line) daily between the model generated 500 mb height fields and the climate mean field.

5. ACKNOWLEDGEMENTS

I wish to thank Prof. Guenter Fischer of the Meteorologisches Institut der Universität Hamburg for his encouragement and hearty hospitality. I am indebted to Dr. Hans von Storch who has given me much help, without whose help I could not have completed this work. Thanks are also due to Mrs. Zinecker for her typing the manuscript.

REFERENCES

Cao, H.X. and G.F. Chen, 1983: Some applications of fuzzy sets to meteorological forecasting, Fuzzy Sets and Systems, 9, 1-12.

Chervin, R.M., 1981: On the comparison of observed and GCM simulated climate ensembles. J. Atmos. Sci., 38, 885-901.

Kandel, A. and S.C. Lee, 1979: Fuzzy switching and automata, Crane Russak, New York, Edward Arnold, London.

Kaufmann, A., 1975: Introduction to the theory of fuzzy sets, Academic Press, INC., London.

Manabe, S., et al., 1965: Simulated climatology of a general circulation model with a hydrological cycle. Mon. Wea. Rev., 93, 769-798.

Rinne, J. and V. Karhila, 1979: Empirical orthogonal functions of 500 mb height in the northern hemisphere determined from a large data sample. Quart. J. Roy. Met. Soc., 105, 873-884.

Tamura, S., 1971: Pattern classification based on fuzzy relation, IEEE Trans. Systems Man Cybernet., 1, 61-66.

Wang, P.Z., 1980: Element of fuzzy sets, Mathematical practice and knowledge, No. 2 of 1980, (in Chinese).

Zhao, R.H. and X.J. Chen, 1981: The universal definition of neartude for the fuzzy subsets and the grade of imprecision, J. Xian Jiaotong University, 15, No. 6, (in Chinese).

ON THE REGULARIZATION OF THE PLUVIOMETRIC SERIES°

Giuseppe Flocchini* and Cristina Bagnasco**

Istituto di Geofisica* e Istituto di Matematica**
Università di Genova, Italia

1. INTRODUCTION

As it is well know, the precipitations show in the time strong irregularities; at Genoa variations in the annual values of about 50% occur very frequently. These variations are mainly due to the variability of the position of the great action centres; moreover, the particular oro-geographic outline, inside is the station, gives rise to the rain phenomena, which reach often in autumn very great intensity and sometime also alluvial character (BOSSOLASCO et Al., 1971; DAGNINO et Al., 1978).

The ordinary Mathematical instruments for the periodicities analysis as the D.F.T. (Discrete Fourier Transform) was already applied to the raw data of climatological series (SCHÖNWIESE, 1974 ; 1975) ; these methods applied to our pluviometric series not permit to emphasize cycles or trends because the strong variability of the precipitations.

For this reason, in order to process the daily data of the time series with the ordinary methods of analysis of the periodicities , attempts were made by us by a mathematical model basing oneself on the Volterra operator (FLOCCHINI et. Al., 1981). The regularization of our pluviometric series by this instrument has allowed to emphasize through the D.F.T. the presence of waves with 11 and 22 year period.

The present work has as main aim the analysis and the use of an regularization instrument different from that previuosly used and this in order to emphasize typical recurrent behaviours of their annual course and possible periodicities.

° Work carried out in the frame of the P. F. "INFORMATICA "Subprog.P2 - Territorio under the grant 82.00804.97

2. DATA

The present paper bases oneself on the monthly totals of the precipitations deduced from the daily values measured continuously and regularly at the Observatory of the Genoa University since from 1833 ; the data here used cover the period January 1833-December 1978. The daily values of our pluviometric series with the informations about their precision and about the Observatory were reported in a previous work (FLOCCHINI et. Al., 1982).

The monthly values of rainfall show oscillations of annual duration around the average value and also considerable advances or delays, variations from year to year in the position and intensity of the rainy periods. In average annual cycles are present but the periods and the intensity seem to be strong " irregular ".

3. MATHEMATICAL MODEL

The ordinary moving averages in presence of strong irregularities in the length of periods and in intensity of the phenomena for every value of the order, $\underline{k}$, supplie new more significant series becoming flat with increasing $\underline{k}$. The algorithm of the moving averages, with fixed degree 1 and variable order $\underline{k}$, which we have marked M_k^1 should lead to find the periods present in a time series having "regular" oscillating components. Therefore, although the ordinary moving averages allows an easy reading of serie's data , they cannot by applied to our series that by this algorithm is not regularized in satisfactory way.

The algorithm here proposed was stimulated by previous studies of economic and social character (WITTSTEIN , 1867 ; GINI et GALVANI, 1938). The applied mathematical

Instrument keeps the advantage of the ordinary moving averages and in the same time it is able to consider in different ways the values which are closer and more distant from the time refered to be the determination of regularized value. It bases oneself on the moving averages with fixed order $\underline{k}$ and variable degree p: $M_k^{(p)}$.

Generally, the equalized process of the moving averages with $\underline{k}=2s+1$ order repeated to the degree $\underline{p}$ generates a time series $M_k^{(p)}$ the elements of this new series are defined in repeated way by the following relation:

$$a_{k,i}^{(p)} = \frac{a_{k,i-s}^{(p-1)} + \ldots + a_{k,i}^{(p-1)} + \ldots + a_{k,i+s}^{(p-1)}}{2s+1} \qquad (1)$$

From the analysis of the relation (1), it appears that, according to the order $\underline{k}$, the $\underline{i}$ term and their adjacents are weighted on the basis of the coefficients depending from the iteration's degree. The Figg. 1a and 1b show the behaviour of the coefficients for $\underline{p}$= 1,2,....,8 and p = 9,10,..., 12 respectively.

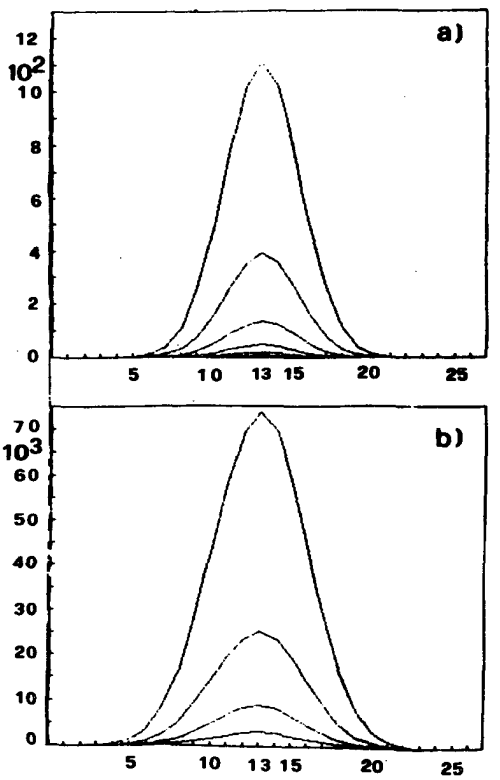


Fig.1 : a) course of the coefficients for p=1,2,....,8
b) course of the coefficients for p=9,10,...,12.

The number of the elements of the originary series necessary to determine the elements of the new series of $\underline{p}$ degree are reported in abscissa while the weight of the coefficients is reported on the ordinate. All the maxima correspond to the same value of the abscissa, that is given by the average value of the variation's field ; the differences among the maxima of the ordinate increase strongly with $\underline{p}$ increasing The behaviour of the curves is similar to a gaussian function :

$$f_i(x) = H.e^{-Kx^2},$$

that for $\underline{i}$ =1,2,...,8 and $\underline{i}$=9,10,.....,12 is reported in Fig.1c and Fig.1d.

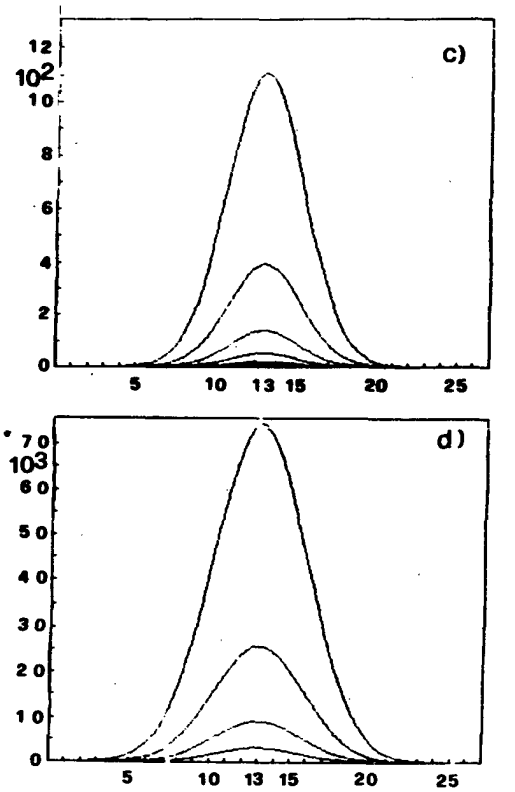


Fig.1: c)course of the theoretical function $f_i(x)$ for $\underline{i}$=1,2,.....,.,8
d)course of the theoretical function $f_i(x)$ for $\underline{i}$=9,10,.....,12.

In order to derive the analytical expression of the curve approaching the behaviours of the coefficients, we have unified the fields and we have determined the H_i and K_i values of the function :

$$f_i(x) = H_i.e^{-K_i(x-a)} \qquad (2)$$

where : $\underline{a}$ is the abscissa of the maxima ;

the translation leads the symmetry axis from the origin to the a point.

If we denote the coefficients of $M_3^{(p)}$ with $c_{i,r}$, we can define the twelve functions :

$$g_i(x) = \begin{cases} c_{i,r} & \text{if } x = (r+12)-i \\ 0 & \text{if } x \neq ((r+12) - i \end{cases} \quad (3)$$

where : r is the number of the iesim iteration coefficients and i = 1,2,....,12.

These functions describe the course of the coefficients and they line the maxima along the single symmetry axis. The values of $g_i(x)$ functions make up the elements of the A(12,25) matrix.Because the H_i must form the maximum ordinates of the $f_i(x)$ functions, it is :

$$H_i = g_i(13) = A(i,13) \text{ with } i = 1,2,..,12 \quad (4)$$

By imposing the crossing x=12,the system :

$$\begin{cases} H_i = A(i,13) \\ H_i . e^{-k_i (12-13)^2} = A(i,12) \end{cases} \quad (5)$$

grows into :

$$\begin{cases} H_i = A(i,13) \\ k'_i = \lg \dfrac{A(i,13)}{A(i,12)} \end{cases} \quad (5')$$

The functions :

$$f_i(x) = A(i,13).e^{-\lg(\frac{A(i,13)}{A(i,12)})(x-13)} \quad (6)$$

reported in Fig.1c and Fig. 1d are really the expressions approaching the functions $g_i(x)$.

To verify wether the courses of the coefficients are gaussian,we have made the hypothesis that the area delimited by the curves is of unit value.By normalizing the $c_{i,r}$ coefficients it is :

$$c'_{i,r} = \frac{g_i(x)}{3^i} \quad \text{with} \quad \begin{matrix} r = 1,2,..,25 \\ i = 1,2,..,12 \end{matrix} \quad (7)$$

and it follows :

$$\sum_{r=1}^{25} c'_{i,r} = 1 \quad \text{with } i = 1,2,..,12 \quad (8)$$

To find if the so determined curves approximate to a limit gaussian,giving rise to a selected behaviour of the weights around the mean point,the procedure of running means was iterated to obtain a significant behaviour of the normalized coefficients for the degree 1,2,3,.....,120 .However, these weights do not have a limit curve ; the maxima of the coefficients approach together.

In order to demonstrate the inexistence of a limit curve ,not zero,we shall investigate the succession of the normalized abscissas of the flex points . Considering the very good approximation of the gaussian functions to the behaviours of the coefficients, we have determined the abscissas of the flex points on the basis of the gaussian functions.

If A is the matrix including the normalized coefficients and k is the position of the central column of this matrix,it is

$$f_i(x) = A(i,j_k).e^{-\lg(\frac{A(i,j_k)}{A(i,j_{k-1})})(x-k)^2} \quad (9)$$

or :

$$f_i(x) = a.e^{-b((x-\alpha)^2} \quad (9')$$

By deriving the equation (9'),we obtain:

$$f'_i(x) = -2ab.\ e^{-b(x-\alpha)^2}$$

$$f''_i(x) = 2b(x-\alpha)^2 - 2ab.e^{-b(x-\alpha)^2}$$

The abscissas of the flex points are given by :

$$x = \alpha \pm (1/2b)^{1/2} \quad (10)$$

The abscissas of the right flex points are therefore :

$$x_{di} = k + (1/2\lg(A(i,j_k)/A(i,j_{k-1}) \quad (11)$$

The abscissas normalized relating to the abscissa of the central maximum,k, are :

$$x_{dNi} = x_{d_i} - k = (1/2\lg(A(i,j_k)/A(i,j_{k-1})^{1/2} \quad (12)$$

The succession of the values defined by the

equation (12) is reported in comparison with the experimental data in Fig.2.

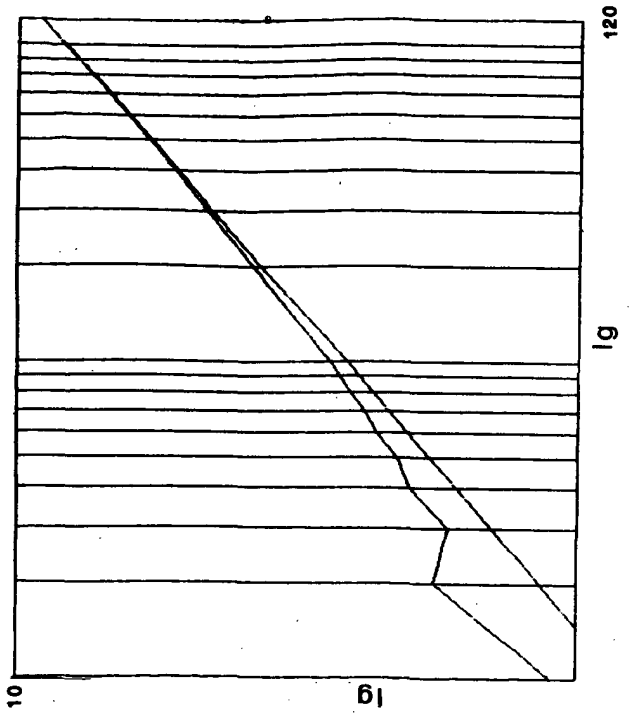


Fig.2 : Diagram pertinent to the course of the abscissas normalized to the right flexes and course of the function :

$$y(x) = 0.822481(x)^{1/2}$$

Except the beginning irregularity ,the behaviour is almost straighforward ;the initial irregularity is mainly due to the strong variation that show the abscissas of the flex points relating to the distribution curves by going from 2th to 4th degree.

The best approximating curve can be symbolized with the following relation :

$$y(x) = C\,x^{a} \qquad (13)$$

The constants C and a were determined by imposing to the equation (13) the crossing the points M (20 .00 ; 3.80) and N(100.00 ;8.23).It has been obtained :

$$\lg 3.80 = \lg C + a \lg 20$$

$$\lg 8.23 = \lg C + a \lg 100$$

By resolving, it is :

$$a = \lg(3.80/8.23)/\lg 8.23 = 1/2$$

$$C = 0.822481$$

The theoretical functions is :

$$y(x) = 0.822481\ (x)^{1/2}$$

The Fig.2 emphasizes also that the succession of the normalized abscissas of the flex points after a certain value superposes on the function y(x) ; the tangent increase with x increasing . This fact shows that a divergency exists in the characters of the normalized abscissa's succesion of the flex points ;this is equivalent to demonstrate by the computer the non-existence of a limit curve not zero.

As far above shortly exposed,it ha been pointed out that the equalization process with moving averages $M_k^{(p)}$ having a fixed order k and variable degree p generates successions of values with fluctuations and so on decreasing as p increasing around the mean value of the averages of the single months ; moreover,with p increasing it approaches therefore the same succession to which is tending the equalization precess $M_1^{(n)}$ with n increasing.

Now, it is necessary to choose the number of the order k and the perequation degree p more suitable for our purpose.Because the data are related to precipitations , the choiche narrows to the values of k = 3 or k =5;in the first circumstance the mediation is influenced by the values directly preceding and successive to which into consideration and in second circumstance by the adjacents values ; for k = 5 regular courses form already at low values of p .This fact should support the choiche of k =5 for the moving averages but the seasonal character of the precipitations in our area ,choiches k = 3.Moreover,if we consider the succession $M_{12}^{(3)}$ and we allow the annual totals of the precipitations,we notice that these values are very near to the measured annual values.

Finally,we remark that the value p =12 supplies an investigation of the i value involving the influence of the twelve preceding and successive data with particular importance for the directly preceding and successive value , which involve "weights" higher fifty per cent of the central term.

4.CHECKING

In order to check the likelihood of the mathematical model above shortly described for the emphasis of possible periodicities

and typical recurrent behaviours, we have computed the monthly totals and the values of the time series obtained by the process $M_{12}^{(3)}$; their courses are reported in Fig.3

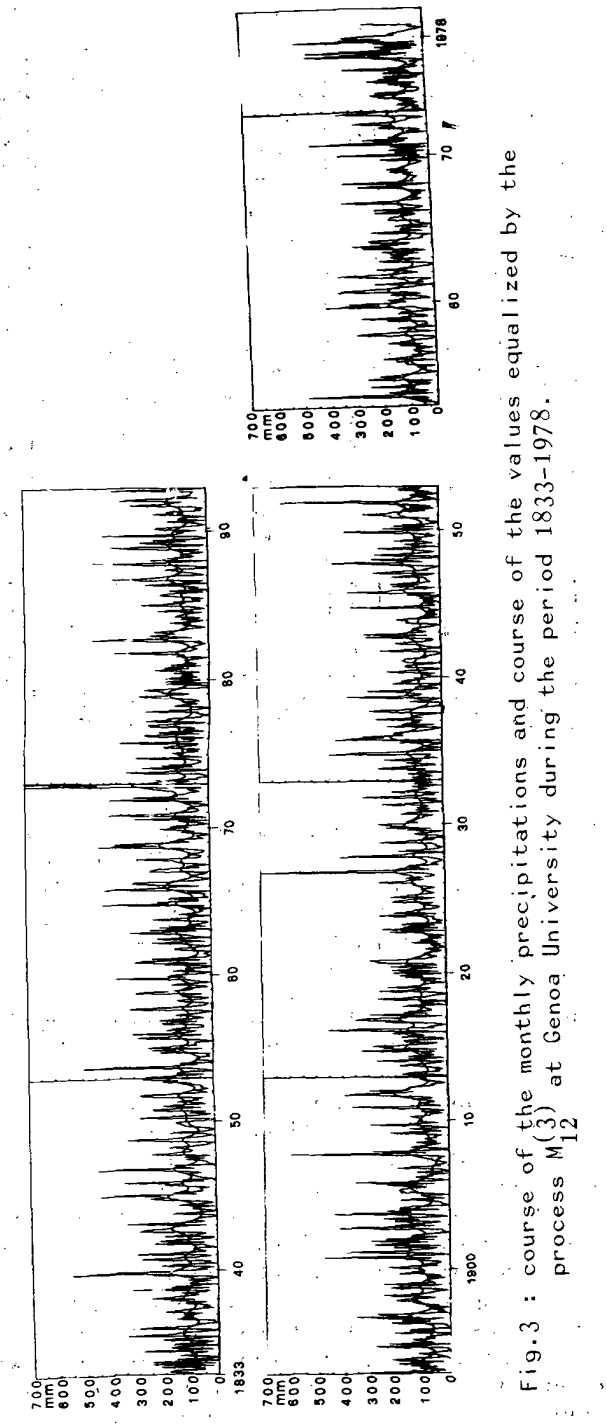


Fig.3 : course of the monthly precipitations and course of the values equalized by the process $M_{12}^{(3)}$ at Genoa University during the period 1833-1978.

Fig.3 shows the effect of the perequation; strong precipitations occurred in a narrow interval of time are collected in a single peack. Indeed, the initial data become more smoothed and some typical behaviours appear ; in average, there are present also cycles with 11 and 22 year periods. The annual precipitations show relative variations of about 10-15 % and therefore few years are reciprocally comparable ; the same average year obtained from the monthly totals for every month over the whole period is not comparable below the 20 % of relative mistake with larger considered annual course.

The investigation of the curve reported in Fig.3 and related to the equalized data suggests the existence of typical courses in the equalized time series and related frequency. On the basis of the running averages with 3th order and 12th degree, we have compared the different years and with the mean year inside a deviation of the 15 %.

If we indicate with $f(x_i)$ the functions giving the mean course of the data, as equalized with the 12th order and 3th degree and having the values reported in the following list

Month	I	II	III	IV	V	VI
P(mm)	115.7	109.2	101.7	95.2	91.2	90.8

Month	VII	VIII	IX	X	XI	XII
P(mm)	94.5	101.3	109.5	116.5	120.5	120.4

and with the $A=f(x_i)$ = 1266.7 mm the sum of the ordinates of the function and if we call with $g_k(x_i)$ the functions giving the annual variations obtained with the series $M_{12}^{(3)}$ and with A the related sums of the ordinates, we have determined the k values for which it is :

$$\frac{\left| g_k(x_i) - f(x_i) A_k / A \right|}{g_k(x_i)} < \varepsilon$$

By putting ε =0.15 , we have 37 values, which represent about the 25 % of the total. The small number of these values suggests the existence of annual courses different among them and from the mean.

The research of these typical waveforms

has been made by an algorithm aggregating on and so on the disconnected classes of the annual behaviours like among them.This algorithm averages also the behaviours of a single class. The annual courses of the period are classified in seven disconnected classes covering only 101 years.

The frequency of the different types of annual course are reported in the following list :

Type	Mean	1	2	3	4	5	6	7
n°	37	7	7	17	11	32	21	6

From this table it appears that there exists some types of annual variations and that the average year does not exists as a typical year. This syntetic year appears therefore as an mathematical abstraction

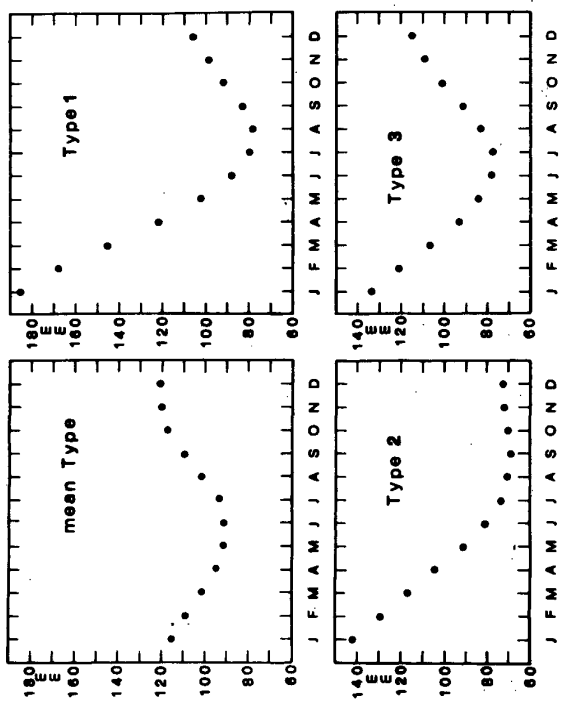


Fig.4a : Mean annual course of the precipitations at Genoa University and courses of the 1th, 2th, and 3th type ;values related to the period 1833-1978.

and it seems to disavow the existence of a typical behaviour in the monthly precipitations at Genoa ; very frequent are also the types 5,6 and 3.
The different types of annual courses are displeyed in the Figg. 4a and 4b.

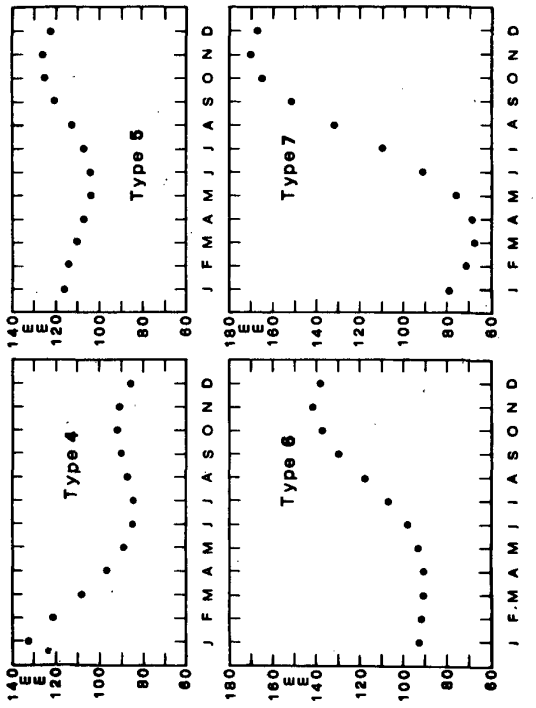


Fig.4b : Mean annual course of the precipitations at Genoa University for the 4th, 5th, 6th and 7th type;Values related to the period 1833-1978.

The 3th and 5th types are similar one to another and to the syntetic year; the 3th type shows the minimum of precipitations in June and july and the maximum in january.The type 5 has the minimum in May-June and the maximum in October-November.
The mean type has the minimum in May-June and the maximum in October-November ; the annual range is minimum for the 5th type and maximum for the 3th type.

A comparison with the solar activity as expressed by sunspots number , shows a peculiar connection among the periods with reduced number of years of the 3th and 5th type.

On this basis we have divided the whole period in four part , according to the level of solar activity and we have

applied the χ^2 test to the distribution of the typical years ; the results of this investigation show that a probability less than 0.5 $^{o}/_{oo}$ exists that such a distribution arises from randoom fluctuation . In this respect, it seems that the type of ye ar is a very sensible indicator of the solar activity.

5.CONCLUSIONS

The proposed model gives a possible answer to the problem of individuate isolating cycles present in the time series only in average. The used technique allows an easy management on time series also of great lenght operating for every value only a segment of the time series also by a desk computer ; moreover, it weights in different way the more accumulated and the more dispersed values to the time to which is the determination of the regularized value of the relatedtime series.

The i term of the original time series and their adjacents are "weghted" by coefficients that depend on the iteration degree.The investigation of the coefficient distribution emphasizes that this distribution for $p > 1$ gives greater "weight" to the values near to the time i and less to the distant values ; moreover, for p sufficiently great ($p > 5$) ,the distribution of the coefficienty is substantially gaussian. For $p = 12$,the i value is influenced by the 12 and 12 successive values with particular importance of the 3 preceding and the 3 successive values ; these have "weights" higher of the 50 % than the wight due to the coefficient of the central term.

The choiche of the iteration degree has been made not deductively with a priviliged value but by optimization.
Suc an instrument applied to the monthly sums of the pluviometric series of Genoa University by using the 12th degree and and the 3th order has emphasized that strong precipitations near between them are "collected " within a single peak; moreover, this instrument emphasizes in our pluviometric time series the presence of typical recurrent behaviours and of regularities that,in average, follow one to another with 11 and 22 year period.

REFERENCES

Bossolasco,M.,et Al.,1971 : Le cause meteorologiche dell'alluvione su Genova del 7-8 Ottobre 1970.Geofis. e Meteor.,XX,122-132.

Dagnino,I.,G.Flocchini and C.Palau,1978 : Le precipitazioni del 6-10 Ottobre 1977 sulla Liguria Centrale.Inst.Suisse de Mètèorologie.,267-270.

Flocchini,G. et Al.,1981 : A Mathematical Model of the Field of Precipitations. Il Nuovo Cimento, Ser.1, 4C, 673-681.

Flocchini,G. et Al.,1981 : Il Regime pluviometrico di Genova.Quaderni del C.N.R.,AQ/5/31,1-40,Roma.

Flocchini,G. et Al.,1982 : I dati pluviometrici della serie storica (1833-1980) di Genova.Quaderni del C.N.R. AQ/5/39,1-132,Roma.

Gini,C. and L.Galvani,1938 : Tavole di mortalità della popolazione italiana.Annali di Statistica,Ser.VI,VIII,

70-310.

Schönwiese, C.D., 1974 : Schwankungsklimatologie im Frequenz- und Zeitbereicht. Wiss. Mitt. Meteor. Institute, 24, 1-115.

Schönwiese, C.D., 1975 : Bemerkungen zur numerischen Bandpass-filterung von Zeitreihen unter besonderer Berücksichtigung meteorologischen Daten. Arch. Met. Geoph. Biokl., Ser. A, 24, 311-325.

Wittstein, T., 1867 : Mathematische Statistik und deren Anwendung auf National-Ökonomie und Versicherungwissenschaft. Mem. Mat., 288, 30-40.

STATISTICAL APPROACH TO CLIMATOLOGICAL NETWORKS OPTIMIZATION

Guy der Megreditchian

Météorologie Nationale

EERM-CRMD - Paris, France

1. INTRODUCTION

The problem of rational planification of a network of climatological stations is extremely complex and involve numerous aspects issuing even from the same finality of its existence. What is the purpose for the creation of such networks, what kind of information it has to give ? And that is the reason for which the optimisation problem may be approached through many points of views : economical, aeronautical, administrative, forecasting, military.

In order to counciliate such various points of views, we will perform a procedure of optimization of the "informativity" which gives the network about the concerned field. The basic principle for an adequate choice of the network is that the latter gives the most faithfull image of the field.

We will consider the field as a statistical object thanks to the deep analogy between the process of recollecting the observations and their generation in a random population. In fact, we will give to the optimization procedure a statistical meaning based on the spatio-temporal statistical characterisation of meteorological fields. That's why it seems useful to remind briefly the main principles of the definition of the statistical structure of meteorological fields.

We will consider here four types of problems linked with the optimisation of climatological networks and will present adequate methods of resolution and also some concrete applications we have already treated.

The first problem is the preferential ranking of the existing stations network. In case of some budget restrictions for example, it's necessary to know which station (or stations) must be closed in order to minimize the loss of information regarding the whole field.

The second problem concerns the optimal localisation of the new stations, which must be placed in the most efficient way, in order to get the maximal increase of information about the field.

The third problem is the choice of the optimum mesh of the regular grid used for the description of the field. How should be realized the best compromise between exhaustivity and economy ?

The fourth problem is that of the optimal definition of the unit in the measuring a meteorological variable. In other words, how to perform the discretisation of a continuous variable without altering its continuous structure ?

Of course, many other questions may arise in connection with network optimisation, but we will consider only those which have just been mentioned above.

2. ECONOMICAL EFFECT OF NETWORK OPTIMIZATION

The optimisation procedure must be based on the definition of some appropriate criterion Such a criterian mut express the compromise resulting from a contradictory situation. The insufficiency of climatological information leads to some losses P for the country's economy (human lives, harvest destruction , transport difficulties and so on) which will decrease with the rationalisation of the climatological observation system. On the other hand, the expenses associated with the maintenance of climatological network will rise together with the increase of the number of stations.

A reasonnable economical criterion may be defined as the sum S of the C cost of the maintenance expenses and the P losses due to our ignorance of the meteorological conditions :

$$S = P + C$$

Thus, the network optimization corresponds then to the research of the network for which the value of the quantity S is minimum :

$$\min S$$
$$\forall \text{ Network.}$$

Nevertheless, it's easy to conceive the pratical difficulties concerning the concrete realisation of that optimization procedure based on an economical criterion in the actual state of our knowledges ; thus, we will mainly consider the network optimization procedure in a more restrictive climatological point of view.

3. THE STATISTICAL STRUCTURE OF METEOROLOGICAL FIELDS

The meteorologist is concerned, not only with the statistical behaviour of individual meteorological elements, but also by their reciprocal influences, their statistical interdependencies. Generally, they are described by the crossing moments of the second order, the definition of which is based on the introduction of an operator of mathematical expectation :

$$E[\psi(X)] = \int_{-\infty}^{\infty} \psi(Z)\, dF_X(Z)$$

which is defined for every scalar function $\psi(X)$ of the random vector X.

4. COUPLE OF RANDOM VARIABLES

To characterize the degree of statistical cohesion between the random variables x and y, the following parameters are introduced :
The non centered coefficients of covariance $s(x,y)$ and structure $\delta(x,y)$:

$$s[x,y] = E[x \cdot y] \qquad \delta[x,y] = E[x-y]^2$$

The centered coefficients of covariance $v(x,y)$ and structure $\gamma(x,y)$

$$v[x,y] = E[\mathring{x} \cdot \mathring{y}] \qquad \gamma[x,y] = E[\mathring{x}-\mathring{y}]^2$$

The normalized and centered coefficients of covariance $r(x,y)$ (correlation coefficient) and structure $g(x,y)$

$$r[x,y] = E[\tilde{x} \cdot \tilde{y}] \qquad g[x,y] = E[\tilde{x}-\tilde{y}]^2$$

When we have a sample of T observations
$x(1), \ldots, x(t), \ldots, x(T)$
$y(1), \ldots, y(t), \ldots, y(T)$
it's possible to compute the corresponding empirical coefficients by means of similar formulas introducing as before the corresponding empirical operator of mathematical expectation

$$\hat{E}[\psi(X)] = \frac{1}{T} \sum_{t=1}^{T} \psi[X(t)]$$

Then, we obtain : $\hat{s}[x,y] = \hat{E}[x \cdot y]$, $\hat{\delta}[x,y] = \hat{E}[x-y]^2$
$\hat{v}[x,y] = \hat{E}[\mathring{x} \cdot \mathring{y}]$, $\hat{\gamma}[x,y] = \hat{E}[\mathring{x}-\mathring{y}]^2$
$\hat{r}[x,y] = \hat{E}[\tilde{x} \cdot \tilde{y}]$, $\hat{g}[x,y] = \hat{E}[\tilde{x}-\tilde{y}]^2$

We can remark here that the covariance coefficients are similarity indexes, while structure coefficients are dissimilarity indexes. From terminological point of view, let's precise once more that the coefficients r (x, y) are usually called correlation coefficients and γ (x, y) variograms.

Some obvious relations link these covariance and structure coefficients in the case of homogenous and isotropic fields :

$$v[x,y] = s[x,y] - E(x) \cdot E(y) \qquad \gamma[x,y] = v[x,x] + v[y,y] - 2v[x,y]$$
$$\gamma[x,y] = \delta[x,y] - [E(x) - E(y)]^2 \qquad g[x,y] = 2[1 - r(x,y)]$$

Each of these coefficients expresses in its own way the intensity of the statistical liaison between x and y variables.

Random vector

Let's take $X = (x_1, \ldots, x_n)$ and $Y = (y_1, \ldots y_m)$ two random vectors. The statistical interliaisons between their components will be described by means of the matrixes formed by the corresponding scalar coefficients for each pair of x_i and y_j :

$$S_{xy} = (s[x_i, y_j]) \qquad \Delta_{xy} = (\delta[x_i, y_j])$$
$$V_{xy} = (v[x_i, y_j]) \qquad \Gamma_{xy} = (\gamma[x_i, y_j]) \qquad R_{xy} = (r[x_i, y_j]) \qquad G_{xy} = (g[x_i, y_j])$$

In particular, if $X = Y$, the statistical cohesion between the vector components may be described by means of the following matrices :

$$S_{xx}, V_{xx}, R_{xx}, \Delta_{xx}, \Gamma_{xx}, G_{xx}$$

The definition of the corresponding empirical matrixes $\hat{S}_{xy}$, $\hat{V}_{xy}$, $\hat{\Delta}_{xy}$, $\hat{\Gamma}_{xy}$, $\hat{G}_{xy}$, $\hat{R}_{xy}$ is based on the same principle and the corresponding computations may be realized if the samples of T observations are available.

$$\begin{matrix} X(1), \ldots, X(t) \ldots X(T) \\ Y(1), \ldots, Y(t) \ldots Y(T). \end{matrix}$$

Thus, we have a lot of adequate parameters for the description of statistical interliaisons between the components of random vectors.

Random fields

It is difficult to characterize in an exhaustive way the statistical dependence between the values of random fields observed at different points because it is necessary to use the joint distribution functions of high order. That's why generally the study is limited by the definition only of the covariance and the structure functions of those fields.

Let's introduce the following notations : $x(t, P_i)$ the value of the meteorological field X observed at the time t, at the point P_i and $y(\tau, R_j)$ the value of the meteorological fiels y observed at the time τ at the point P_j.

The spatio-temporal statistical structure of the fields x and y is then described by the spatio-temporal covariance $v_{xy}(t, \tau, P_i, P_j)$ and structure $\gamma_{xy}(t, \tau, P_i, P_j)$ functions defined accordingly to the corresponding definitions of the same coefficients for a pair of random variables by means of the following relations : $v_{xy}[t, \tau, P_i, P_j] = v[x(t, P_i), y(\tau, P_j)]$

$$\gamma_{xy}[t, \tau, P_i, P_j] = \gamma[x(t, P_i), y(\tau, P_j)].$$

From such general definitions, we obtain in particular that of spatial covariance $v_{xy}(t, t, P_i, P_j)$ and structure $\gamma_{xy}(t, t, P_i, P_j)$ functions putting $t = \tau$ and also the definition of temporal covariance $v_{xy}(t, \tau, P, P)$ and structure $\gamma_{xy}(t, \tau, P, P)$ functions putting $P_i = P_j = P$.

In a similar way, we can introduce normalized spatio-temporel covariance $r_{xy}(t, \tau, P_i, P_j)$ and structure $g(t, \tau, P_i, P_j)$ functions, and their purely spatial $r_{xy}(t, t, P_i, P_j)$, $g_{xy}(t, t, P_i, P_j)$ or purely temporal $r_{xy}(t, \tau, P, P)$, $g_{xy}(t, \tau, P, P)$ expressions.

We will distinguish also the cross-covariance or cross-structure functions if $x \neq y$ and autocovariance or autostructure functions if $x = y$.

The definition of the statistical structure of random fields is mainly simplified if the fields are homogeneous and isotropic, that means a fact that their spatial characteristics associated with the point R of the field are constants for each point R, and their spatial characteristics associated with a pair of points P_i and P_j are depending only from the distance

between these points.

We will have, for the homogeneous and isotropic fields, the following relations :

$$m_i[x(p)] = m_i(x) ; \sigma[x(p)] = \sigma[x] \qquad \gamma_x[p_i,p_j] = \gamma_x[d_{ij}] ; g_x[p_i,p_j] = g_x[d_{ij}]$$
$$v_x[p_i,p_j] = v_x[d_{ij}] ; r_x[p_i,p_j] = r_x[d_{ij}]$$

In the case of temporal structure, the notions of homogeneity and isotropy are replaced by the notions of stationarity imposing the following relations :

$$m_i[x(t)] = m_i[x] ; \sigma[x(t)] = \sigma[x] \qquad \gamma_x[t_i,t_j] = \gamma_x[\tau_{ij}] ; g_x[t_i,t_j] = g_x[\tau_{ij}]$$
$$v_x[t_i,t_j] = v_x[\tau_{ij}] ; r_x[t_i,t_j] = r_x[\tau_{ij}] ; \qquad \text{où} \quad \tau_{ij} = t_i - t_j$$

In practical applications, the patterns of the points corresponding to the values of these functions for given d_{ij} or τ_{ij} are adjusted by means of adequate theoretical functions depending of one or several parameters.

5. ABOUT THE NOTION OF INFORMATIVITY OF CLIMATOLOGICAL OBSERVATION NETWORK

Let's consider an adequate networks of observation stations for the study of some meteorological element. Let Γ_i be the station number "i" of the network,

$$(\Gamma_i)_{i=1}^{N}$$

the network of N stations, $\xi_i(t)$ the value of the anomalie of the meteorological element observed of the station Γ_i at the time t.

The meteorological field of anomalies of such meteorological element defined by the N values $\xi_i(t)$ observed in these N stations at the time t will be described by the N dimensional vector :

$$\xi(t) = \{\xi_1(t), \ldots, \xi_N(t)\}$$

Let's choose a stochastic model of the natural mechanism generating the observations, according to which the anomalies observed at the time t are the result of a random sampling realised in a statistical population defined by some probabilistic measure P. Then, the statistical structure of the system of random variables $\xi_i(t)$ is defined by the joint distribution function :

$$F_{\xi(t),\ldots,\xi(\tau)}[X(t),\ldots,X(\tau)] = P[\xi(t) < X(t), \ldots, \xi(\tau) < X(\tau)]$$

where

$$X(t) = \{x_1(t), \ldots, x_N(t)\}$$

This distribution function determine completely the quality of the corresponding network of observational stations in the frame of adopted model.

Let us study in particular the correlational redundancy of the network.

The statistical interdependances between random variables $\xi_i(t)$ and $\xi_j(t)$ are characterized by the covariance coefficients :

$$v_{ij}(t,\tau) = E[\xi_i(t)\,\xi_j(\tau)],$$

or for the normalized variables by the correlation coefficients

$$r_{ij}(t,\tau) = \frac{v_{ij}(t,\tau)}{\sqrt{v_{ii}(t)\,v_{jj}(\tau)}}$$

Our knowledge of the spatio-temporal interdependences between random variables $\xi_i(t)$ is characterized by the matrix :

$$V(t,\tau) = (v_{ij}(t,\tau)) \qquad R(t,\tau) = (r_{ij}(t,\tau))$$

Now let's impose two important restrictions. First, we will consider a gaussian model and secondly a stationary model. Otherwise, the density probability function of the vector of anomalies can be expressed by the following formulae :

$$f_\xi(X) = \frac{1}{(2\pi)^{N/2}|V|^{1/2}} e^{-\frac{1}{2}X'V^{-1}X}$$

where V designes the spatial covariance matrix of the field and $E(\xi) = 0$, as we consider the vector of anomalies.

Now let's introduce a very important notion of "equivalence" (in anomalicity) between two observational networks $(\Gamma_i)_{i=1}^{N}$ and $(\gamma_j)_{j=1}^{n}$ measuring the values of the same meteorological field. For this purpose, we are going to use an important characteristic of the observed field, the anomalicity coefficients K and $\dot{K}$:

$$K = \frac{1}{N}\sum_{i=1}^{N}\xi_i^2 \quad \text{ou} \quad \dot{K} = \frac{1}{N}\sum_{i=1}^{N}\frac{\xi_i^2}{v_{ii}}$$

which describe the field variability in regard to its mean behaviour. We will point out that the coefficients are associated to the network $(\Gamma_i)_{i=1}^{N}$ generating the observations $x_i(t)$ by introducing respectively the following notations :

$$K(\Gamma_N), \dot{K}(\Gamma_N), K(\gamma_n), \dot{K}(\gamma_n)$$

So we can now define the notion of equivalence in anomalicity between two networks $(\Gamma_i)_{i=1}^{N}$ and $(\gamma_j)_{j=1}^{n}$.

Definition 1

We say that the networks $(\Gamma_i)_{i=1}^{N}$ and $(\gamma_j)_{j=1}^{n}$ are "equivalent" in anomalicity if the distribution functions of their anomalicity coefficients are equal :

$$F_{K(\Gamma_N)}(x) = F_{K(\gamma_n)}(x)$$

In other words, we will consider that two observational networks are equivalent if the statistical characterisation of their variability in regard to their mean behaviour is stochasticaly identical. We may say that this definition of "equivalence" is given in its strong sense. It's naturely too restrictive and it's necessary to give a less restrictive definition that we will call an "equivalence" in the weak sense.

Definition 2

We will say that the networks $(\Gamma_i)_{i=1}^{N}$ and $(\gamma_j)_{j=1}^{n}$ are "equivalent" of the order "s" in the weak sense if the first "S" moments or cumulants ν_j of random variables $K(\Gamma_N)$ and $K(\gamma_n)$ are equal. In other workds, the two networks are equivalent if, for j = 1,S, we have the equality :

$$m_j[K(\Gamma_N)] = m_j[K(\gamma_n)] \quad j = 1,s \qquad \nu_j[K(\Gamma_N)] = \nu_j[K(\gamma_n)] \quad i = 1,s$$

The theory of the distribution of quadratic forms from normal random variables allow us to get the characteristic function of anomalictiy coefficients and therefore the explicit from of their cumulants :

$$F_K(x) = 1 - \sum_{i=1}^{N} \left(\prod_{\substack{j=1 \\ j \neq i}}^{N} \frac{\lambda_j}{\lambda_j - \lambda_i} \right) e^{-\frac{x}{\lambda_i}}$$

It's not possible to compute directly the corresponding density probability function by direct application of the inversion formulae, but we can get an approximate formula (when N is odd) :

where λ designes the eigenvalues of the matrix V. Then, we can define the return period T (X) of an anomalicity of given strengh ($>$ x) :

$$T(x) = \frac{1}{1 - F_K(x)}.$$

We can now use the introduced notion of the equivalence of two networks to get a new definition ; the concept of the number of statistically independant observations n_{equ} (t) "equivalent" in the defined sense to the N correlated observations. Here, in fact, it's a matter of some fictive independent observations ; the most important fact is that this number allows to get an objective measure of the quality of the network from which the observations are obtained. Such a quality concerns the localization (or the choice of the time sampling) of corresponding stations where the observations are realised.

In fact, it makes possible the elimination of network correlational redundancy. For that, we suppose that the network $(\gamma_j)_{j=1}^{n}$ is characterized by the covariance matrix $V = \sigma^2 I_n$, where I_n is the unit matrix of order M.

In such a way, we get a good-looking formula obtained from the definition of the equivalence of order 2 in the weak sense :

$$n_{equ} = \frac{[tr\, V]^2}{tr\, V^2}$$

where $t_r(V)$ signifies the trace of the V matrix. When we use for the notion of networks' equivalence the normalized anomalicity coefficient K, we obtain the following very simple formula :

$$n^*_{equ} = \frac{[tr\, R]^2}{tr\, R^2}$$

We can easily verify that in trivial cases such a definition is reasonable. Thus, we have

if $R = I_N$: $n^*_{equ} = N$

if $R = U_N U_N'$ $n^*_{equ} = 1 \quad (U_N = \{1 \ldots 1\})$

Furthermore, if the last eigenvalues of the matrix V are 0, that is the case when "K" linear relations exist between the corresponding components of the vector X, we have the relation :

$$n_{equ} < N - k$$

It is interesting to point out that the notion of equivalence of order 3 will lead to the similar formula :

$$n_{equ} = \frac{(tr\, V^3)^2}{(tr\, V^2)^3}$$

We will use the introduced notion of equivalent number of independent observations to resolve various problems, related with some aspects of network optimization. Here are some examples.

The fig. 1 shows the temporal autocorrelation function of daily minimum temperature for Bordeaux stations.

On the fig. 2 we give the temporal autocorrelation coefficients of daily minimum and maximum temperatures for a time lag from 0 to 30 days, and the corresponding equivalent number of independant observations (the last row of the table).

It is possible, as well, to make a correction to the definition of the number of equivalent independant observations in the case when some measurement errors are altering the vector X of observations.

In the classical hypothesis, when the observed vector $\hat{X} = X + \varepsilon$ has a normal distribution and the ε errors are independant from each other and also from the X, it is possible to obtain a simple relation between the real equivalent number n_{equ} and its empirical analoguous $\hat{n}_{equ}$ calculated from the matrix $\hat{V} = V_{\hat{X}\hat{X}}$:

$$n_{equ} = \frac{\hat{n}_{equ} N}{[N - \hat{n}_{equ}][1 + q]^2 + \hat{n}_{equ}}$$

where $q = \frac{\sigma_\varepsilon^2}{\sigma_x^2}$ characterizes the noise-signal ratio.

6. RATIONAL LOCALISATION OF THE STATIONS BASED ON THE MINIMISATION OF THE NETWORK CORRELATIONAL REDUNDANCY

Two problems arise according to the fact that we want to eliminate the most redundant stations or, on the contrary, to add globally the most interesting station for the knowledge of the meteorological field.

It is possible to solve the first problem by means of a backward stepwise procedure, based on the minimisation of the network correlational redundancy. For that we use the empirical correlation matrix $\hat{R}$ calculated from the observed sample.

At the first step, we search the less interesting station in the following way :

We compute for each subset of N-1 stations the equivalent number of independant stations n_{equ} (i) , where "i" is the number of the station we eliminate from the initial set to obtain the subset of N-1 stations.

We search the value "i_1" which satisfy the following condition :

$$n_{equ}(i_1) = \max_i \; n_{equ}(i)$$

The station number i_1 is realy the less interesting due to the fact that the loss of information is minimised when we eliminate it.

At the second (and further) step, we proceed in an analoguous way with the remaining stations after the elimination of the station i_1 and we restart this procedure of preferential ranking.

Thus, we get a preferential ranking of the N stations in the order of increasing interest

$$i_1, i_2, \ldots, i_s, \ldots, i_n.$$

It is also possible to apply a backward stepwise procedure based on the successive choice of the most interesting station associated with the minimization of the equivalent number :

$$n_{equ}(j_1) = \min_i n_{equ}(i)$$

On the fig. 3, we show an example of the utilization of such procedure for ranking 48 French stations, describing the field of atmospheric ground pressure obtaining by means of a backward stepwise selection.

The result seems to be logical enough. The 12 best stations are localized on the periphery of France and one on the center. We found more stations in the mountainous regions (Alps, Pyrenees) than in the plains. The 12 following stations fill the empty regions and the last 12 stations are situated in the flat regions. Let's underline that the problem is more complex, when the field is a vectorial one, but it's also possible to propose some reasonable solutions.

The most simple variant consists to associate to each station a rank corresponding to the relative importance of this station obtained for the scalar field corresponding to a given component of the vector field. After this, we define for the vector field a global rank corresponding to the relative importance of given station for the global knowledge of the vector field on the base of a new rank. This new global rank starts from the sum of individual ranks obtained for each scalar field.

On the fig. 4 the geographical position of each station shows with evidence its relative importance for the whole network.

It is also possible to use a more sophisticated method, the so-called "experts method", or "preference analysis" to obtain a global ranking which in some sense for a particular optimisation criterion, gives the best concordance with the whole of particular rankings obtained for each scalar field corresponding to each scalar component of the vector field.

The second problem of the "optimal" localization of a new station ( ) requires the knowledge of the statistical structure of the corresponding meteorological field. In the hypothesis of an homogenous and isotropic field, the networkd optimization depends mainly of its spatial density characterized by the matrix of interdistances between observations stations :

$$D = (d_{ij}) \quad ; \quad d_{ij} = d[\Gamma_i, \Gamma_j]$$

In this case, the spatial autocorrelation function of the field depends only of these interdistances $r[\xi_i, \xi_j] = \psi(\ldots)$

The most frequently met forms of such autocorrelation function are given by the following formulae :

$$\psi(d) = e^{-ad} \quad \psi(d) = [1+ad]\cdot e^{-ad}$$

$$\psi(d) = e^{-a^2d^2} \quad \psi(d) = ad\,K(ad) \quad \psi(d) = (ad)^{\frac{1}{2}} K_{\frac{1}{2}}(ad) \quad \psi(d) = e^{-ad} I_0(ad)$$

$$\psi(d) = \frac{a^2}{[a^2+d^2]^{\frac{1}{2}}}$$

Thus, the equivalent number of independant observations, takes the following from :

$$n^*_{equ} = \frac{N^2}{\sum_{i,j=1}^{N} \psi^2(d_{ij})}$$

which shows clearly the explicit form of its dependency from the network spatial density.

For the optimal localization of a new station, we consider a fictive station Γ_{N+1} at the point defined by the coordinates (x, y) and by the distances $d[\Gamma_i, \Gamma_{N+1}] = \delta_i(x,y)$ of the other stations. The quality of this new system of N + 1 stations is then determined by an equivalent number of independent stations which become now a function of the coordinates s and y :

$$n^*_{equ}(x,y) = \frac{(N+1)^2}{\sum_{i,j=1}^{N} \psi^2(d_{ij}) + 2\sum_{i=1}^{N} \psi^2[\delta_i(x,y)] + 1}$$

Thus, tne optimal localization of a new station leads to the research of a global extremum of this function :

$$n^*_{equ}(x,y)$$

for all the points (x,y) belonging to the considered domain D. In such a way, we search the point (x_o, y_o) which verify the relation :

$$n_{equ}[x_o, y_o] = \max_{x,y \in D} n_{equ}[x,y].$$

Various numerical algorithms to obtain the extremum are available, for example the gradient or conjugate gradient methods. In such cases, the most delicate element is the danger to find a local extremum instead of a global one for the given domain D.

7. NETWORK OPTIMISATION BASED ON THE MAXIMISATION OF MULTIPLE CORRELATION COEFFICIENT BETWEEN THE OBSERVED VALUE ON ONE POINT AND THE VALUE ON THE SAME POINT RECONSTITUTED BY MEANS OF THE SET OF EXISTING STATIONS.

It seems natural to consider that for the preferential ranking of the stations, the less informative station is the one which may be in the best way explicated by the others. In the frame of a multiple linear forecast, the corresponding quality criterion may be the multiple correlation coefficient between the really observed value x_i at this station and the value $\tilde{x}_i$ forecasted by means of the set of the N-1 other stations. We will admit for this multiple correlation coefficient the notation $r^2_{n-1}[x_i, \tilde{x}_i]$

Thus, the aim of the optimisation procedure is to search the index i_1 such as

$$r^2_{n-1}[x_{i_1}, \tilde{x}_{i_1}] = \max_i r^2_{n-1}[x_i, \tilde{x}_i].$$

This maximum is easy to find by means of the inversion of the correlation matrix R_{XX}, which expresses the spatial correlations between the values of the field observed of the existing stations. In fact, it is not difficult to prove the following lemma, which define the foundamental property of the inverse correlation matrix.

Lemma - let be $R_{xx} = (r_{ij})$. $R^{-1}_{xx} = (r^{ij})$, $X[i] = \{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n\}$

Then, we have the following result :

$$r^{ii} = \frac{1}{1 - r^2_{n-1}[x_i, \tilde{x}_i(X[i])]}$$

	Temperatures Mini							Temperatures Maxi						
Δt +1	Bordeaux	Lyon	Le Bourget	Rennes	Strasbourg	Marignane	Orly	Bordeaux	Lyon	Le Bourget	Rennes	Strasbourg	Marignane	Orly
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	.6123	.6249	.6194	.5818	.6523	.5853	.6365	.6799	.7264	.7233	.7150	.7267	.7065	.7219
3	.3966	.4427	.4196	.3704	.4617	.3937	.4326	.4557	.5165	.5145	.5114	.5229	.5137	.5134
4	.2905	.3520	.3115	.2662	.3564	.3072	.3221	.3242	.3986	.3847	.3818	.3945	.3947	.3788
5	.2176	.2843	.2420	.2048	.2764	.2386	.2471	.2500	.3158	.3057	.3046	.3133	.3233	.2966
6	.1555	.2252	.1972	.1500	.2267	.1882	.1961	.1995	.2489	.2526	.2496	.2488	.2699	.2402
7	.1176	.1773	.1527	.1155	.1872	.1505	.1502	.1571	.2042	.2144	.2092	.1973	.2195	.1980
8	.0957	.1438	.1206	.0874	.1582	.1322	.1174	.1340	.1679	.1882	.1837	.1581	.1873	.1686
9	.0818	.1199	.1075	.0815	.1329	.1156	.1039	.1260	.1415	.1648	.1644	.1282	.1665	.1457
10	.0650	.1002	.0933	.0669	.1105	.1070	.0899	.1206	.1289	.1475	.1541	.1105	.1455	.1307
11	.0612	.0796	.0762	.0576	.0940	.0901	.0707	.1007	.1139	.1334	.1429	.0940	.1192	.1201
12	.0526	.0650	.0590	.0375	.0677	.0732	.0515	.0822	.0919	.1149	.1191	.0757	.1016	.1039
13	.0286	.0476	.0435	.0180	.0504	.0722	.0424	.0673	.0776	.0895	.1001	.0681	.0818	.0804
14	.0305	.0363	.0382	.0222	.0357	.0680	.0359	.0585	.0695	.0742	.0903	.0565	.0763	.0675
15	.0243	.0157	.0293	.0177	.0240	.0494	.0235	.0544	.0587	.0676	.0783	.0426	.0613	.0584
16	.0161	.0182	.0158	.0055	.0286	.0501	.0132	.0578	.0466	.0604	.0709	.0350	.0627	.0538
17	.0173	.0134	.0025	.0034	.0189	.0474	.0053	.0554	.0499	.0627	.0733	.0281	.0585	.0561
18	.0143	.0178	- .0008	- .0034	.0294	.0495	.0032	.0538	.0582	.0559	.0700	.0299	.0539	.0510
19	.0262	.0182	.0013	.0040	.0390	.0565	.0027	.0565	.0617	.0573	.0732	.0328	.0601	.0488
20	.0197	.0190	.0073	.0075	.0370	.0567	.0078	.0464	.0596	.0576	.0708	.0310	.0562	.0508
21	.0241	.0218	.0186	.0123	.0372	.0542	.0161	.0450	.0459	.0610	.0769	.0350	.0586	.0538
22	.0255	.0310	.0239	.0145	.0431	.0526	.0153	.0488	.0366	.0551	.0684	.0379	.0534	.0507
23	.0211	.0323	.0262	.0230	.0506	.0476	.0224	.0451	.0324	.0511	.0557	.0392	.0447	.0449
24	.0115	.0310	.0288	.0208	.0520	.0451	.0247	.0316	.0276	.0445	.0561	.0345	.0374	.0371
25	.0090	.0210	.0192	.0225	.0486	.0405	.0232	.0361	.0266	.0449	.0572	.0246	.0312	.0420
26	.0019	.0193	.0283	.0175	.0436	.0397	.0261	.0322	.0249	.0419	.0580	.0208	.0351	.0386
27	.0003	.0183	.0283	.0094	.0462	.0470	.0203	.0404	.0243	.0405	.0568	.0209	.0387	.0353
28	- .0018	.0193	.0252	.0095	.0363	.0500	.0136	.0372	.0233	.0394	.0554	.0228	.0380	.0346
29	- .0034	.0209	.0199	.0113	.0285	.0442	.0080	.0266	.0200	.0333	.0512	.0187	.0342	.0249
30	- .0045	.0178	.0059	.0065	.0256	.0410	- .0019	.0136	.0124	.0333	.0425	.0212	.0285	.0225
31	- .0044	.0173	.0060	.0066	.0256	.0410	- .0019	.0136	.0125	.0334	.0425	.0212	.0285	.0225
n_{equ}	7.1067	6.5775	6.8519	7.3651	6.3980	6.9683	6.7477	6.4372	5.8351	5.8125	5.8125	5.8286	5.8222	5.9104

Fig. 2 - Temporal autocorrelation coefficients of daily extreme temperature for 6 French stations.
Pressure network optimization.

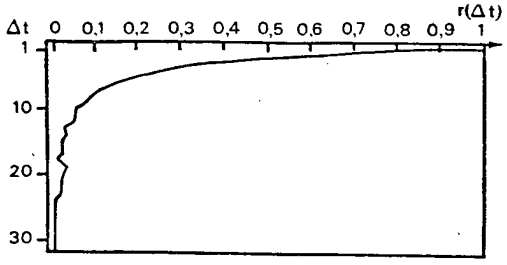


Fig. 1 - Temporal autocorrelation function of daily minimum temperature. Station Bordeaux - 1949-1976.

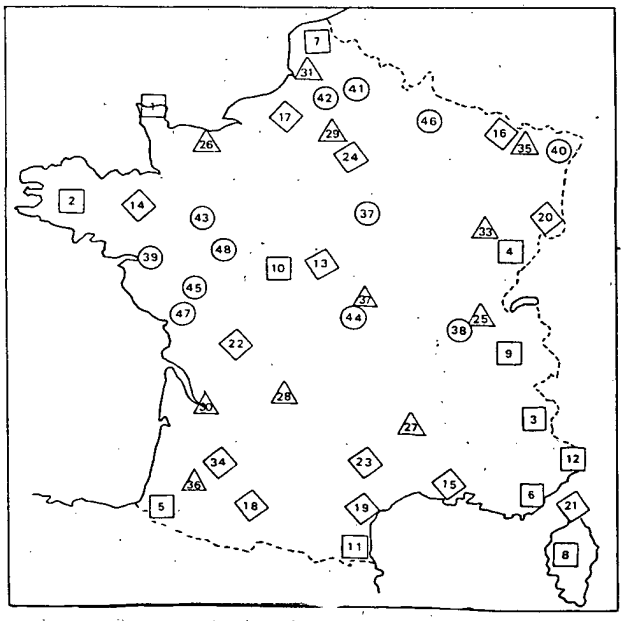


Fig. 4 - Pressure network optimisation. Geographical representation of preferential ranking of 48 stations.

□ de 1 à 12 ; ◇ de 13 à 24 ; △ de 25 à 36 ; ○ de 37 à 48.

	0 h	6 h	12 h	18 h	Σ	Rang global	Général
LE BOURGET	27	26	39	16	108	25	29
REIMS	39	30	29	45	143	37	46
BORDEAUX	23	28	38	35	124	30	30
BEAUVAIS	40	46	34	43	163	43	42
CHERBOURG	1	1	1	1	4	1	1
DIJON	33	34	32	34	133	34	33
CLERMONT-FERRAND	43	29	28	26	126	33	44
LIMOGES	20	31	25	21	97	21	22
LYON	25	25	33	42	125	31	38
AGEN	28	43	37	37	145	38	34
PERPIGNAN	10	9	11	10	40	9	11
RENNES	15	16	15	18	64	14	14
STRASBOURG	18	40	36	31	125	31	40
TOULOUSE	17	19	20	25	31	18	18
TOURS	48	36	46	28	158	42	48
AJACCIO	8	8	2	8	26	7	8
NANTES	41	41	43	38	163	43	39
MARIGNANE	16	17	19	15	66	15	19
METZ	12	20	21	13	66	15	16
MONTELIMAR	26	23	27	23	99	22	27
ABBEVILLE	29	37	40	33	139	36	31
NANCY	46	7	45	47	145	38	35
ROUEN	47	11	17	7	82	19	17
ST QUENTIN	42	44	41	41	168	46	41
TOULON	6	6	6	6	24	6	6
AUXERRE	21	45	44	39	149	40	37
NEVERS	14	14	14	14	56	13	13
LILLE	7	15	7	19	48	11	7
MILLAU	19	18	18	17	72	17	23
BESANCON	4	4	4	4	16	3	4
POITIERS	38	47	35	48	168	46	45
BOURGES	11	10	10	11	42	10	10
CAEN	35	27	31	29	122	29	26
COGNAC	45	39	47	44	175	48	47
VICHY	32	42	42	40	156	41	32
MONTPELLIER	30	32	13	27	102	24	19
LE MANS	24	48	48	46	166	45	43
MT DE MARSAN	36	35	26	32	119	28	36
NICE	13	13	12	12	50	12	12
GOURDON	34	24	30	30	118	27	28
ORLY	44	33	23	36	136	35	24
AMBERIEU	31	38	24	20	113	26	25
BASTIA	22	21	22	22	87	20	21
CHALES-LES-EAUX	9	12	9	9	39	8	9
EMBRUN	3	3	8	3	17	4	3
MULHOUSE	37	22	16	24	99	22	20
ROSTRENEN	2	2	3	2	9	2	2
DAX	5	5	5	5	20	5	5

Fig. 3 - Preferential ranking of 48 French stations.

The prove is based on specific inversion formulae of a matrix partitioned into 9 blocks which generalize the well-known Frobenius formula.

$$r^{i_1 i_1} = \max_i r^{ii}$$

Then, the iterative procedure of preferential ranking of existing set of stations may be formulated as follows : first, we compute R_{XX} , then the inverse matrix R_{XX}^{-1} . After that, we search the maximum value of the diagonal elements $\max r^{ii}$. Let the index associated whith this maximum be i_1, so the station i_1 will be the most redundant one, i.e. the best explicated by the others. So, we eliminate in the matrix R the row and the column number i_1 ; thus, we get a new matrix R (1) of M-1 order. Now, we perform the inversion of the matrix R (1), which gives $R_{(2)}^{-1} = (r^{ij}_{(2)})$. We search again the maximum of the diagonal elements $\max r^{ii}_{(1)}$ and we denote i the index corresponding to the maximum. The elimination of the row and the column number i_2 leads us to a new matrix of order N 2, that will be designed by $R(2)$. After that, we repeat successively the described procedure and we obtain a set of ranks

$$i_1 \alpha i_2 \alpha i_3 \alpha \ldots \alpha i_n$$

which define the preferential ranking of the stations in the increasing order of importance.

When the number of stations is too high, or when the correlation matrix is bad conditioned, the inversion procedure is replaced by a pseudo-inversion, for example by means of the formula

$$\left[R_{xx} + \varepsilon I_n\right]^{-1}$$

8. NETWORK OPTIMIZATION BY MEANS OF CANONICAL CORRELATION.

The canonical correlation coefficient between two sets of variables $x_1, \ldots x_n$ and $y_1, \ldots y_m$ may be used for our optimization problem. For that, let's introduce the notation

$X = (x_1, \ldots x_i, \ldots x_N)$ $\quad X(i_1, \ldots i_K) = (x_{i_1}, \ldots x_{i_K})$ $\quad Z(i_1, \ldots i_K) = (x_1, \ldots x_{i_1} \ldots x_N) - X(i_1, \ldots i_K)$

We apply a stepwise procedure, which at each iteration, consists to form two groups of stations X ($i_1, \ldots, i_K$) the group of K stations to be eliminated, as the best explicated by the others, and the group of N-K remaining stations $Z (i_1, \ldots, i_k) = \{ x_{j_1}, \ldots, x_{j_S}, \ldots, x_{j_{N-K}} \}$ where $j_S \neq i_t$.

Iteration 1. We perform the operations :

1. $\rho_1[i] = \rho [X(i)/Z(i)] \quad i = 1,N$

2. $\max_i \rho_1[i] = \rho_1[i_1]$

The station number i_1 is that one, which is the best explicated by the N-1 others, so it's the less interesting station for the whole set and the first one to be eliminated, as the most redundant.

Iteration "K + 1"

After the iteration "K", we have got two groups of stations, the K stations to be eliminated $X[i_1, \ldots, i_K]$ and the N-K remaining stations $Z [i_1, \ldots, i_K]$. Thus, we performs the following operations :

3. $\rho_{K+1}[i] = \rho\left[X(i_1, \ldots i_K, i)/Z(i_1, \ldots i_K i)\right]$

4. $\max_i \rho_{K+1}[i] = \rho_{K+1}[i_{K+1}]$

Carrying the iterations from 1 to N-1, we obtain a preferential ranking of the set of stations in the increasing order ; the first station is the less interesting and the last one the most interesting.

The basic information is given by the correlation matrix R_{XX} (N×N) for the whole system os stations. The iterative algorithm implies the use of canonical correlation subroutine, which computes the canonical correlations coefficient $\rho [E/R]$ between the group E of eliminated stations and the group R of remaining stations. Such computation is based on the elaboration of four blocks R_{EE}, R_{ER}, R_{RR}, R_{RE}, from the initial global matrix R_{XX} and the computation of the value :

$$\rho^2\left[E/R\right] = \lambda_1\left[R_{EE}^{-1} R_{ER} R_{RR}^{-1} R_{RE}\right]$$

To attenuate some insufficiency of the canonical correlation, which is rather an index for testing the independancy between two groups of variables than an index for measuring the degree of dependency, we have got as informativity criterion the sum of some canonical correlation coefficients and not only the first of them .

9. NETWORK OPTIMIZATION BY MINIMISATION OF THE OPTIMAL INTERPOLATION ERROR.

One of the important aims of networks implantation for the knowledge of a meteorological field is the necessity to perform an objective analysis of this field, i.e. to interpolate its values at the points of a regular grid by means of the values observed at the stations.

This step is the first state of the resolution in finite differences of the system of differential equations (Navier-Stokes equations) characterizing the evolution of atmosphere in the frame of hydrodynamical model.

There are many variants for the realization of such an objective analysis, but the classical variant of Gandin presents many advantages and in particular gives an explicit formula for the interpolation error in each point of the field.

Then it is possible to draw the curves of isovalues of optimal interpolation error. It seems logical to use the extreme values of this field for the rational localization of the stations. We realize the optimal localization of new stations putting them at the points corresponding to the maximum of errors, as this has been realized by Machkovic (5) for example.

10. THE PRINCIPLE OF NETWORK OPTIMIZATION BY MEANS OF THE METHOD OF UNIVERSAL KRIGING.

The basic ideas of the meteorologists for the objective analysis of a field have been, at the origin of the development of a very interesting method of fields analysis, which is the so-called the method of Kriging. The Fontainebleau school under the direction of the Professor Matheron has considerably generalised, systematised and developed the methods of optimal interpolation. Without considering the technical details of this theory, we can say that the methodology of Kriging solve the problem of objective analysis in less restrictive conditions, which means in a more realistic frame, nearer to real problems (12, 13).

Consequently, this apparatus can be applicated in similar way as optimal interpolation to the problem of network optimization. The only difference will be in the quality of the estimation of Kriging operator passing through the definition of structure functions (variograms) or generalised covariance functions of the field.

11. OPTIMIZATION OF THE NETWORK MESH.

We are going to consider this problem on an example of application concerning the choice of the optimum intervall (spatial or temporal) between the observations.
Let us suppose that the autocorrelation function has the following form :

$$\Psi(d) = e^{-\alpha d},$$

where $\alpha = 1/d_c$
and d_c is the correlation radius.
Let us suppose that it concerns a spatial correlation studied on a linear domain of length, with unit step equals to 1.

We have then $N = \frac{L}{l} + 1 \; ; \; r = e^{-\frac{l}{d_o}}$

The formula of the equivalent number becomes :

$$n_{equ} = \frac{N^2 [1-r^2]^2}{N[1-r^4] - 2r^2[1-r^{2N}]}$$

and when $l \to 0$, we obtain

$$\lim_{l \to 0} n_{equ} = \frac{2N_o^2}{2N_o - (1 - e^{-2N_o})}$$

where

$$N_o = \frac{L}{d_o}.$$

This limit value corresponds to a finite number, which indicates properly that the amount of the information furnished by the network does not increase indefinitely with the multiplication of the measuring points. Consequently, it is possible to foresee a rational choice of the number of observational stations.

This statistical approach allows to come to the problem of network mesh optimization. For this purpose, we study the behaviour of the equivalent number of stations corresponding to the network mesh. The mesh l, the dimension L of the domain and the number of stations N are related as follows :

$$l = \frac{L}{N-1}$$

We study the behaviour of expressing the network informativity when N increased, that is when l decreases. The shape of the curve is characteristical for the correlation functions that we have studied. With the reduction of the network, its informativity expressed by $n_{equ}(N)$ first increases, and then becomes stationary ; the further reduction of the mesh brings no more profit in informativity.

The optimum mesh l_{opt} may be defined by a reasonable value for which the maximum value of informativity has already been reached. We can precise that, by admitting $l_{opt}(1-\varepsilon)$ the value of the mesh for which we have

$$\frac{n_{equ}[N]}{\max_l n_{equ}[N]} = 1 - \varepsilon$$

Through analogy with the classical thresholds of mathematical statistics, we could use the mesh optimal at the level $l_{opt}(0,95)$; $l_{opt}(0,99)$:

Such a study has been ralized for different autocorrelation functions of the field of which here are four examples :

On the fig. 5, the correlation function is exponential :

$$r(d) = e^{-d/d_o}$$

On the graphic it has been pointed the optimal values of the mesh as being $l_{opt}(0,95)$ and $l_{opt}(0,99)$. The corresponding values of the equivalent number are noted $n_{equ}(0,95)$ or $n_{equ}(0,99)$.

If the cost of the network maintenance was known in function of the number of stations, the notion of the optimal mesh could be precised in economical terms. This optimal mesh would have been defined through a reasonable compromise between two contradictory tendancies : the increase of the stations number has got a positive effect (increase of informativity), but also a negative effect (increase of the cost of network maintenance).

Here it is interesting to underline that if the cost of the network maintenance is a function of the stations number $C = C(N)$, on the contrary the losses P due to the network unsufficiency are a decreasing function not of N, but of $n = n_{equ}(N)$, the equivalent number of independent stations : $P = P(n_{equ})$

On the fig. 6, 7, we have taken the autocorrelation function of the sea-level pressure field. the case of validity of the hypothesis of homogeneity, spatial isotropy and temporal stationarity of this pressure field its spatio-temporal autocorrelation function has the form :

$$r[d,\tau] = [1+\Delta] \cdot e^{-\Delta}$$

where

$$\Delta^2 = \left(\frac{d}{1005}\right)^2 + \left(\frac{\tau}{30}\right)^2$$

d in km and τ in hours.

This formula expresses the correlation between two observations of the pressure field realized in two stations distant one from the other of d km at two moments, different of T-hours.

We have separately studied the spatial correlation (fig. 6) putting $\tau = 0$:

$$r[d] = \left[1 + \frac{d}{1005}\right] \cdot e^{-\frac{d}{1005}}$$

For this purpose, we have varied the dimenstions of the spatial domain D, on which we can consider that are positioned a certain amount of stations regulary separated from each other. In this example D = 9000 km.

On fig. 7 we have separately studied the temporal correlation by admitting d = 0 :

$$r[\tau] = \left[1 + \frac{\tau}{30}\right] \cdot e^{-\frac{\tau}{30}}$$

and thus we have varied the dimensions of temporal domain T, during which a certain amount of observations are realized at regular intervals. In this example T = 700 hours.

On the fig. 8 we have studied a well-known spatial autocorrelation function

$$r[d] = \frac{\sin 0{,}0015/d}{0{,}0015/d} \cdot e^{-d/4\,000}$$

where d is in km.

On the fig. 9 we show the values of the equivalent number for the radiance field measured by satellite on a domain of the South West part of the South Atlantic.

The curve marked (-) corresponds to the horizontal direction and the curve noted (.) to the vertical direction.

The difference between the correlational structure in these two directions appears in the distinct behaviour of the equivalent number reaching more quickly the state of saturation for the horizontal direction.

That's why the optimal grid should have been rectangular, the choice of the optimal mesh $l_{opt}(1-\varepsilon)$ thus being defined starting from the same principle : for example we can take the value $\varepsilon = 0{,}05$ for both directions. In this example the signal noise ratio was $\frac{\sigma_e}{\sigma_x} = 0{,}1$.

A similar study of course may be applicated for all the autocorrelation functions known in the technical litterature.

For all these curves the shape is characteristical : a fast increase, followed by a stagnatinon corresponding to the predominancy of the redundancy on the informational contribution.

Another interesting approach was proposed by Sneyers (9) in the simple case of linear interpolation, for the spatial domain corresponding to the linear part of the autocorrelation function of the field. The minimization of the variance of global error, composed of both measurement and interpolation errors leads to define three variants of network density : insufficient, superabundant and optimal.

Thus the relations $r = \frac{\sigma_x^2 - \sigma_e^2}{\sigma_x^2}$
define the correlation between two points separated from each other by a distance equal to the optimal mesh. Here σ_e^2 denotes the error variance and σ_x^2 of the meteorological element.

12. DEFINITION OF THE OPTIMAL MEASUREMENT UNIT OF A METEOROLOGICAL VARIABLE.

Here we shortly represent some ideas formulated in signal theory and being able to be applied to meteorological measurement.

Let's consider the problem of optimal dicretisation of a continuous signal g (τ). Thus the spectral density of this signal is defined by the formula

$$Q(\omega) = \frac{1}{2\pi} \int e^{i\omega\tau} g(\tau)\, d\tau$$

and then we have the inversion formula

$$g(\tau) = \int e^{i\omega\tau} Q(\omega)\, d\omega$$

If $Q(\omega) = 0$ for $\omega > A$
the function g (τ) is then called a function with limited specter. For such a class of functions, a fundamental result has been obtained by Shanon, Kotelnikov, Ignatiev ( ).

Theorema

If $g(\tau)$ is a function with limited specter, then it could be exactly reconstituted thanks to a discretisation with a path equal to

$$\Delta\tau = \frac{1}{2A}$$

This result may be considered as a base to optimal discretisation.
For example : if we wish that the choosen path permits the exact knowledge of the structure of the studied meteorological variable, in the sense of the autocorrelation function , then we can use the relations

$$R(\tau) = \int e^{i\omega\tau} Q(\omega)\, d\omega \qquad Q(\omega) = \frac{1}{2\pi} \int e^{-i\omega\tau} R(\tau)\, d\tau .$$

If the function Q (w) don't vanish , as it is often the case in practice when $T \to 0$, we can introduce the notion of quasi-limited specter putting $Q(\omega) < \varepsilon$ if $\omega > A(\varepsilon)$
and defines a quasi-optimal discretisation of ε order associated with the path

$$\Delta\tau = \frac{1}{2A(\varepsilon)} .$$

Another approach consists to take as a criterion the quality of the estimation of a statistical moment, for example that of the mean value. A simple result of this sort has been obtained by Vilenkin (10) to define the optimal internal in the sense of the minimization of the variance of the arithmetical mean of the data. This optimal interval Δt must verify the following equation :

$$8R\left(\frac{\Delta t}{2}\right) - 5R(\Delta t) - 3 = 0 .$$

For the exponential autocorrelation functions of the type $R(\Delta t) = e^{-\alpha|\Delta t|}$, we thus obtain

$$\Delta t_{opt} = \frac{1.022}{\alpha} .$$

CONCLUSION.

The rationalisation of a climatological observations network is a rather complicated problem, the resolution of which is related to the choice of an adequate optimisation criterion defined by the purpose of the elaboration of such network. The general methodology is based on a very simple idea to use the knowledge of the statistical structure of concerned meteorological fields to improve the network informativity.

The solutions that we proposed are not unique and may be different in function of the adopted criterion or strategy. Anyway they must consider only as reasonable recommendations and not as imperative directives.

ACKNOWLEDGMENT.

The author thanks all his collaborators, students and military staff who participated to the elaboration of the present work : A. Yessayan, J. Boulanger, O. Dubrule, G. Le Goff, B. Renaude and H. Torgomian.

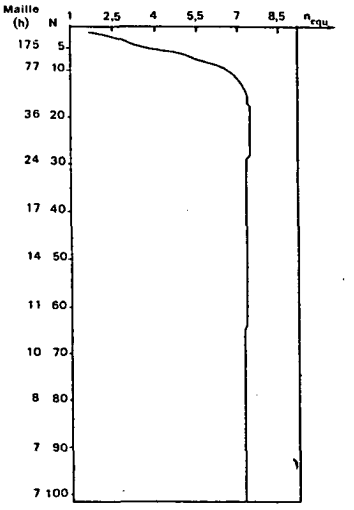


Fig. 7 - Optimization of temporal intervall

$$r[r] = \left[1 + \frac{r}{30}\right] \cdot e^{-\frac{r}{30}}$$

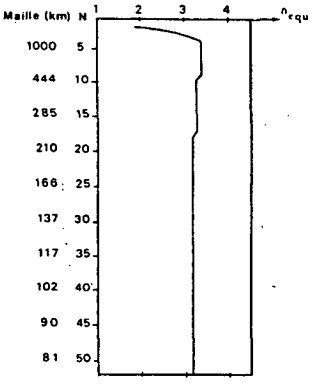


Fig. 8 - Stations of the network's optimal mesh.

$$r[d] = \frac{\sin 0,0015/d}{0,0015/d} \cdot e^{-d/4\,000}$$

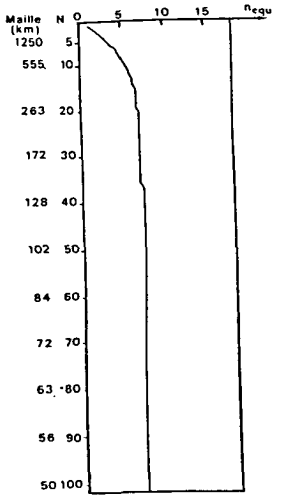


Fig. 5 - Network mesh optimization.

$$r[d] = e^{-d/d_o} \; ; \; d_o = 500\ km \; ; \; n_{equ}[0,95] = 10,06 \; ; \; n_{equ}[0,99] = 10,48 \qquad (d \text{ en km})$$

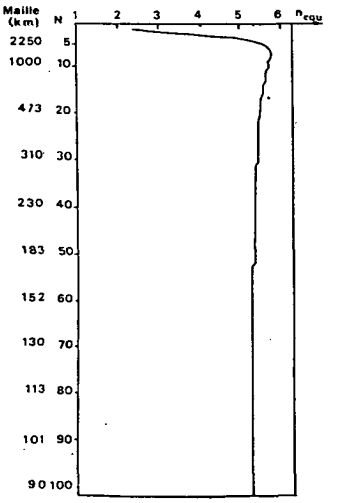


Fig. 6 - Network mesh optimization.

$$r[d] = \left[1 + \frac{d}{1005}\right] \cdot e^{-\frac{d}{1005}}$$

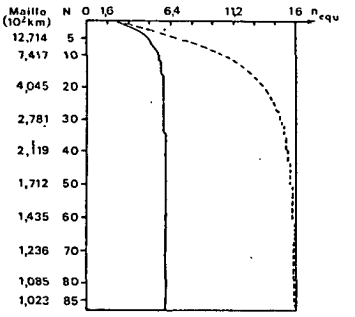


Fig. 9 - Equivalent number for the radiances'

Curve 1 - Horizontal direction

Curve 2 - Vertical direction

$$\alpha = \frac{\sigma_e}{\sigma_x} = 0,1$$

FITTING GENERALIZED GAMMA DISTRIBUTION TO WIND SPEED DATA

F. Reimão Ferrão and Luís Pessanha
Instituto Nacional de Meteorologia e Geofísica
Lisboa - Portugal

1. INTRODUCTION

The probability models more frequently proposed lately to be fitted to wind speed data are Weibull, gamma, Rayleigh and hyper gamma distributions (e.g., Hennessey, 1977; Gustus, 1977; Stewart and Essenwanger, 1978; Van der Auwera et al., 1980; Suzuki, 1980). The acceptable fitting of these different models frequency distributions of wind speeds may be explained by the large variability of the characteristics of this climatic element caused by differences in periods and locations. Taking into account a more general description of the statistical features of the underlying event, the purpose of this study is to suggest the use, as a first fitting approach, of the generalized gamma distribution (GGD), of which the abovementioned distributions are particular cases.

In order to reduce the number of parameters we will choose later, among the simpler models, the one which fits the sample best.

2. GENERALIZED GAMMA DISTRIBUTION (GGD)

The GGD is defined by the probability density function

$$f(x|\theta) = \theta_3[\theta_2\Gamma(\theta_4)]^{-1}.[\theta_2^{-1}(x-\theta_1)]^{\theta_3\theta_4-1} \cdot \exp\{-[\theta_2^{-1}(x-\theta_1)]^{\theta_3}\}$$

Where Γ is the gamma function and the vector θ has the components $(\theta_1, \theta_2, \theta_3, \theta_4)$, $-\infty<\theta_1<+\infty$, $\theta_2, \theta_3, \theta_4>0$ and $f(x|\theta) = 0$ if $x< \theta_1$.

The cumulative distribution function is

$$F(x|\theta)=\Gamma(\theta_4)^{-1}.\Gamma_t(\theta_4)$$

where $t = [\theta_2^{-1}(x-\theta_1)]^{\theta_3}$ and $\Gamma_t(\theta_4)$ is the incomplete gamma function.

In order to simplify the following equations we difine the function

$$\phi(p,q) = y^p(\ell n\, y)^q \text{ ,with } y = \frac{x-\theta_1}{\theta_2},$$

Whose asymptotic expectation was derived, taking the form

$$\mu(p,q) = [\theta_3^q\Gamma(\theta_4)]^{-1}.\ \Gamma^{(q)}(\theta_4+p\theta_3^{-1})$$

where $\Gamma^{(q)}$ is the qth derivative of Γ. Using this function the general expressions of the rth moment about the origin and the rth central moment, obtained directly from the definitions, are respectively:

$$\begin{cases} \mu'_r = \sum_{j=0}^{r}\binom{r}{j}\ \theta_1^{r-j}\ \theta_2^j\ \mu(j,0) \\ \mu_r = \theta_2^r\sum_{j=0}^{r}\binom{r}{j}[-\mu(1,0)]^{r-j}.\mu(j,0) \end{cases}$$

Mean, variance and coefficients of skewness and kurtosis of GGD can be calculated from these equations and estimates of the parameters can be obtained using the method of moments (Appendix 1). In this study the GGD parameters were estimated by using the maximum likelihood technique.

The GGD includes various probability models but only the special cases of the gamma ($\theta_3=1$), Weibull ($\theta_4=1$) and Rayleigh ($\sqrt{2}\theta_2$, $\theta_3=2, \theta_4=1$) are considered in this texte.

3. TESTS APPLIED

Two commonly used tests were selected for evaluating the goodness of fit of the

TABLE 1. VALUES OF THE PARAMETERS AND OF THE TESTS OF FIT

	4 PARAM. GGD	3 PARAM. GGD	GAMMA	WEIBULL	RAYLEIGH
			JANUARY		
θ_1	0.877	0.877	0.877	0.877	0.877
θ_2	6.243	6.547	8.287	19.135	19.852
θ_3	0.994	1.014	1	1.741	2
θ_4	2.699	2.635	2.155	1	1
χ^2	15.646	16.391	19.552	37.629	62.704
K-S	0.028	0.030	0.035	0.050	0.094
			17.927	759.99	
X			3 PARAM. GGD	3 PARAM. GGD	
			FEBRUARY		
θ_1	—0.465	—0.465	—0.465	—0.465	—0.465
θ_2	19.707	14.097	4.355	23.616	23.148
θ_3	1.880	1.881	1	2.224	2
θ_4	1.325	2.490	4.688	1	1
χ^2	23.742	339.54	43.651	27.568	28.449
K-S	0.028	0.140	0.071	0.048	0.038
X					0.296 WEIBULL
			JUNE		
θ_1	2.094	2.094	2.094	2.094	2.094
θ_2	9.547	7.484	9.751	14.944	15.472
θ_3	1.306	1.272	1	1.750	2
θ_4	1.646	2.197	1.576	1	1
χ^2	14.602	40.661	62.382	45.525	43.266
K-S	0.030	0.061	0.078	0.057	0.084
X			80.422 3 PARAM. GGD	418.04 3 PARAM. GGD	0.708 WEIBULL

distributions to the data: χ^2 and Kolmogorov--Smirnov (K-S) tests. The χ^2 statistics was calculated considering a number of class intervals approximately equal to

$$k = 1+3.3\ \ell n\ n,$$

with $\ell n\ n$ the natural logarithm of the sample size (Sturges, 1926). The K-S statistic applied is the maximum absolute deviation of the sample distribution $F(x)$ from the continuous hypothetical distribution $F_0(x)$: $D_n = max|F(x)-F_0(x)|$. It is important to note that these statistics were used essencialy to choose the model that fits the data best.

Concerning the reduction of the number of parameters, the method used consists in verifying whether the more general distribution with k parameters $\hat{\theta}$ and covariance matrix (v_{ij}) is identical to another simpler distribution with parameters θ_0, using the statistic

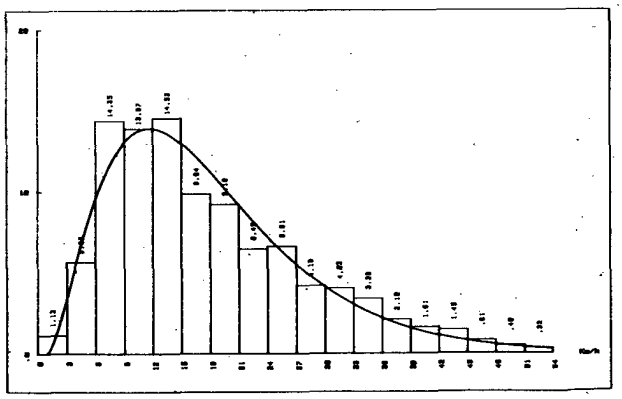
$$X = (\hat{\theta}-\theta_0)^T.(v_{ij})^{-1}.(\hat{\theta}-\theta_0)$$

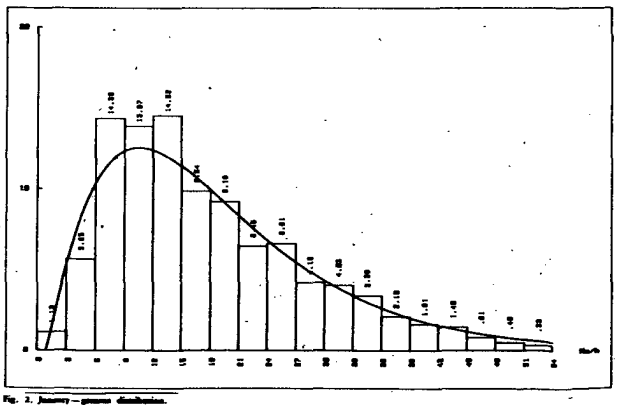
whose asymptotic distibution is approximately a χ^2 with k degrees of freedom (Sneyers, 1975).

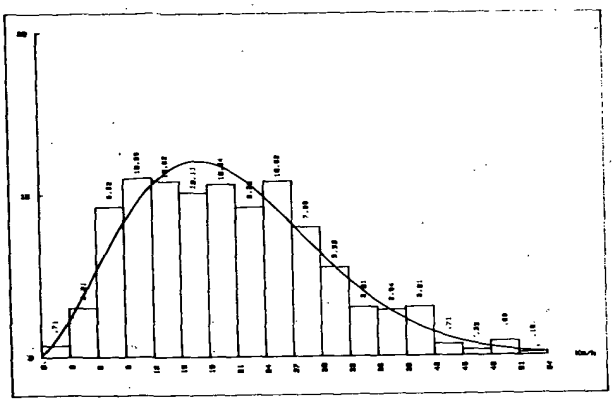
4. ANALYSIS OF ACTUAL EXAMPLES

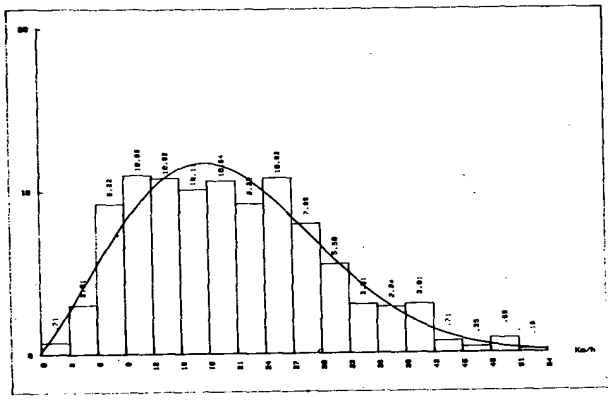
The three sets of wind speed data used in this study are 6 hourly averaged wind speeds at Sines in January, February and June, from 1975 to 1979. The parameter estimates and the values of statistics of the tests applied are presented in Table 1.

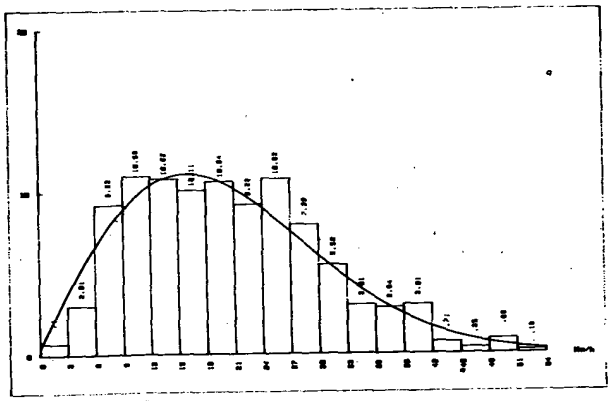
We began to estimate the parameters for 4 param. GGD for each sample solving the set of four maximum likelihood equations. The covariance matrix is not positive definite in the three cases, but as the χ^2 and K-S tests show a good fitting to the data, we have kept the value of the location parameter θ_1 in estimating the parameters of the other distributions.











It has been noted that, in January, the 4 and 3 param. GGD show pratically identical fittings considering the values of the χ^2 and K-S statistics and that the gamma and 3 param. GGD parameter estimates do not differ significantly according to the value of X statistic. The data are well fitted within the gamma branch of the GGD distribution, Figures 1 and 2.

In February both Weibull and Rayleigh models fit well the wind speed sample and are identical according to the value of X statistic, Figures 3,4 and 5.

The choice of the GGD branch that fits better to the sample is more difficult in the case of June data. The gamma distribution is nearer to the 3 param. GGD than to Weibull distribution but, on the other hand, it shows a worse fitting to the data than the latter. The 3 param. GGD seems to be the best alternative, Figure 6 to 9. It should be noted that Rayleigh distribution, although its fitting to these data is not good, shows a non significant value of X statistic comparatively to Weibull distribution, as in the case of February.

Finaly it should be stressed that, in the three examples the 4 param. GGD is the one which shows the best fitting to the data.

5. CONCLUDING REMARKS

The results described in this study suggest that the GGD is supple enough to enable the best fitting for wind speed frequencies in appreciable regions of variability of these frequencies in what concerns their skewness and kurtosis features.

Furthermore, it has been noted that the values of the statistics of the χ^2 and K-S tests are not significant in any of the examples at the 1% level and that, at the 5% level, only the case of February is significant. However, its value is fairly near to the critical value of 22.362.

In some cases, such as those presented in this text, the method of the maximum likelihood enables a good estimation of the GGD parameters, but, on the other hand, the matrix of the asymptotic variances and covariances of the parameters usually is not positive definite. Nevertheless the authors have noted that, in the cases where the maximum likelihood estimates are not good, the method of the moments enables on acceptable parameters estimation.

Thus, because many problems still arise, we think, that this distribution, perhaps with a better parameterization, requises special attention in future.

6. APPENDIX

1. METHOD OF MOMENTS

The values of the parameter estimates $\hat{\theta}_3$ and $\hat{\theta}_4$ are given by the solutions of the equations

$$\begin{cases} m_3 . m_2^{-1.5} = A^{-1.5} [2\mu^3(1,0) - 3\mu(1,0)\mu(2,0) + \mu(3,0)] \\ m_4 . m_2^{-2} = A^{-2} [\mu(4,0) - 4\mu(1,0)\mu(3,0) + 6\mu^2(1,0) . \\ \qquad . \mu(2,0) - 3\mu^4(1,0)] \end{cases}$$

where $A = \mu(2,0) - \mu^2(1,0)$, the funtion μ is defined in the text and

$$m_\ell = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^\ell \quad , \quad \ell = 1, 2, \ldots$$

are the sample central moments.

Once the values of $\hat{\theta}_3$ and $\hat{\theta}_4$ are calculated, then $\hat{\theta}_1$ and $\hat{\theta}_2$ are determined as

$$\begin{cases} \hat{\theta}_2 = s A^{-0.5} \\ \hat{\theta}_1 = \bar{x} - \hat{\theta}_2 \mu(1,0) \end{cases}$$

where $\bar{x}$ and s are, respectively, the sample mean and standard desviation.

2. METHOD OF MAXIMUM LIKELIHOOD

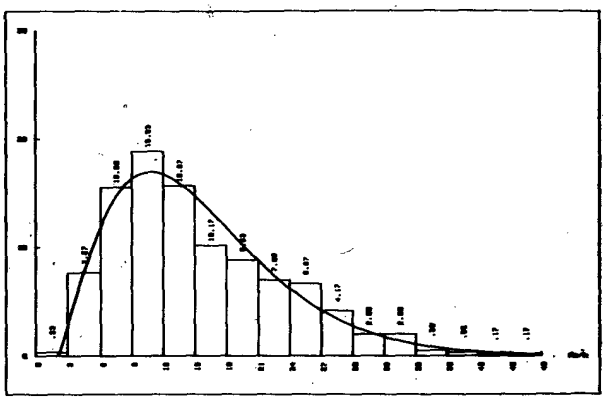
Being $x_n = (x_1, \ldots, x_n)$ a random sample of n observations distributed according to GGD, the natural logarithm of joint density of x_n is given by the equation

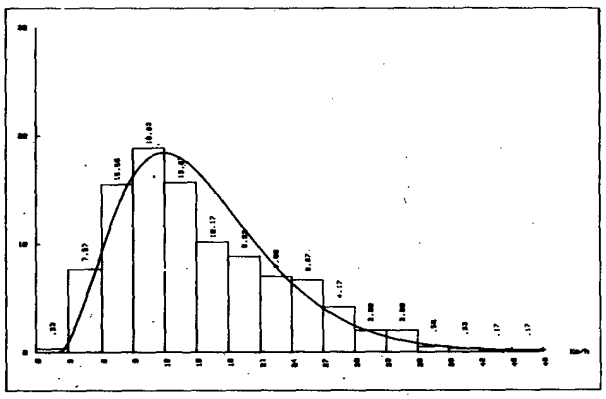
$$L(\theta) = \ell n \prod_{i=1}^{n} f(x_i | \theta) =$$

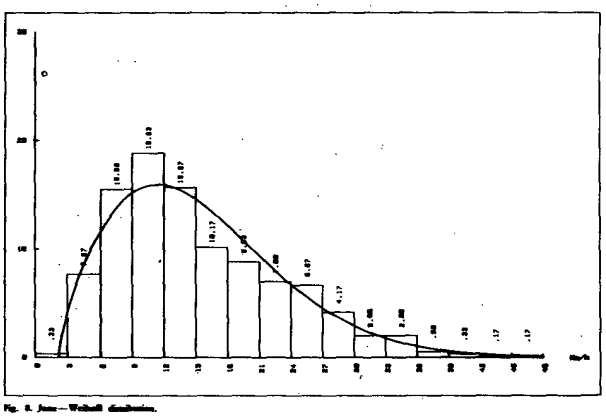
$$= n[\ell n \theta_3 - \ell n \theta_2 - \ell n \Gamma(\theta_4) + (\theta_3 \theta_4 - 1) m(0,1) - m(\theta_3, \theta)]$$

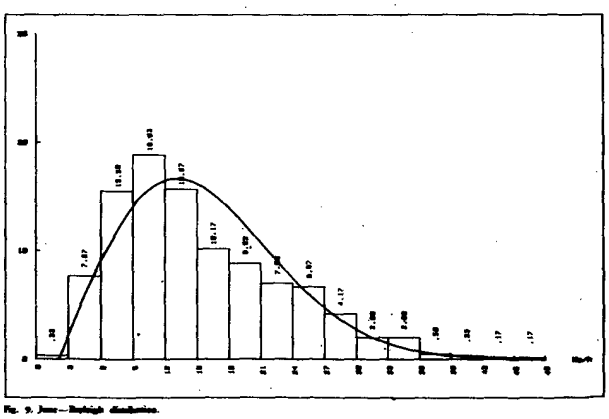
The maximum likelihood estimates of the parameters are given by the solutions of the set of equations

$$\frac{\partial L(\theta)}{\partial \theta} = 0, \ i.e.:$$









$$\begin{cases} (1-\theta_3\theta_4).m(-1,0)+\theta_3.m(\theta_3-1,0) = 0 \\ m(\theta_3,0)-\theta_4 = 0 \\ \theta_3^{-1}+\theta_4.m(0,1)-m(\theta_3,1) = 0 \\ \theta_3.m(0,1)-\psi(\theta_4) = 0 \end{cases}$$

where ψ is the digamma function and

$$m(p,q)=\frac{1}{n}\sum_{i=1}^{n} y_i^p(\ell n\ y_i)^q$$

with $y_i=\frac{x_i-\theta_1}{\theta_2}$ is the impirical value of the function $\mu(p,q)$.

3. COVARIANCE MATRIX

This matrix is given by the expression

$$(cov\ (\theta_i,\theta_j))=\left[-E\left(\frac{\partial^2 L(\theta)}{\partial\theta^2}\right)\right]^{-1}=\frac{1}{n}\ I^{-1}$$

where I is a matrix whose elements are the following:

$$i_{11}=n\theta_2^{-2}\left[\theta_3(\theta_3\theta_4-2)+1\right]\mu(-2,0)$$

$$i_{12}=n\theta_2^{-2}\theta_3(\theta_3\theta_4-1)\mu(-1,0)$$

$$i_{13}=-n\theta_2^{-1}\mu(-1,0)\{(\theta_4-\theta_3^{-1})[\psi(\theta_4-\theta_3^{-1})+1]-\theta_4+1\}$$

$$i_{14}=n\theta_2^{-1}\theta_3\mu(-1,0)$$

$$i_{22}=n\theta_2^{-2}\theta_3^2\theta_4$$

$$i_{23}=-n\theta_2^{-1}[\theta_4\psi(\theta_4)+1]$$

$$i_{24}=n\theta_2^{-1}\theta_3$$

$$i_{33}=n\theta_3^{-2}\{\theta_4[\psi^2(\theta_4)+\psi'(\theta_4)]+2\psi(\theta_4)+1\}$$

$$i_{34}=-n\theta_3^{-1}\psi(\theta_4)$$

$$i_{44}=n\psi'(\theta_4)$$

where ψ' is the trigamma function.

7. REFERENCES

Sturges, H.A., 1926: the choice of a class interval. Journal of American Statistical Association, 21, 65-66.

Sneyers, R., 1975: Sur l'analyse statistique des séries d'observation. Note Technique Nº 143, Org. Met. Mondiale.

Hennessey. J.P., 1977: Some aspects of wind power statistics. J. Appl. Meteor., 16, 119-128.

Stewart, D.H. and Essenwanger, O.M., 1978: Frequency distribution of wind speed near the surface. J. Appl. Meteor., 17, 1633-1642.

Van der Auwera, L., F. de Meyer and L.M. Malet, 1980: The use of the Weibull three-parameter model for estimating mean wind power densities. J.Appl. Meteor., 19, 819-825.

Suzuki, E., 1980: A summarized review of theoretical distributions fitted to climatic factors and Marcov chain models of weather sequences, with some examples. Statistical Climatology, Elsevier Sc. Pub. Comp.

THE GENERALIZED CIRCULAR DISTRIBUTION APPLIED TO WIND DIRECTIONS AS MEANS FOR URBAN CLIMATOLOGY

Konrad Cehak

Institute for Meteorology and Geophysics of the University of Vienna
and
Central Office for Meteorology and Geodynamics in Vienna
A 1190 Vienna, Austria

Von Mises (1918) has defined a circular distribution, which he took as the counterpart of the normal distribution for directional data. This circular distribution has one mode within its definition range of $0 \leq x \leq 2\pi$. It is evident that such a distribution can be used for meteorological purposes, say for the representation of variables with an annual or diurnal variation, or for the representation of the direction of upper winds, which usually are unimodal at least above a certain altitude (Cehak, 1969). The theory of this distribution has been given by Gumbel et al. (1953).

For wider application in Climatology the circular distribution had to be generalized in order to be capable to describe the distribution of variables having more than one mode in the interval $0 \leq x \leq 2\pi$. This has been done by Sneyers and Van Isacker (1979), who introduced their new distribution on the occasion of the First International Meeting on Statistical Climatology in Tokyo. Sneyers (1980) has shown the application of this generalized circular distribution to wind data and, at the same time, combined it with his selective harmonic analysis (Sneyers, 1975).

In this paper the generalized circular distribution shall be applied for the comparison of wind data of two stations in order to investigate the possible change of climatic conditions in a growing city.

Usually, climatic changes are investigated by means of temperature and precipitation data. Such changes are often caused by the growth of the city, in which the station is situated, or, at least, the climatic change is overlain by the urban influence. Exemples for such investigations for the city of Vienna, Austria, are given by several papers in the climatological literature (e.g. Böhm und Gabl, 1978, Böhm, 1979). Vienna is not situated in a flat country, the question may arise whether the climatic features of the city and its environments are due to urban or orographical influences.

Especially in Vienna, where the growth of the city is not as distinct as in many of the other large towns (Böhm und Gabl, 1978), every possible detail of the climatic situation may be investigated in order to get the best possible insight in the urban climate and its changes. So for the first time, the wind conditions are taken up for comparison with a neighbour station in a small village, which has not taken part in the modern trend of growing and industrialization to an extent, which one would notice in the climatic data. This small willage is Gross Enzersdorf, which is situated east of Vienna in a distance of about 20 km from the city centre of Vienna or 25 km from the meteorological observatory Vienna - Hohe Warte. Furthermore, Gross Enzersdorf lies in the flat part of the surroundings of Vienna, so that orography does not play any role in the distribution of climatic elements there.

Gross Enzersdorf is only a secondary climatological station, therefore, only the distribution of wind direction could be chosen for the aims of this paper. The data for both stations are taken from the climatological annals and archives of the Austrian Meteorological Service (Zentralanstalt für Meteorologie und Geodynamik, 1936-80). In the Austrian annals wind directions are grouped in eight sectors of 45^{o} width. The group centered around the direction north has been chosen as zero-direction.

According to Sneyers and Van Isacker (1979) the generalized circular distribution is given by the probability density function

$$f(x) = C^{-1} \exp \sum_{j=1}^{N} (a_j \sin jx + b_j \cos jx) \qquad (1)$$

If the full circle is divided into M parts (here: M wind direction groups), x is a multiple of $2\pi/M$. The constant C is determined such that the condition

$$\int_0^{2\pi} f(x)dx = 1 \qquad (2)$$

is fullfilled. The quantities

$$A_j = \frac{1}{n} \sum_{i=1}^{n} \sin jx_i \qquad (3)$$

$$B_j = \frac{1}{n} \sum_{i=1}^{n} \cos jx_i \qquad (4)$$

in which the frequencies of the values x_i are taken from the data series, are exhaustive statistics and the vector Θ of the 2N parameters a_j and b_j (the number N of the indices j has to chosen suitably) can be determined using the maximum likelihood method by differentiating the likelihhod function

$$L = [C^{-1} \prod_{i=1}^{n} \exp \sum_{j=1}^{N} (a_j \sin jx_i + b_j \cos jx_i)] =$$

$$= [C^{-1} \exp \sum_{j=1}^{n} (a_j A_j + b_j B_j)]^n \qquad (5)$$

or its logarithm. This yields the equations

$$\frac{\partial \ln L}{\partial z_k} = 0 \qquad (6)$$

if z_k is one of the 2N components of Θ and the matrix

$$\Delta = (E \frac{\partial^2 \ln L}{\partial z_j \partial z_k}) \qquad (7)$$

leads to the covariance matrix for the estimates $\hat{\Theta}$ of Θ:

$$\text{cov } \hat{\Theta} = (-\Delta)^{-1} \qquad (8)$$

The necessary calculations start with the determination of A_j and B_j by means of a Fourier analysis of the wind directions. The analysis was made for consecutive periods of five years length for both stations starting with the pentade 1936 to 1940 as the first full period, in which both stations were observing. The years 1945 and 1946 had to be omitted on account of too many missing data in the station of Gross Enzersdorf. The distribution of the wind directions were expressed in permille numbers from the observations with wind blowing, calmes were excluded. These numbers then were used for the Fourier analysis, the results of which are given in table 1.

The wind roses of both stations show two distinct main directions (NW and SE or S in Gross Enzersdorf, W or NW and SE in Vienna), therefore, the Fourier analysis shows the domination of the second wave in the analysis, which explains about half of the variance. Further, a significant part of the variance is explained by the first harmonic component. These two components suffice for the representation of the frequency distribution, which led to the selection of N = 2 in the definition of the generalized circular distribution.

On the basis of the reults of the initial Fourier analysis the first two terms of the exponential in the definition (5) of the generalized circular distribution were used. The calculation of the entities A_j and B_j using (3) and (4) yields the values presented in table 2. These values were used for the solution of equation (6). As a direct solution of this equation is not possible, the maximum of the likelihood-function ln L was obtained by the means of a numerical method, the method of steepest descent. The solutions a_j and b_j are also contained in table 2. As in the Fourier analysis one can transform the sums of sine and cosine of the same arguments in equation (1) such that one gets

$$f(x) = C^{-1} \exp \sum_{j=1}^{N} c_j \cos (jx - \phi_j) \qquad (9)$$

In our case, N = 2, we can resolve the exponential of a sum into the product of two exponentials and see that our generalized circular distribution is a double wave, the amplitude of which itself is modulated by a simple wave. Having found the maximum likelihood solution for the constants a_j and b_j one has to calculate the values of $C.f(x_i)$ and then the constant C can be derived from a summation in analogy to equation (2). Equation (9) is independent from the number of direction sectors used in the maximum likelihood estimation, however, the value of C has to be recalculated for each directional partitioning.

An exemple for the fit of data by means of the generalized circular distribution is shown in table 3 and figure 1. The data are taken from the pentade 1976 to 1980 from the station of Vienna. The fit is acceptable for all practical purposes. The equation of the frequency density is:

$$f(x) = \frac{1}{8.9268} \exp 0.3479 \cos (x - 283) \cdot \exp 0.5439 \cos (2x - 222) \qquad (10)$$

The maxima of both waves are approximately at the same angle, therefore, the distribution has such a strong peak as it can be seen in figure 1. Figure 2 represents the cumulative frequency function, from which the distributions for arbitrarily selected groupings of the direction could be derived. In addition, with the generalized circular distribution frequencies for other groupings than that, from which the density function had been derived, can easily be calculated by changing the scale of x in equation (10), having in mind that x is a multiple of 2π:M, and recalculation of C. The frequency function in figure 2 is not yet a smooth curve. If one would use smaller steps for x (a larger value of the number of groups M), the curve would approach a smooth unbroken line.

The interpretation of the numbers given in table 2 yields the climatological results. The maxima of the distributions are near to the phase angle k_1, the value of this maximum is determined by both amplitudes. A comparison of the values of d_i and k_i (i = 1, 2) in the different pentades shows the existence of more or less distinct trends. A rough characterization of these trends is provided by linear regression lines of these parameters. Table 4 gives the coefficients of these regression lines

$$y = a\,x + b \qquad (11)$$

in which y is one of the parameters d_i or k_i and x is the time rank of the pentade counted from the medium pentade 1956 to 1960. The amplitudes in Vienna show a decreasing trend, the same holds true for the first component in Gross Enzersdorf, however, the second component increases. The phase angles in Vienna decrease, those of Gross Enzersdorf increase. All the a-coefficients are too small to be statistically significant, nevertheless they indicate that the prevailing wind

Table 4. Coefficients of the regression lines for phase angles and amplitudes with pentade number (equation (11)).

		a	b
Vienna	k_1	-0.7667	280
	k_2	-1.1667	238
	d_1	-0.0129	0.4366
	d_2	-0.0055	0.6307
Gross Enzersdorf	k_1	5.9833	275
	k_2	4.8833	275
	d_1	-0.0350	0.3693
	d_2	0.0027	0.7803

direction in Gross Enzersdorf turned from more south-westerly to more north-westerly directions whereas at the same time in Vienna a slight change of direction in the opposite sense was observed.

Following the course of the data through the entire series one finds that the first component in Vienna shows but a slight change from north-westerly to westerly winds for the maximum direction with only small deviations from the regression line. In Gross Enzersdorf, however, there is a steep ascent in the opposite direction and with large deviations from the regression line. In the second component the same principal difference in the main direction of change is to be seen, a slight more southerly component at the end of the observation period in Vienna, a strong northerly component in Gross Enzersdorf. The deviations from the regression line have at first the same sign in both stations, however, after a large deviation in the pentade 1961 to 1965 they stay on the northerly side in Gross Enzersdorf, while they cross to the southerly side in Vienna.

There is not much to say about the amplitudes of the first component, they fall slightly in both stations and the deviations from the regression line have opposite sign. The deviations are also on different sides of the regression line in the second component. This means that the wind in Vienna is more evenly distributed at the same time when in Gross Enzersdorf larger differences between the frequencies of the single directions can be observed. This effect is most dominent between 1945 and 1965.

The dominance of northerly components in the wind rose of Gross Enzersdorf, especially at the end of the observation period, makes on think that the wind tried to flow around the rough obstacle of the city the more the city extended and got more and more high buildings. A strong local wind system did not develop, however, a feeble one may be existent.

The length of the series is short with its 45 years. It is not possible to make statements on the long-range climatic changes and the corresponding changes in the features of the wind distribution. For this medium-range period the changes are by no means drastic, but they may be of interest for the urban climatologist, nevertheless.

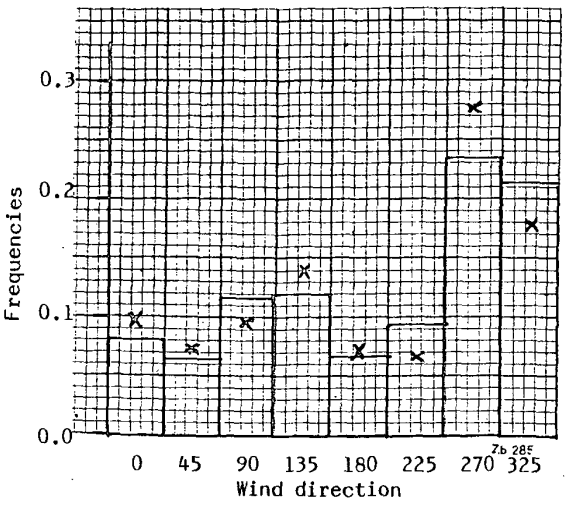


Fig. 1. Frequencies of wind direction in Vienna, 1976-1980. Full line: Generalized circular distribution. Crosses: Observed values.

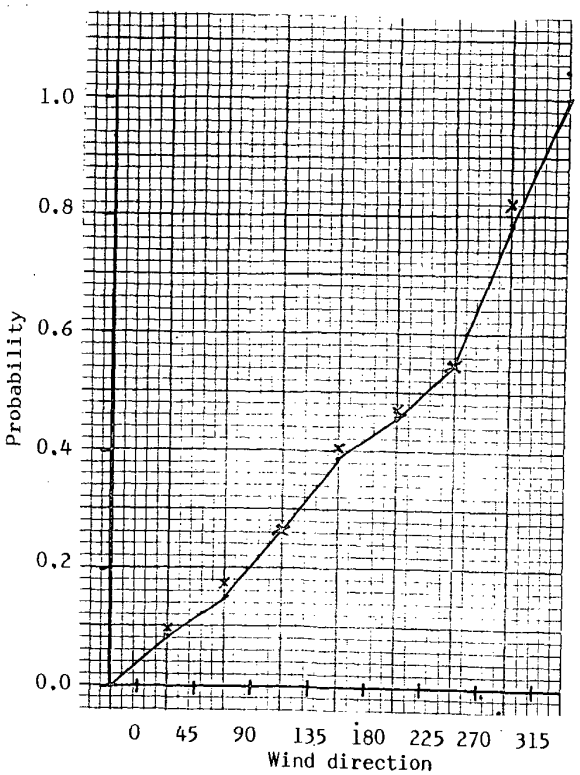


Fig. 2. Distribution function of wind direction in Vienna, 1976-1980. Full line: Generalized circular distribution. Crosses: Observed values.

Table 1. Results of the Fourier analysis of wind directions in Vienna and Gross Enzersdorf. The Fourier series is defined by:
$y = \bar{x} + c_1 \cos(x-k_1) + c_2 \cos(2x-k_2) + c_3 \cos(3x-k_3) + c_4 \cos 4x$.
R_i is the variance reduction by the i-th component, s is the standard deviation of the frequencies about their mean.

		c_1	k_1	R_1	c_2	k_2	R_2	c_3	k_3	R_3	c_4	R_4	s
Vienna	1936-40	67.5	292	28	91.8	240	52	53.2	95	18	12.8	2	89.8
	1941-44	71.8	293	35	85.8	237	50	44.1	106	13	10.8	2	85.7
	1947-50	74.4	286	37	83.2	231	46	51.3	101	17	1.5	0	86.9
	1951-55	71.8	286	38	78.9	235	46	45.6	90	15	5.0	0	82.2
	1956-60	64.3	292	40	67.9	247	44	39.1	85	15	9.0	2	72.2
	1961-65	69.2	295	45	68.3	261	43	34.1	98	11	8.0	1	73.3
	1966-70	71.2	284	34	82.7	230	47	52.0	98	18	6.5	1	85.7
	1971-75	48.1	281	19	86.7	228	61	49.5	95	20	6.0	1	78.6
	1976-80	53.2	286	31	67.7	221	50	39.9	92	17	10.0	2	67.8
Gross Enzersdorf	1936-40	61.0	282	34	81.4	280	61	23.6	144	5	4.3	0	74.0
	1941-44	61.4	278	33	82.2	242	59	30.2	133	8	4.8	0	75.8
	1947-50	59.4	292	16	114.9	262	59	9.2	69	0	52.8	25	105.8
	1951-55	55.9	290	16	115.3	252	67	19.0	91	2	38.8	15	99.4
	1956-60	55.0	290	21	104.7	255	74	27.2	93	5	2.5	0	85.8
	1961-65	47.1	320	14	117.4	290	85	12.3	324	1	7.0	1	90.1
	1966-70	60.8	300	27	98.6	296	72	7.5	12	0	1.5	0	82.1
	1971-75	47.9	289	27	70.5	312	58	22.6	58	6	19.5	9	64.1
	1976-80	33.2	314	13	73.9	292	66	33.9	37	14	15.8	6	64.0

Table 2. Parameters of the generalized circular distributions for wind direction in Vienna and Gross Enzersdorf. The probability density function is defined by:
$C.f(x) = \exp d_1 \cos(x-k_1) \cdot \exp d_2 \cos(2x-k_2)$.

		A_1	B_1	A_2	B_2	d_1	k_1	d_2	k_2
Vienna	1936-40	-0.2500	0.1019	-0.3163	-0.1857	0.4275	283	0.7648	240
	1941-44	-0.2654	0.1113	-0.2893	-0.1846	0.4604	286	0.5192	238
	1947-50	-0.2857	0.0815	-0.2580	-0.2097	0.4939	277	0.6747	233
	1951-55	-0.2754	0.0807	-0.2577	-0.1823	0.4857	276	0.6346	237
	1956-60	-0.2378	0.0972	-0.2499	-0.1055	0.4413	284	0.5280	250
	1961-65	-0.2511	0.1168	-0.2696	-0.0409	0.4968	284	0.5269	266
	1966-70	-0.2761	0.0689	-0.2532	-0.2124	0.4772	274	0.6763	233
	1971-75	-0.1891	0.0380	-0.2558	-0.2333	0.3157	269	0.7412	229
	1976-80	-0.2045	0.0589	-0.1770	-0.2051	0.3476	283	0.5420	222
Gross Enzersdorf	1936-40	-0.2384	0.0528	-0.3210	0.0549	0.5680	263	0.7136	286
	1941-44	-0.2432	0.0361	-0.2910	-0.1529	0.4683	261	0.6943	246
	1947-50	-0.2201	0.0914	-0.4539	-0.0664	0.4353	269	0.9950	263
	1951-55	-0.2096	0.0768	-0.4394	-0.1407	0.3869	268	1.0237	252
	1956-60	-0.2077	0.0738	-0.4049	-0.1103	0.3909	269	0.9041	255
	1961-65	-0.1211	0.1451	-0.4413	0.1637	0.2691	311	1.0547	293
	1966-70	-0.2092	0.1224	-0.3562	0.1703	0.5117	277	0.8851	301
	1971-75	-0.0952	0.0925	-0.2734	0.1135	0.2184	305	0.6122	294
	1976-80	-0.0953	0.0923	-0.2735	0.1133	0.2182	305	0.6123	294

Table 3. Representation of the distribution of wind directions in Vienna, 1976 to 1980, by a generalized circular distribution. The probability density function is:
$f(x) = \frac{1}{8.9268} \exp 0.3479 \cos (x - 283) \cdot \exp 0.5439 \cos (2x - 222)$.

Direction	0	45	90	135	180	225	270	315
Observed	0.097	0.074	0.095	0.140	0.070	0.067	0.277	0.179
Calculated	0.081	0.065	0.119	0.120	0.069	0.094	0.236	0.216

REFERENCES

Böhm, R. und K. Gabl, 1978: Die Wärmeinsel einer Großstadt in Abhängigkeit von verschiedenen meteorologischen Parametern. Archiv Meteor. Geoph. Biokl., B 26, 219-237.

Böhm, R., 1979: Untersuchungen des urbanen Einflusses auf das Niederschlagsfeld in der Umgerbung einer Großstadt mit Hilfe verschiedener Methoden. Archiv Meteor. Geoph. Biokl., B 27, 225-246.

Cehak, K., 1969: Beiträge zur statistischen Meteorologie. Lecture given at the ETH Zurich.

Gumbel, E.J., J.A. Greenwood and D. Durand, 1953: The circular normal distribution: theory and tables. J. Amer. Stat. Assoc., 48, 131-152.

Mises, R.v., 1918: Über die Ganzzahligkeit der Atomgewichte und verwandte Fragen. Phys. Z., 19, 490-500.

Sneyers, R., 1975: Sur l'analyse statistique des séries d'observations. OMM, Note technique No. 143, 135.

Sneyers, R. and J. Van Isacker, 1979: A generalized circular distribution. In: Ed. S. Ikeda, Proc. 1st Intern. Conf. on Stat. Clim., 21-25.

Sneyers, R., 1980: Sur l'emploi de la répartition circulaire généralisée. Archiv Meteor. Geoph. Biokl., B 28, 217-226.

Zentralanstalt für Meteorologie und Geodynamik in Wien, 1936-1980: Jahrbücher.

PROGRESSIVE ANALYSIS FOR DEPICTING TRENDS OR JUMPS IN CLIMATOLOGICAL SERIES USING KENDALL'S COEFFICIENT

J. Casimiro Mendes and F. Reimão Ferrão
Instituto Nacional de Meteorologia e Geofísica
Lisboa - Portugal

1. INTRODUCTION

On carriyng out the statistical analysis of climatological samples we find examples of an acceptable adjustment of a probability distribution to a set of data which present significant heterogeneities not only as a result of the sistematic errors produced by changes in the observation conditions, but also as a result of a continuous man-made or natural transformation in environment.

Due to this fact the methodology advised by Sneyers [1] has been followed by the authors, first by testing the hypothesis of fitting a simple random model to the series of observations, considering as alternative hypothesis the positive serial correlation of consecutive observations and the correlation between the observed values and time.

However, it has often been noted that this analysis to the total series is insuficient to know the characteristics of the series as to its homogeneity. In some cases we accept the hypothesis of randomness to the total series when the history of the meteorological station mentions changes in the observation conditions, which result in heterogeneities within the series. In other cases, when the nule hypothesis has to be rejected due to the sgnificance of the correlation between the series and time, the study does not provide any information on the form of the heterogeneity and its location.

In this conditions it seems indispensable to complete always the study of the randomness of the total series by analysing it progressively by means a test of time correlation.

Some examples with heterogeneities under the form of constant jumps or linear trends will be presented and analysed in this text. Other simple or complex forms of heterogeneity will be analysed later. No theoretical analysis of the problem is presented, use being made just of a graphic process capable of showing the sensibility of the progressive analysis in the distinction between the two mentioned heterogeneity types and their locations, applying Kendall's coefficient. For this purpose the comparative study has been made between generated series and climatological series observed in meteorological stations whose operational history is known.

2. TESTS APPLIED

2.1. TEST AGAINST POSITIVE SERIAL CORRELATION OF LAG 1.

Generally the two following tests are applied [2]: Wald-Wolfowitz's test [3] if the series makes the statistics of a test of normality significant and Anderson's test [4,5] if it does not. If the sample size n is sufficiently large, such as in the examples presented, the test of the serial correlation is used under the approximate form [1]

$$u(r) = \frac{(n-1)r+1}{\sqrt{n-1}}$$

where r is the sample estimate of the serial correlation coefficient and $u(r)$ is conside

red normally distributed with mean zero and variance unity.

2.2. TEST AGAINST TIME CORRELATION

Given a series of n observations

$$x_1 \; x_2 \ldots . \; x_i \ldots . x_j \ldots . x_n$$

let be $n_j = \sum_{i=1}^{j-1} v_{ji}$, with $j > 1$, $n_1 = 0$

and

$$v_{ji} = \begin{cases} 1 & \text{if } x_j > x_i \\ 0,5 & \text{if } x_j = x_i \\ 0 & \text{if } x_j < x_i \end{cases}$$

For each partial series $x_1 \ldots . x_k$, with $k \leq n$, is obtained Kendall's correlation coefficient

$$t_k = \sum_{j=1}^{n} n_j$$

and the value $u(t_k) = \dfrac{4t_k - n(n-1)}{4} \cdot \sqrt{\dfrac{72}{n(n-1)(2n+5)}}$ whose asymptotic distibution is a reduced normal one.

According to Sneyers [1] the series can be progressively analysed by means of two sequences of values $u(t_k)$:

a) a sequence of values u corresponding to the partial series beginning in x_1 and ending successively in x_k, with $k=1,\ldots,n$;

b) a sequence of values u' corresponding to the series beginning successively in x_k and ending in x_n.

The analysis of the graph of these two sequences will enable us to draw some conclusions on the type and location of simple heterogeneities.

3. ANALYSIS OF GENERATED AND ACTUAL SAMPLES

3.1. GENERATED SAMPLES

Four hypothetical cases were selected for comparison with actual examples which will be presented later.

A sample of 100 random independent normal numbers ε_i with mean zero and variance unity has been generated previously and after the following changes were introduced separately into this series: a constant jump so that $x_i = \varepsilon_i + 1$ from $i = 30$ to $i = 100$, another constant jump so that $x_i = \varepsilon_i + 1$ from $i = 75$ to $i = 100$ and a linear trend in the form $x_i = \varepsilon_i + 0,02i$ from $i = 30$ to $i = 100$.

In order to test the normality of these four series the probability α_Δ associated with the sample estimate of the coefficient of skewness was obtained by means of the D'Agostino's normal transformation [6] and a probability α_k associated with the sample estimate of the coefficient of kurtosis was estimated by means of the table of D'Agostino and Tietjen [7]. For each series these two probabilities were globaly tested by means of Fisher's test

$$\chi = -2(\ell n \alpha_\Delta + \ell n \alpha_k)$$

which has a chi square distribution with 4 degrees of freedom. As a result of these tests, the hypothesis of normality is acceptable for the four examples at a significance level of 0.05.

Relating to the total series, the probabilities associated with the statistics of asymptotic Wald-Wolfowitz's and Kendall's tests and with Fisher's test relating to those two tests are shnown in Table 1.

Concerning the progressive and regressive analysis of the series, Figures 1 to 4 show the graphical representation of the standardized values of Kendall's coefficient u (unbroken line) and u' (dashed line).

These graphs suggest the following remarks:

Figure 1 - Characteristical curves u and u' of a simple random series which does not show parts affected by trends or jumps. For a significance level of 0.05 the whole curves are confined within the confidence interval.

Figures 2 and 3 - In first place it should be stressed that these two series

admit the adjustment of a reduced normal model, although showing heterogeneities that make significant the statistics of the serial correlation test and of the time correlation test. Yevjevich [8] refers to the need of introducing a constant jump of $\delta > 2$ so that the series modified may be no longer unimodal, whatever the place where the jump is introduced.

As it can be noted through the graphs, the introduction of a constant jump with a value high enough to make significant the statistics of the time correlation test relating to the total series results in a good definition of the location of the jump from the point where sequences u and u' cross.

It can also be noted that in both examples there are parts of curves u and u' corresponding to the modified and unmodified partial series, respectively, which result of translations of the same curves corresponding to the simple random series ε_i and that these parts begin in points that are symetrical relative to the crossing point of u and u'. If the jump is located fairly near the beginning of the series the progressive values u show a graph with a steep short slope, extending then in a form similar to the graph of u corresponding to the simple random series ε_i. If the jump is near the end of the series a steep slope and a well developed section of stationarity appear in the regressive sequence u'.

Figure 4 - In this case, where the trend introduced persists up to the end of the series, the sequences of values u and u' do not show the sections of stationarity mentioned in the previous examples and the crossing point of the graphs of these sequences does not define the beginning of the trend satisfactorily.

3.2. ACTUAL EXAMPLES OF CLIMATOLOGICAL SERIES

Figures 5 to 8 show four examples of climatological series whose homogeneity is going to be analysed by comparing the graphs of the values u and u' for each example with the generated series already studied. The upper part of each figure shows the observed values of the meteorological event and the lower part shows the curves for the values u (unbroken line) and u' (dashed line).

Figure 5 - Series of observations respecting the highest values of maximum daily temperature of the air (^{o}C) in Santarém in August from 1924 to 1981. According to the values shown in Table 1 and to the comparison of u and u' curves in this figure with those shown in Figure 1, the conclusion is drawn that the hypothesis of a simple random series is acceptable.

The history of this meteorological station does not mention significant changes since the beginning of its operation.

Figure 6 - Series of mean daily temperature values (^{o}C) respecting November in Lisbon from 1856 to 1981. Considering the values in Table 1, the rejection of the hypothesis of a simple random series is to be attributed mainly to the significance of Kendall's coefficient. The curves of u and u' show clear stationarity sections, similar to what happens in the case of the series shown in Figure 3.

Therefore, it is to be accepted that this climatological series shows a constant jump from a date between 1935 and 1945. In fact, the history of the station mentions a change in the observation site in 1940 and it is known that, up to that date, the thermometer was placed at a height of $h_t = 23.1$ m and from 1941 to 1980 it is a height $h_t = 1.6$ m.

Figure 7 - Series of mean values of maximum daily temperature of the air (^{o}C) respecting August in Lisbon from 1856 to 1980. Like in the previous example, the rejection of the hypothesis of randomness is to be attributed to the significance of the statistics of Kendall's test.

The curves u and u' do not show sections of stationarity, and so the existance of a trend is to be admitted, in comparison with Figure 4. The year 1917 seems to

	WALD.WOLF.	KENDALL	FISHER
Fig. 1	0.4"4	0.596	0.640
Fig. 2	0.042	0.001	0.000
Fig. 3	0.020	0.002	0.000
Fig. 4	0.022	0.000	0.000
Fig. 5	0.460	0."'8	0.726
Fig. 6	0.130	0.005	0.005
Fig. "	0.272	0.000	0.000
Fig. 8	0.004	0.63"	0.019

TABELA 1—SIGNIFICANT LEVELS FOR RANDOMNESS

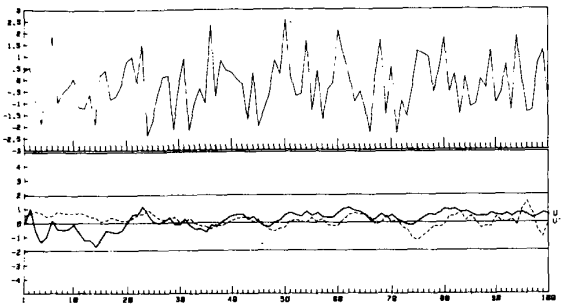


Fig. 1—Generated sample e_i.

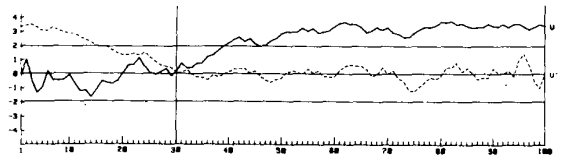


Fig. 2—Generated sample $e_i + 1.0$, $i = 30$ to $i = 100$.

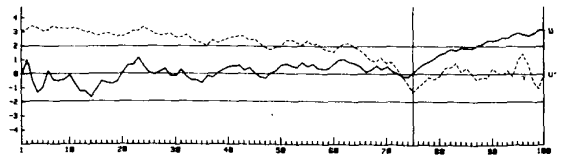


Fig. 3—Generated sample $e_i + 1.0$, $i = 75$ to $i = 100$.

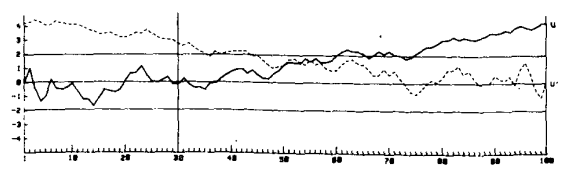


Fig. 4—Generated sample $e_i + 0.02i$, $i = 30$ to $i = 100$.

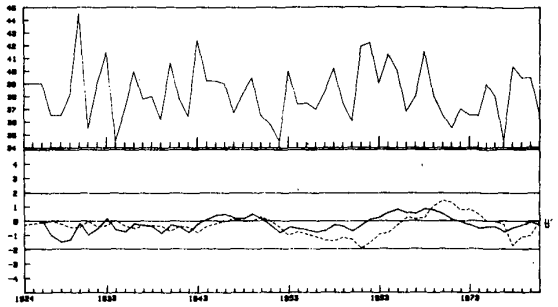


Fig. 5—Higher values of daily maxima air temperature (°C) in Auguste—Santarém.

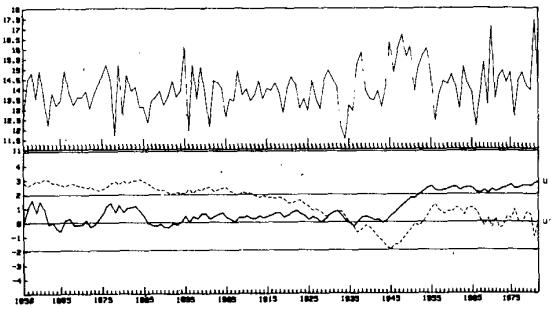


Fig. 6—Mean values of daily air temperature (°C) in November—Lisboa.

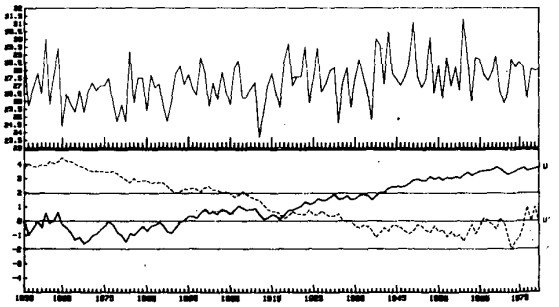


Fig. 7—Mean values of daily maxima air temperature (°C) in Auguste—Lisboa.

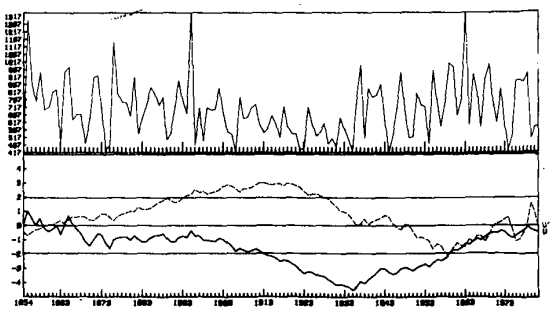


Fig. 8—Annual amountes of precipitation (mm) in Lisboa.

be a first approximation to the beginning of this trend, nevertheless it must have occurred before that date. On analysing the progressive curve u in more detail it can be noted that it shows a section between 1940 and 1950 with a steeper slope, which probably corresponds to a jump caused by a change in the observation conditions in 1940, as referred to in the previous example. For a more datailed study of the trend, concerning its causes and a more accurate location of its beginning, it is necessary to obtain more information from other sources, especially from other climatological series. It should be underlined, in first place, that Sneyers [1] has found a heterogeneity of this type with similar features for the series of mean annual values of air temperature in Bruxelles-Uccle.

Figure 8 - Series of the annual values of precipitation amount in Lisbon from 1854 to 1982. Although Kendall's coefficient is not significant for the whole series, it shows significant values within the series. The central part of the graph shows a significant divergence of the curves u and u', corresponding to a succession of years in which a clear change occurs in the series of the values observed. The information available is that there has been a change in the height of the raingauge place at this station in 1930. Up to 1929 the raingauge was placed at a height $h_r = 22.8$ m and from 1930 to 1982 it was placed at $h_r = 1.6$ m. It should be underlined that about 1930 the u' values became significant and the u values, which revealed a negative trend, began to undergo a change in the opposite direction that makes the curve take progressively a form corresponding to a simple random series. This climatological series cannot be satisfactorily analysed by comparison with the graphs of the generated series presented in this work. This series has been presented just to show that the statistical analysis of a series as a whole does not enable, in some cases, to detect the alternative hypotheses when the hypothesis of a simple random series is to reject.

4. CONCLUDING REMARKS

From the examples presented it can be concluded that the progressive analysis of climatological series using Kendall's coefficient is a fairly sensitive statistical method, which provides good information on the form and location of heterogeneities. Therefore, we consider its use most desirable on taking a decision on the random features of a series. This text includes only a graphical process of series modified by a constant jump or a linear trend, which are particularly simple modifications in the various cases actually occurring. A deeper study of this matter will certainly enable qualitative and quantitative results to be obtained, providing more accurate statistical analysis of heterogeneities, not only of the simple types presented, but of more complex ones as well.

5. REFERENCES

1. Sneyers, R., 1975: Sur l'analyse statistique des series d'observations. Note Technique Nº 143, Org. Met. Mondiale.

2. Mitchell, G.M. et al.: Climatic change. Technical Note Nº 79, World Met. Organization.

3. Wald, A. and Wolfowitz, G., 1943: An exact test for randomness in the non-parametric case based on serial correlation. Annals Math. Stat., 14, 4, 378-388.

4. Andersin, R.L., 1942: Distribution of the serial correlation coefficient. Annals. Math. Stat., 13, 1, 1-13.

5. Dixon, W. 1944: Further contributions to the problem of serial correlation. Annals Math. Stat., 14, 2, 119-144.

6. D'Agostino, R.B., 1970: Transformation

to normality of the null distribution of g_1. Biometrika, 57, 679-681.

7. Pearson, E.S. and Hartley, H.O., 1962: Biometrika tables for statisticians, Vol. 1, Cambridge University Press.

8. Yevjevich, V. and Jeng, R.I., 1969: Properties of non-homogeneous hydrologic series. Hydrology papers, Colorado State University, 32.

ERRATA

page		instead of	read
10	dellete the 3rd. line		
11	Session 9 paper 9.6	9.1.1-6	9.1.1-4
11	Session 10 paper 10.2	10.2.1-5	10.2.1-6
13	Session 11 paper 11.8	11.8.1-8	11.8.1-6
14	Session 14 paper 14.1	**	14.1.1-8
16	6th. line	Chiarmen	Chairmen
2.2.1	1st. line	Genrealized	Generalized
16.7.1			16.6.1
16.7.2			16.6.2

20421780

ENV. CAN. LIBR./BIB. DOWNSVIEW
QC 980 I583 1983 c.1
II International Meeting on Statistical.
0 0002 04217804 1

QC
980
I583
1983

LIBRARY
ATMOSPHERIC ENVIRONMENT SERVICE
4905 DUFFERIN ST.
DOWNSVIEW, ONTARIO
CANADA M3H 5T4

Date Due

MAR	4 1986		
MAR	4 1986		
JUN 3	1986		
8/6/06	24		
OCT 29	1986		

NLR 174

MONTAGEM E IMPRESSÃO NA DIVISÃO GRÁFICA DO
INSTITUTO NACIONAL DE METEOROLOGIA E GEOFÍSICA
Rua C Aeroporto 1700 LISBOA — PORTUGAL