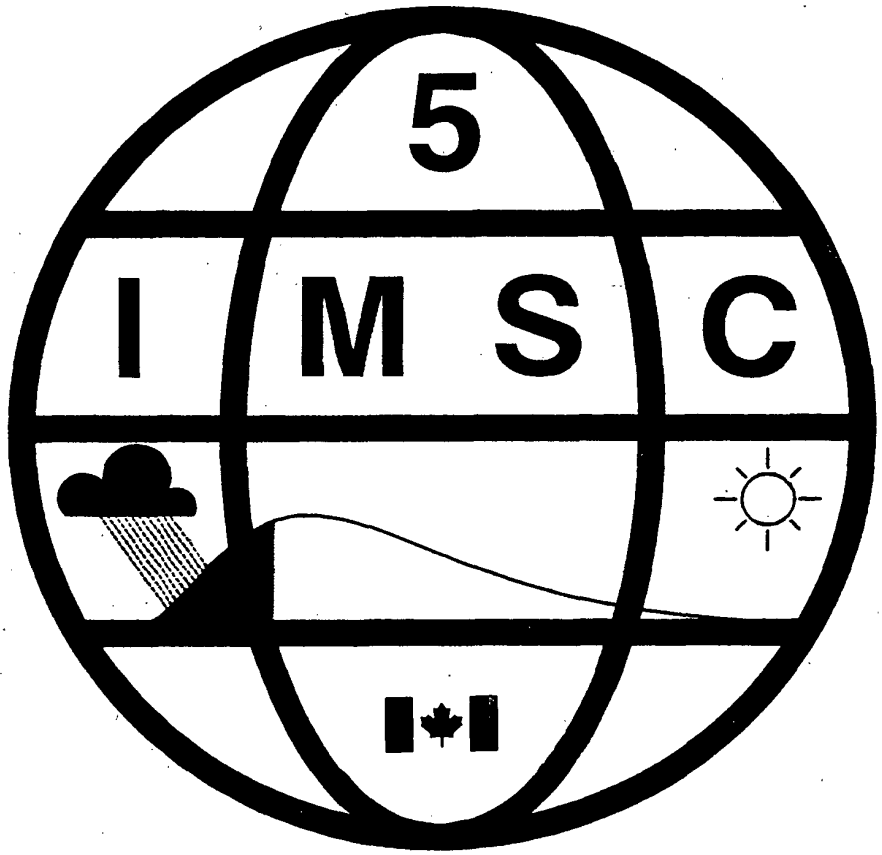
Environment Canada
Environnement Canada

Atmospheric Environment Service
Service de l'environnement atmosphérique

5th International Meeting on Statistical Climatology



22-26 June 1992
Toronto, Canada

Canada

Note: Co-authors will appear in *italics*.

5th International Meeting
on
Statistical Climatology

22-26 June 1992

Toronto, Canada

Organizer

The Steering Committee
for
Interational Meetings on Statistical Climatology

Cosponsors

American Meteorological Society
American Statistical Association
Atmospheric Environment Service of Canada
Bernoulli Society
Canadian Meteorological and Oceanographic Society
Statistical Society of Canada
Sun Microsystems of Canada
World Meteorological Organization

Steering Committee
International Meetings on Statistical Climatology

Allan Murphy (Chair - Oregon State University)
George Gruza (Institute for Global Climate and Ecology - Russia)
Neville Nicholls (Bureau of Meteorology Research Center - Australia)
Ignacio Rodriguez-Iturbe (Instituto International de Estudios Avanzados - Venezuela)
Shyamvir Singh (Indian Institute of Tropical Meteorology)
Raymond Sneyers (Royal Meteorological Institute of Belgium)
Francis Zwiers (Canadian Climate Centre)

Program Committee
5th International Meeting on Statistical Climatology

Francis Zwiers (Chair - Canadian Climate Centre)
Peter Bloomfield (North Carolina State University)
David Brillinger (University of California - Berkeley)
Nicholas Fisher (CSIRO)
Robert Livezey (Climate Analysis Center - Washington)
Roland Madden (NCAR)
Paul Mielke (Colorado State University)
Allan Murphy (Oregon State University)
Nancy Reid (University of Toronto)
Benjamin Santer (Max Planck Institute for Meteorology)
Raymond Sneyers (Royal Meteorological Institute of Belgium)
Andrew Solow (Woods Hole Oceanographic Institution)
Hans von Storch (Max Planck Institute for Meteorology)

Notation Convention for Paper Numbering

Papers listed in the table of contents are numbered using the following conventions: Contributed papers which were accepted for oral presentation are listed with prefix "C"; Contributed papers which were accepted for poster presentation are listed with prefix "P"; Reserve papers, which were accepted for presentation if time becomes available, are listed with prefix "R"; Invited papers are designated with prefix "I"; Papers presented jointly to both 5IMSC and the 12th Conference on Probability and Statistics in the Atmospheric Sciences (12PSAS) are listed with prefix "J". All joint sessions were organized by invitation. The joint papers appear at the back of this volume.

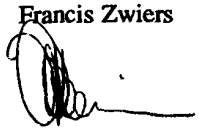
Preface

This meeting, the fifth in the sequence of International Meetings on Statistical Climatology, represents the fruition of an idea that was born about 15 years ago. The concept which was the basis of the first meeting in 1980 was that statistics is an essential part of climatology and that a full understanding of our climate and its mechanisms depends very much on the development of sophisticated and powerful statistical methodology. The size of this volume, the number of participants in this meeting, the disciplinary diversity of the participants and their wide geographical dispersion all provide evidence of the extent to which this idea has matured and flowered.

The IMSC meetings are the children of a small steering committee of like minded scientists that is chaired by Prof. Allan Murphy. To the credit of the steering committee, each IMSC meeting has been allowed to have its own distinct characteristics which are coloured by the research priorities of the day, the interests of the program chair and the meeting venue. All three have certainly had an impact on this meeting.

Many people and organizations contributed to the planning and execution of this meeting and the production of this preprint volume. Because any attempt to acknowledge everyone involved is certain to result in ommissions I will not be exhaustive in this regard. I am genuinely appreciative of the efforts and contributions of all the individuals and organizations who have contributed to this meeting and hope that you will accept my thanks. None the less, there are some individuals and groups whom I would like to single out for special thanks. First, the members of the program committee and their designates deserve kudos for the superb work which they have done. Secondly, this meeting enjoys the support of several financial backers who have enabled us to bring a substantial number of invited and international participants to this meeting. These backers are the Atmospheric Environment Service of Canada, the Natural Sciences and Engineering Research Council of Canada, the World Meteorological Organization and the Canadian Meteorological and Oceanographic Society. Thirdly, I would like to thank Derita Wittke for her invaluable assistance in compiling this volume and in performing the myriad of organizational and clerical tasks that are part of organizing any meeting of this type. Fourthly, I would like to thank Evelyn Mazur and the American Meteorological Society for the logistical support which they have provided. Finally, I would like to thank my colleagues in the Canadian Climate Centre for the support and encouragement they provided during the evolution of this endevour.

Francis Zwiers



Program Chairman

TABLE OF CONTENTS

5th INTERNATIONAL MEETING
ON
STATISTICAL CLIMATOLOGY

Contributed and Invited Papers

Periodic Correlation in Meteorological Time Series

Robert Lund
Department of Statistics
University of North Carolina
Chapel Hill, NC 27599-3260

Harry Hurd
Center for Stochastic Processes
University of North Carolina
Chapel Hill, NC 27599-3260

Peter Bloomfield
Department of Statistics
North Carolina State University
Raleigh, NC 27695-8203

ABSTRACT

A fifty-year time series of monthly stratospheric ozone readings from Arosa, Switzerland is analyzed. The time series exhibits the properties of a periodically correlated (PC) random sequence with annual periodicities. A test to detect periodic correlation is presented. An ARMA model with periodically varying coefficients (PARMA) is fit to the data in two stages. First, a periodic autoregressive model (PAR) is fit to the data. This fit yields residuals that are stationary, but non-white. Next, a stationary ARMA model is fit to the residuals and the two models are combined to produce a larger model for the data. The combined model turns out to be a PARMA model and yields residuals that have the correlation properties of white noise.

1. Introduction

The seasonal nature of weather imparts a seasonal structure into many meteorological time series. Frequently, the seasonal structure is in the form of periodic correlation. A random sequence $\{X_n\}$ with finite second moments is called periodically correlated (PC) with period T if $\mu(n) = \mathbb{E}[X_n]$ and $C(m,n) = Cov(X_m, X_n)$ are periodic with period T:

$$\mu(n+T) = \mu(n) \quad \text{and}$$

$$C(m+T, n+T) = C(m,n). \tag{1.1}$$

To avoid ambiguity, the period T is taken as the smallest positive integer such that (1.1) holds. When $T = 1$, $\{X_n\}$ is covariance stationary and will be referred to as stationary for short. One can always assume that $\mu(n) \equiv 0$ by examining $\{X_n - \mu(n)\}$; in practice, the periodic sample mean is subtracted from the data.

This paper is concerned with modeling the correlation structure of stratospheric ozone data. A model is developed for a data set, plotted in Figure 1, that contains 50 years of monthly observations from Arosa, Switzerland. Adequate models for ozone data are important in the prediction of future values and in the analysis of possible trends (Hill *et al.*, 1986). A natural choice for the period is $T = 12$; this choice will be statistically justified by a test presented in Section 2. Figures 2 and 3, which plot the monthly sample mean and standard deviation of the data set in Dobson units, clearly indicate that the data is nonstationary.

2. A test to detect periodic correlation

A test for detecting periodic correlation in $\{X_n\}$ against the stationary null hypothesis was recently presented by Hurd and Gerr (1991). A variant of this test will be used in our numerical work that follows. The test we propose uses the discrete Fourier transform of the data sample $\{X_0, X_1, \ldots, X_{N-1}\}$ defined by

$$I_N(\lambda_j) = \frac{1}{\sqrt{2\pi N}} \sum_{n=0}^{N-1} X_n e^{-in\lambda_j} \tag{2.1}$$

at the Fourier frequencies $\lambda_j = 2\pi j/N$ for $j = 0, 1, \ldots, N-1$. The test is based upon the following fact: if $\{X_n\}$ is stationary, then $I_N(\lambda_j)$ and $\overline{I_N(\lambda_k)}$ are uncorrelated except when $\lambda_j = \lambda_k$; if $\{X_n\}$ is PC with period T, then $I_N(\lambda_j)$ and $\overline{I_N(\lambda_k)}$ are uncorrelated except when $\lambda_j = \lambda_k + 2\pi h/T$ for some $h = 0, \pm 1, \ldots, \pm(T-1)$ (Gladyšhev, 1961). Here, the overline denotes complex conjugation. Thus, a statistic designed to find correlations in the discrete Fourier transform should prove useful in detecting periodic correlation.

For $h \in \{0, \pm 1, \ldots, \pm(N-1)\}$, define the M-point sample coherence

$$\left|\gamma_{h,M}(\lambda_j)\right|^2 = \frac{\left|\sum_{m=0}^{M-1} I_N(\lambda_{j+m})\overline{I_N(\lambda_{j+h+m})}\right|^2}{\sum_{m=0}^{M-1}\left|I_N(\lambda_{j+m})\right|^2 \sum_{m=0}^{M-1}\left|I_N(\lambda_{j+h+m})\right|^2}. \tag{2.2}$$

The quantity $|\gamma_{h,M}(\lambda_j)|^2$ is called Goodman's squared coherence statistic (Goodman, 1965) and takes values

in the interval [0,1] only. The squared coherence statistic is symmetric about the main diagonal in the bifrequency square $[0,2\pi)$x$[0,2\pi)$:

$$\left|\gamma_{h,M}(\lambda_j)\right|^2 = \left|\gamma_{-h,M}(\lambda_{j+h})\right|^2. \qquad (2.3)$$

Hence, it is sufficient to consider squared coherence statistics for $h \geq 0$ only.

For simplicity, assume that the data record contains d full years; that is, assume $d = N/T$ is an integer. Suppose that $\{X_n\}$ is PC with period T. Then a small value of $|\gamma_{h,M}(\lambda_j)|^2$ is anticipated unless h is an integer multiple of d. For terminology, let the hth diagonal line be the subset of $[0,2\pi)$x$[0,2\pi)$ that contains the linear diagonal lattice of points $(\lambda_j, \lambda_{j+h})$.

To detect periodic correlation in $\{X_n\}$, Goodman's squared coherence statistic is computed for all $h \in \{0,1, 2, \ldots, (N-1)\}$ and λ_j such that $0 \leq \lambda_j < 2\pi$. Next, the percentage of squared coherence statistics exceeding a preset threshold is computed along the hth diagonal line for each $h > 0$. The exceedance percentage is plotted against the diagonal line index h. *If $\{X_n\}$ is PC with period T, then the exceedance percentage should be small whenever h is not a multiple of d and should be large for some h's that are multiples of d. Thus, this diagonal exceedance percentage plot should reveal large values at some multiples of d.* If $\{X_n\}$ is stationary, no large values should appear in the diagonal exceedance percentage plot.

To determine the preset squared coherence threshold, the distribution of $|\gamma_{h,M}(\lambda_j)|^2$ must be known under the null hypothesis that $\{X_n\}$ is stationary. The squared coherence distribution is known explicitly only for the case where $\{X_n\}$ is mean zero Gaussian white noise:

$$\mathbb{P}[\,|\gamma_{h,M}(\lambda_j)|^2 > x\,] = (1-x)^{M-1}, \quad 0 \leq x \leq 1, \qquad (2.4)$$

(Goodman, 1965). A simulation study by Hurd and Lund (1991) has shown the beta type distribution in (2.4) to be very robust against departures from both normality and white noise. The threshold

$$t_M = 1 - (.05)^{\frac{1}{M-1}} \qquad (2.5)$$

will be used in the numerical work that follows; this threshold provides approximately a 95% degree of statistical confidence. We refer the reader to Hurd and Gerr (1991) for remarks on the selection of M.

Figure 4 displays the diagonal exceedance percentage plot for the Arosa data with the periodic sample mean removed and $M = 8$. The plot clearly shows a large exceedance percentage at $h = 50$. The squared coherence statistic exceeds the 95% threshold of $t_8 = .348$ at 41.0% of the frequencies along the line $h = 50$. With 50 years of monthly data ($N = 600$), one has evidence that $N/T = 50$ or that $T = 12$.

3. Models for PC random sequences

Periodic correlation can be introduced into an ARMA model when the coefficients of the model are allowed to vary periodically with time. ARMA models with periodically varying coefficients are frequently called PARMA models.

For clarity, the notation of Vecchia (1985) is adopted and $\{X_t\}$ is indexed by year and season: $X_{nT+\nu}$ refers to the time series during the νth season of year $n \geq 0$. The total number of seasons per year is T and the seasonal index ν satisfies $1 \leq \nu \leq T$. We say that $\{X_t\}$ follows a PARMA model if

$$X_{nT+\nu} - \sum_{k=1}^{p(\nu)} \phi_k(\nu) X_{nT+\nu-k} = \sum_{k=0}^{q(\nu)} \theta_k(\nu) \epsilon_{nT+\nu-k} \qquad (3.1)$$

where $\{\epsilon_t\}$ is mean zero white noise with $\mathbb{V}ar(\epsilon_t) \equiv \sigma^2$. The coefficients $\phi_k(\nu)$ and $\theta_k(\nu)$, and the orders $p(\nu)$ and $q(\nu)$ depend on the season and are extended periodically in the variable ν to all integers.

Model 3.1 has a total of $T + 1 + \sum_1^T \{p(\nu) + q(\nu)\}$ parameters. In the analysis of the Arosa data, the simple first order periodic autoregressive model

$$X_{nT+\nu} = \phi(\nu) X_{nT+\nu-1} + \sigma(\nu)\epsilon_{nT+\nu} \qquad (3.2)$$

will be used. Here, $\{\epsilon_t\}$ is mean zero white noise with $\mathbb{V}ar(\epsilon_t) \equiv 1$. Model 3.2 has $p(\nu) \equiv 1$, $q(\nu) \equiv 0$, $\phi_1(\nu) = \phi(\nu)$, $\theta_0(\nu) = \sigma(\nu)$, and a total of $2T$ parameters. A special case of model 3.2 with a constant $\phi(\nu) \equiv \phi$ was used to analyze ozone data by Reinsel and Tiao (1987).

Properties of model 3.2 can be found in Bloomfield *et al.* (1992). In particular, it is shown there that model 3.2 generates a PC time series with period T whenever $|\phi(1)\phi(2) \ldots \phi(T)| < 1$. This result holds even when $\{\epsilon_t\}$ is stationary. Explicit forms for the periodic variance and the seasonal correlation functions can also be found there.

We now consider parameter estimation for model 3.2. Approximate maximum likelihood estimates can be obtained under the assumption of a normally distrib-

uted error sequence $\{\epsilon_t\}$. The approximation arises by setting $X_0 = 0$. For notation, let $\vec{X} = (X_1, X_2, \ldots, X_N)'$ be the data vector and let $L(\vec{X}; \vec{\sigma},\vec{\phi})$ denote the approximate likelihood function evaluated at $\vec{X}$ when the parameter vectors are $\vec{\phi} = (\phi(1), \phi(2), \ldots, \phi(T))'$ and $\vec{\sigma} = (\sigma(1), \sigma(2), \ldots, \sigma(T))'$ respectively. Bloomfield *et al.* (1992) show that $\vec{X}$ has a N-variate normal distribution and that

$$-\log\left\{L(\vec{X}; \vec{\phi},\vec{\sigma})\right\} = \frac{N}{2}\log(2\pi) + d\sum_{\nu=1}^{T}\log\{\sigma(\nu)\}$$
$$+\frac{1}{2}\sum_{n=0}^{d-1}\sum_{\nu=1}^{T}\left(\frac{X_{nT+\nu} - \phi(\nu)X_{nT+\nu-1}}{\sigma(\nu)}\right)^2. \quad (3.3)$$

Explicitly minimizing (3.3) via differentiation produces the approximate maximum likelihood estimates

$$\hat{\phi}(\nu) = \frac{\hat{\gamma}_\nu(1)}{\hat{\gamma}_{\nu-1}(0)}, \qquad \hat{\sigma}^2(\nu) = \hat{\gamma}_\nu(0) - \hat{\phi}(\nu)\hat{\gamma}_\nu(\nu),$$

$$\text{where} \quad \hat{\gamma}_\nu(j) = \frac{1}{d}\sum_{n=0}^{d-1} X_{nT+\nu}X_{nT+\nu-j}. \quad (3.4)$$

In the computation of $\hat{\gamma}_\nu(j)$, one takes $X_i = 0$ for $i \le 0$; $\hat{\gamma}_\nu(j)$ is also interpreted periodically in the variable ν. When restrictions exist between the parameters, one can always numerically minimize the negative log likelihood in (3.3) over the variables of interest. This issue will arise in Section 4 when the total number of parameters modeling the Arosa data set is reduced.

To evaluate the exact likelihood function, the covariance matrix of $\vec{X}$ must be computed explicitly. In principle, this covariance matrix can be obtained from Theorem 1 in Bloomfield *et al.* (1992); in practice, this computation is very tedious and can be bypassed completely with the approximation $X_0 = 0$.

4. Model development for the Arosa data

Figures 5 and 6 plot $\hat{\phi}(\nu)$ and $\hat{\sigma}(\nu)$ for the mean subtracted Arosa data as computed from (3.4). A negative log likelihood of 2408.535 was obtained. This 24-parameter model will be called the full model. Notice the approximate sinusoidal shape of $\hat{\sigma}(\nu)$ in Figure 6. To evaluate the fit of the full model, the residuals

$$\hat{\epsilon}_{nT+\nu} = \frac{X_{nT+\nu} - \hat{\phi}(\nu)X_{nT+\nu-1}}{\hat{\sigma}(\nu)} \quad (4.1)$$

with $\hat{\epsilon}_1 = X_1/\hat{\sigma}(1)$ are examined. If the full model fit is good, the residuals should have correlation properties similar to those of white noise.

Figure 7 displays the diagonal exceedance percentage plot of the residuals with $M = 8$. The plot shows that the large exceedance percentage at $h = 50$ in Figure 4 has been removed; this provides statistical evidence that the residuals are from a stationary sequence. Figure 8 plots the periodogram of these residuals. The periodogram has a U-shaped feature which indicates that the residuals may not be white noise. To test this hypothesis statistically, the portmanteau test (see Brockwell and Davis, 1987) is applied to the residuals over the first 25 lags. This test sums the square of the residual's sample correlation function over the first 25 lags. Under the null hypothesis that the residuals are white noise, one anticipates a small test statistic. In this case, the value of the portmanteau test statistic was 68.823 which gives a p-value of $5.75\text{x}10^{-6}$ when compared to the chi-squared, 25 degrees of freedom, asymptotic, null hypothesis distribution. Thus, there is statistical evidence that the full model has removed the periodic correlation in the data, and that the error terms belong to a stationary, but non-white random sequence.

One can attempt to reduce the number of parameters in the full model by parametrizing $\phi(\nu)$ and $\sigma(\nu)$ as a short Fourier series (Jones and Brelsford, 1967). Parametrize $\phi(\nu)$ and $\sigma(\nu)$ via

$$\sigma^2(\nu) = \alpha\left\{1 + \rho\cos\left(\frac{2\pi(\nu-\tau)}{12}\right)\right\}, \quad \phi(\nu) \equiv \phi. \quad (4.2)$$

This four-parameter model will be called the reduced model. Approximate maximum likelihood estimates of the parameters in (4.2) can be obtained by numerically minimizing the negative log likelihood in (3.3). This was performed and produced $\hat{\alpha} = 231.579 \pm 14.603$, $\hat{\rho} = 0.787 \pm .031$, $\hat{\tau} = 1.786 \pm .160$, and $\hat{\phi} = .293 \pm .038$. A negative log likelihood of 2421.057 was produced. Uncertainties are one standard error and were calculated from an approximation to the observed information matrix. The estimates of $\phi(\nu)$ and $\sigma(\nu)$ for the reduced model are plotted against the full model estimates in Figures 5 and 6. Note the approximate "least squares" fit of the reduced model estimates to the full model estimates.

Figure 9 presents the diagonal exceedance percentage plot of the reduced model's residuals with $M = 8$. Again, the large exceedance percentage at $h = 50$ is absent. Figure 10 plots the periodogram of the reduced model's residuals. This periodogram displays the same U-shaped feature encountered with the full model's residuals. Applying the portmanteau test to the reduced model's residuals over the first 25 lags produces a test statistic of 67.756 and a p-value of $8.25\text{x}10^{-6}$. Thus, there is statistical evidence that the reduced model's residuals are stationary, but not white noise.

The 24-parameter full model and the four-parameter

reduced model both produced residuals that were stationary, but non-white. Twice the difference between the negative log likelihoods of the reduced and full models is 25.044. This produces a p-value of .200 when compared to the chi-squared distribution with 20 degrees of freedom. Hence, we find the reduced model preferable to the full model.

A stationary ARMA model is next fit to the residuals of the reduced model. The optimal ARMA model for these residuals as selected by both the AIC and BIC criterions (Brockwell and Davis, 1987) is the ARMA(2,1) model

$$\hat{\epsilon}_l - \gamma_1 \hat{\epsilon}_{l-1} - \gamma_2 \hat{\epsilon}_{l-2} = \omega_l + \beta_1 \omega_{l-1}. \qquad (4.3)$$

The maximum likelihood estimates of the parameters in (4.3) and their approximate standard errors are $\hat{\gamma}_1 = .644 \pm .080$, $\hat{\gamma}_2 = .206 \pm .042$, and $\hat{\beta}_1 = -.738 \pm .077$. The estimated white noise variance of $\{\omega_n\}$ is 0.943. The model fitted in (4.3) yields $\mathbb{E}[\hat{\epsilon}_l^2] \equiv 1.0004$ which is roughly consistent with the model 3.2 assumption that $\mathbb{E}[\epsilon_l^2] \equiv 1$.

The fitted ARMA model in (4.3) and the fitted reduced model can be algebraically combined to produce a PARMA model with the seasonal orders $p(\nu) \equiv 3$ and $q(\nu) \equiv 1$. The coefficients of this model can be obtained from

$$\hat{\theta}_0(\nu) = \hat{\sigma}(\nu), \quad \hat{\theta}_1(\nu) = \hat{\beta}_1 \hat{\sigma}(\nu), \quad \hat{\phi}_1(\nu) = \hat{\phi} + \frac{\hat{\gamma}_1 \hat{\sigma}(\nu)}{\hat{\sigma}(\nu-1)},$$

$$\hat{\phi}_2(\nu) = \hat{\sigma}(\nu)\left(\frac{\hat{\gamma}_2}{\hat{\sigma}(\nu-2)} - \frac{\hat{\phi}\hat{\gamma}_1}{\hat{\sigma}(\nu-1)}\right),$$

$$\hat{\phi}_3(\nu) = \frac{-\hat{\phi}\hat{\gamma}_2\hat{\sigma}(\nu)}{\hat{\sigma}(\nu-2)}, \qquad (4.4)$$

for $\nu = 1, 2, \ldots, T$ and the estimated parameters in (4.2) and (4.3). Recall that all parameters in (4.4) are interpreted periodically in the variable ν.

As a final check, the residuals of the combined model are analyzed. These residuals are computed recursively via

$$\hat{\omega}_{nT+\nu} = \qquad (4.5)$$

$$\frac{X_{nT+\nu} - \sum_{k=1}^{p(\nu)} \hat{\phi}_k(\nu) X_{nT+\nu-k} - \sum_{k=1}^{q(\nu)} \hat{\theta}_k(\nu) \hat{\omega}_{nT+\nu-k}}{\hat{\theta}_0(\nu)},$$

where $X_i = 0$ and $\hat{\omega}_i = 0$ for $i \leq 0$.

Figure 11 displays the diagonal exceedance percentage plot of the combined residuals when $M = 8$. The plot shows no large exceedance percentages; hence, the combined model has removed the periodic correlation in the data. Figure 12 plots the periodogram of the combined model's residuals. This periodogram does not appear to deviate sharply from a white spectrum. Applying the portmanteau test over the first 25 lags to the combined residuals produces a test statistic of 22.728. The p-value of this test is .593. Thus, there is statistical evidence that the combined residuals are white noise. Hence, the combined model appears to fit the data well.

One can check that $|\hat{\phi}(1)\hat{\phi}(2) \ldots \hat{\phi}(T)| < 1$ for both the full and reduced models. Thus, the combined fitted PARMA model is PC with period T.

Acknowledgements. R. Lund and H. Hurd were supported by the Office of Naval Research, Contract No. N00014-86-C-0227. P. Bloomfield was supported by the National Science Foundation, Grant No. DMS-8610127.

REFERENCES

Bloomfield, P., H. L. Hurd and R. B. Lund, 1992: Periodic Correlation in Stratospheric Ozone Data, submitted to *Journal of Time Series Analysis.*

Brockwell, P. J., and R. A. Davis, 1987: *Time Series: Theory and Methods,* Springer-Verlag, 519 pp.

Gladyšhev, E. G., 1961: Periodically Correlated Random Sequences, *Soviet Math,* **2**, 385-388.

Goodman, N. R., 1965: *Statistical Tests for Stationarity Within the Framework of Harmonizable Processes,* Rocketdyne Research Report No. 65-28.

Hill, W. J., G. W. Oehlert and G. C. Reinsel, 1986: Trend Analysis Sensitivity Studies of Dobson Total Ozone Data Through 1984, *Journal of Geophysical Research,* **91**, 14515-14520.

Hurd, H. L., and N. L. Gerr, 1991: Graphical Methods for Determining the Presence of Periodic Correlation, *Journal of Time Series Analysis,* **12**, 337-350.

Hurd, H. L., and R. B. Lund, 1991: *A Sensitivity Study of Goodman's Coherence Statistic Via Simulation,* Technical Report, Harry L. Hurd Associates.

Jones, R. H., and W. M. Brelsford, 1967: Time Series with Periodic Structure, *Biometrika,* **54**, 403-408.

Reinsel, G. C., and G. C. Tiao, 1987: Impact of Chlorofluoromethanes on Stratospheric Ozone, *Journal of the American Statistical Association* **82**, 20-30.

Vecchia, A. V., 1985: Periodic Autoregressive-Moving Average (PARMA) Modeling with Applications to Water Resources, *Water Resources Bulletin,* **21**, 721-730.

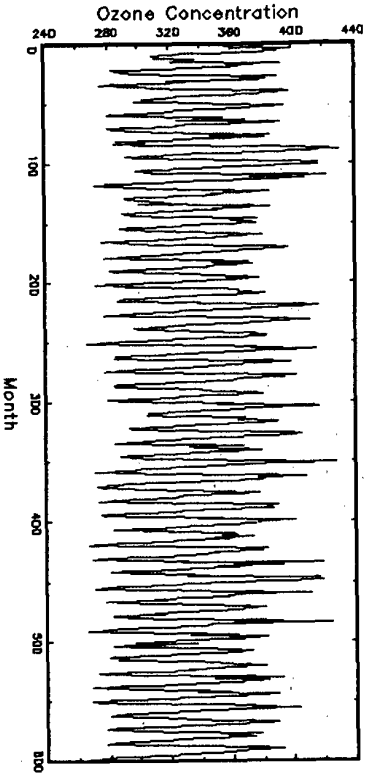


Figure 1: The Arosa Data

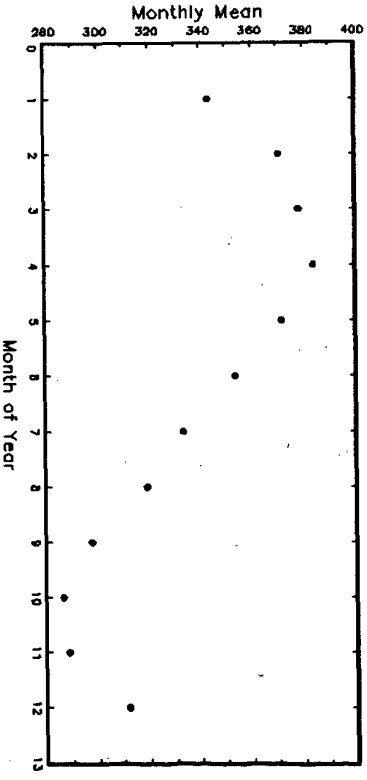


Figure 2: Monthly Mean of the Data

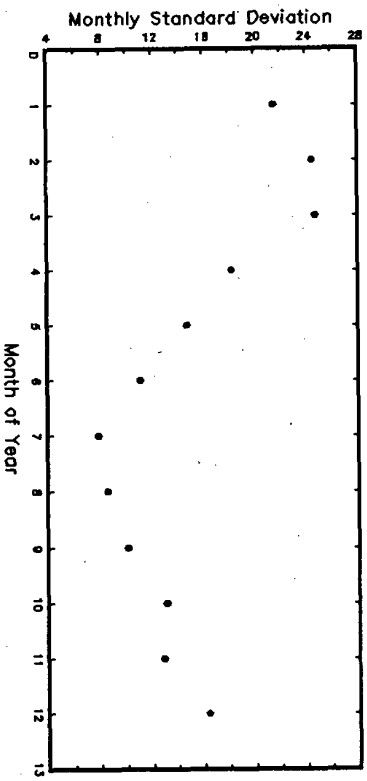


Figure 3: Monthly Standard Deviation of the Data

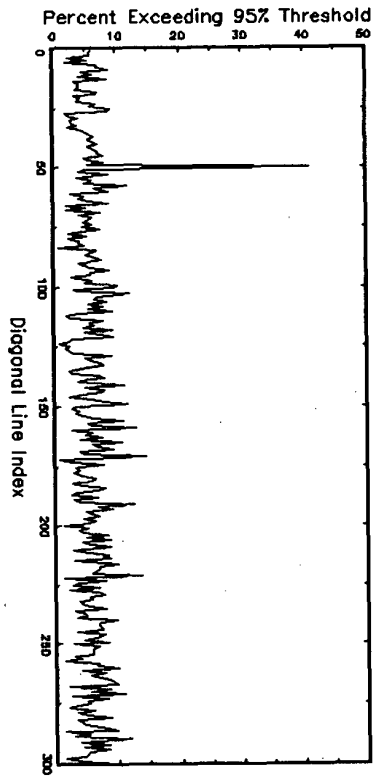


Figure 4: Diagonal Exceedance Percentage Plot for the Arosa Data

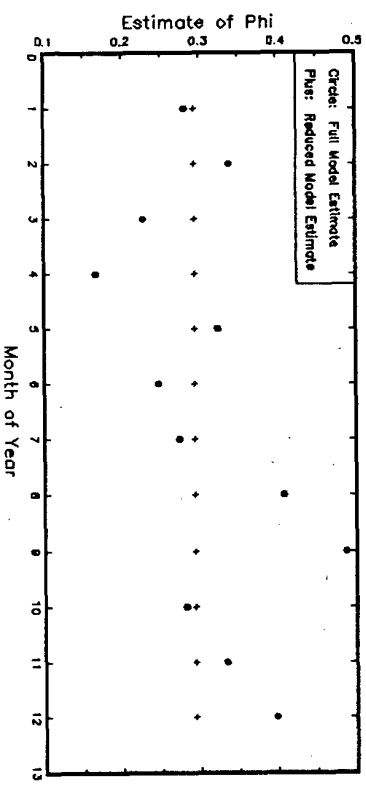


Figure 5: Estimates of Phi

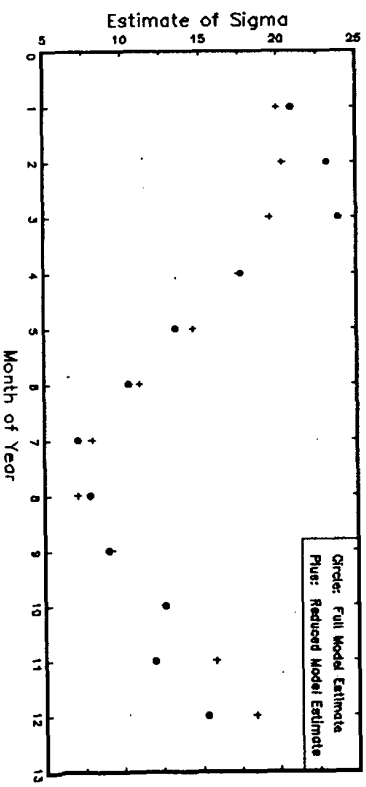


Figure 6: Estimates of Sigma

Figure 7: Diagonal Exceedance Percentage Plot of the Full Model's Residuals

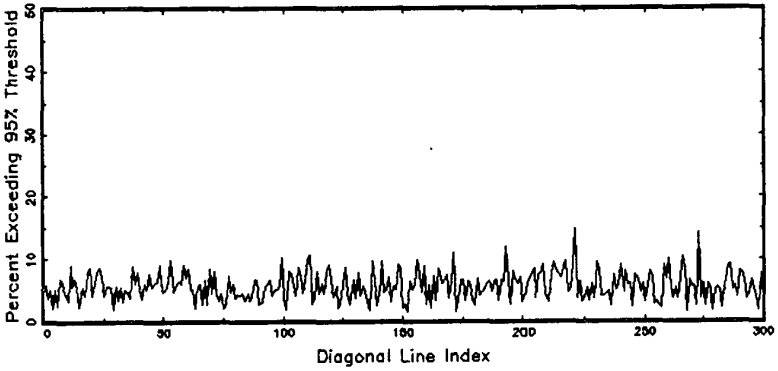


Figure 8: Main Diagonal Periodogram of the Full Model's Residuals

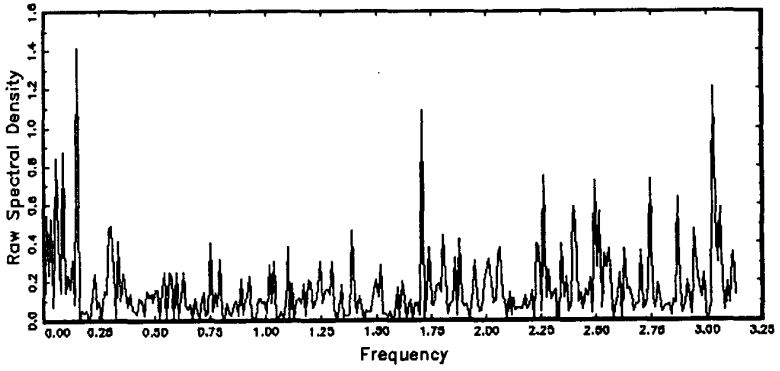


Figure 9: Diagonal Exceedance Percentage Plot of the Reduced Model's Residuals

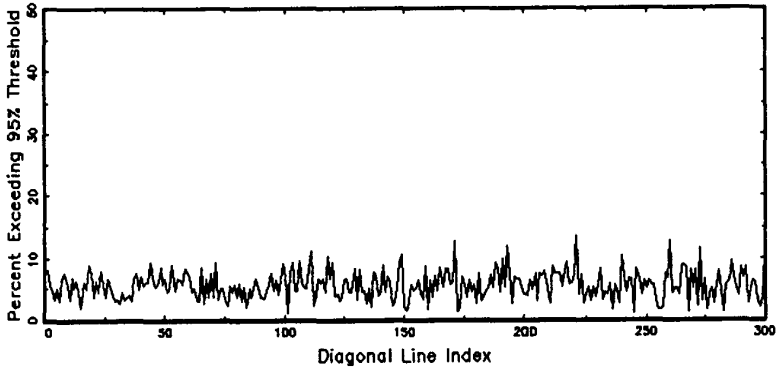


Figure 10: Main Diagonal Periodogram of the Reduced Model's Residuals

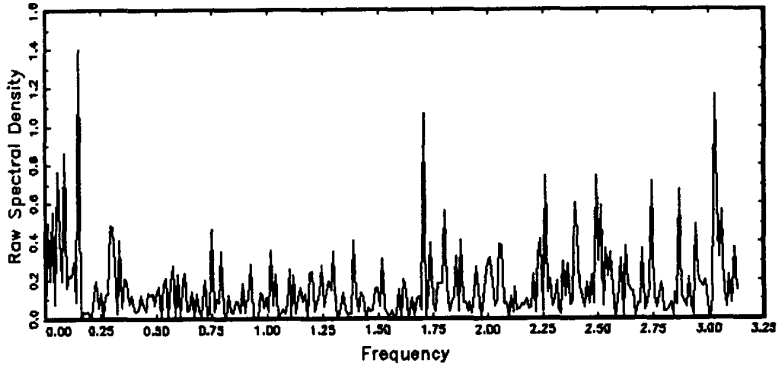


Figure 11: Diagonal Exceedance Percentage Plot of the Combined Model's Residuals

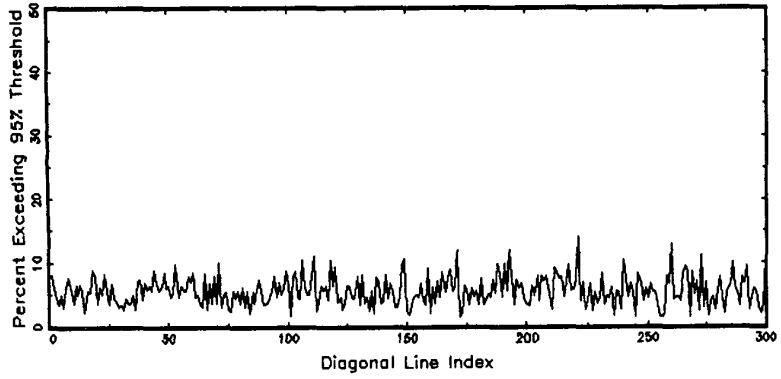
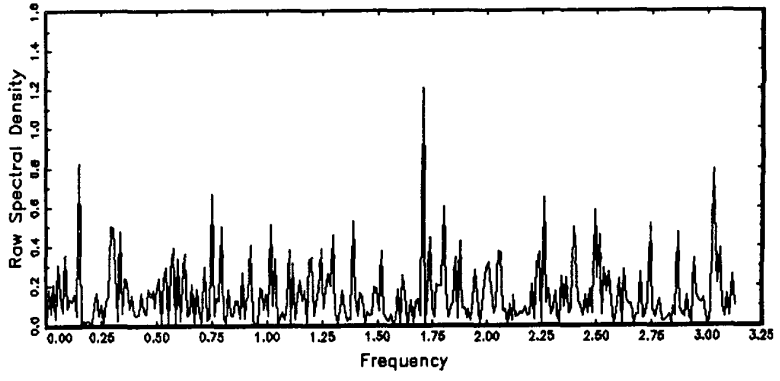


Figure 12: Main Diagonal Periodogram of the Combined Model's Residuals



EMPIRICAL ESTIMATE OF THE RESPONSE OF GLOBAL SURFACE TEMPERATURES (1867-1988) TO DIFFERENT CLIMATE FORCINGS

A.J. Coops, Institute for Marine and Atmospheric Research
State University of Utrecht, The Netherlands

1. INTRODUCTION

Climate variations have internal and external causes, which operate simultaneously. This fact seems problematical, when one would like to assign a cause-effect relationship to observed variability (the so-called "attribution"-problem; it is one of the problems in detecting the enhanced greenhouse effect). Fortunately, different climate factors are operating at different time-scales. This phenomenon is utilized here to distinguish between the forcings of different factors.

El Niño and the Southern Oscillation are important examples of climate fluctuations at a short time-scale (in the order of one year). The influence of ENSO on yearly and globally averaged surface temperatures has been estimated and eliminated by Jones (1988).

The sunspot cycle is operating at an intermediate time-scale (in the order of one decade). During sunspot maxima the solar radiation is a little bit more intense than during minima. Although the influence of a volcanic eruption upon climate is restricted to a few years at most, eruptions frequently occur in clusters, whose influence can extend over the same decadal time-scale. A method of estimation and elimination of the temperature response to the forcings of these external factors is described and the residual series is interpreted: air pollution is considered to be a forcing at a long time-scale.

2. SOLAR RADIATION AND VOLCANISM

2a. Solar radiation

Variations in solar activity most probably are reflected by fluctuations in the sun's radiation intensity on earth. High activity is accompanied by many flares and spots, places on the surface of the sun with relatively high and low temperatures, respectively. The positive effect of extra flares on the solar intensity overcompensates the negative effect of extra spots (Foukal and Lean 1990). A continuous record of yearly means of sunspot numbers from the year 1700 is available (Waldmeier 1961) and has been updated until 1990.

2b. Volcanism

The influence of a volcanic eruption upon climate is dependent upon many factors like place, time, intensity and composition of the eruption. Ultimately, the negative effect of an eruption on surface temperature is mainly determined by the amount, the spreading and the residence time of sulphuric aerosols in the stratosphere. The amount is dependent upon the altitude, the intensity and the composition of the eruption. The spreading is dependent upon the stratospheric circulation at the latitude of the eruption, which is a function of the time of the year. The residence time is restricted to some years.

The Volcanic Explosivity Index (VEI) of Simkin et al. (1981) has been discussed by Newhall and Self (1982). Their chronological list of largest explosive eruptions since 1500 is based on purely geological data and is confined to eruptions, whose contents are assumed to have reached the stratosphere. Since the time span between many successive eruptions is less than a few years, effects at a longer time-scale may be expected, which may be more pronounced than effects from singular events.

2c. Estimation and elimination of solar and volcanic effects

Since both forcings are operating at a comparable time-scale, their effects are quantified and removed alternately and iteratively. At first, the solar influence is estimated. Namely, the 11-year sunspot cycle is less irregular than the occurrence of volcanic eruptions, which looks like a Poisson process (De la Cruz - Reyna 1991). A first guess of the solar effect is eliminated from the original series. Next, the volcanic influence is investigated by using this series (after removal of the first estimate of the solar effect). This is the second part of the first iteration step. The second step consists of re-estimating the solar effect (after the first guess of the volcanic influence has been eliminated), followed by a second estimate of the volcanic effect. This iteration process is stopped, when two successive steps lead to the same results.

A two-stage temperature model for each effect is considered: low and high. The construction of merely a dichotomy (bifurcation; two levels) is justified by the "proxiness" of the forcing functions in question. The trans-

ition from one stage to the other is determined according to the method of Pettitt (1979): each change-point successively is chosen in such a way that the difference between the two levels in the adjacent subperiods is at its maximum in a statistical sense.

The number of change-points is determined by the number of sunspot cycles and the number of volcanic clusters, respectively. During each solar cycle two transitions are imposed by the model: an upward jump during the rise and a downward jump during the drop of the cycle. Hence, the dates of extremes of the solar cycle are the boundary conditions for the choice of the transition-points. In principle, it has been tried to split up the period between the occurrences of two successive volcanic eruptions into two subperiods, in such a way that an upward temperature jump arises. This can be done only, if the distance (in time) between these occurrences amounts to 4 years at least. (Namely, yearly values are used and it is required that each part of the subdivision of a period consists of two elements at least in order to reduce the possible effects of occasional extremes of singletons.) So, in most cases it is necessary to combine different eruptions into groups. A downward jump is imposed in the year of the first eruption of a cluster or in the next year.

During every iteration step the position of each transition successively is optimized (for both factors separately), until no longer shifts in change-points appear. Then the temperature values in each "affected" state are adjusted with the temperature difference between the average level during this state and the time-weighted average during the adjacent "unaffected" states.

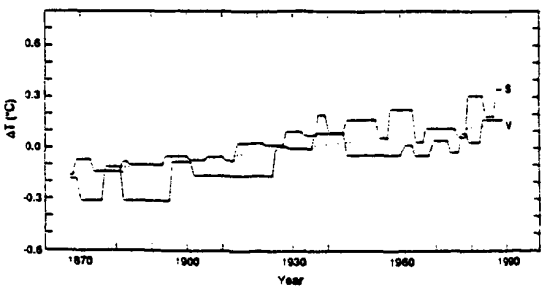


Fig. 1. End solution of the iteration process. Pattern of solar (S) and volcanic (V) effects (ΔT; values relative to the 1867-1988 temperature average). The straight line represents a linear fit to the residual series (after elimination of both effects).

2.d. Results

In Figure 1 the end solution of the iteration process is represented. The influence of the solar cycles after the Second World War (WW II) is stronger than that of those before. In general, the opposite holds for the volcanic effect. A net rise of 0.09 °C may be attributed to these two effects concerned here (see Table 1). This is only a small fraction of the rise of about 0.5 °C, as observed in the long-term component of global surface temperatures (Jones 1988)! The average temperature differences between the two states are also indicated in Table 1, for both factors, as is the average duration of both effects. The period of rise in the sunspot cycle has an average duration of about 4 years, the period of fall lasts around 7 years. The sun-affected temperature state starts ±1½ years before the solar maximum is reached and ends ±4½ years after this extreme, on average. So, the temperature response to the solar cycle has a time-lag of order one year. The volcanism-affected temperature state starts a few months after the first eruption of each group and ends some 2 or 2½ years after the last one, on average.

Table 1.a. Average solar (S) and volcanic (V) effects in °C.
b. Average difference between the two levels, for each effect, in °C
c. Average duration of the "affected" states in years (percentage of time).

Forcing factor	a.Effect:	Before WWII	After WWII	Increase
S	+ 0.045	+ 0.02	+ 0.085	+ 0.065
V	- 0.085	- 0.095	- 0.07	+ 0.025
S+V	- 0.04	- 0.075	+ 0.015	+ 0.09
	b. Difference:			
S	0.08	0.05	0.12	+ 0.07
V	0.13	0.14	0.11	- 0.03
	c. Duration:			
S	6 (55)	5 (45)	7.5 (70)	
V	10 (65)	13 (65)	7 (65)	

The average temperature deviation during each affected state is multiplied by its duration in order to get a cumulative temperature departure. These cumulative deviations are related to the yearly means of sunspot numbers (cumulated over each of the 11 cycles) and to the Smithsonian Volcanic activity Indices (SVI, cumulated over each of the 8 clusters), respectively. The SVI is an exponential transformation of the VEI and is proportional to the total amount of volcanic material erupted into the atmosphere. The significance of the relations is tested by the rank correlation statistic of Mann-Kendall.

Both relations are significant at the 5% level. The relationships are shown in Figures 2a and b. The relation with the cumulative sunspot numbers is linear, with the cumulative SVI log-linear.

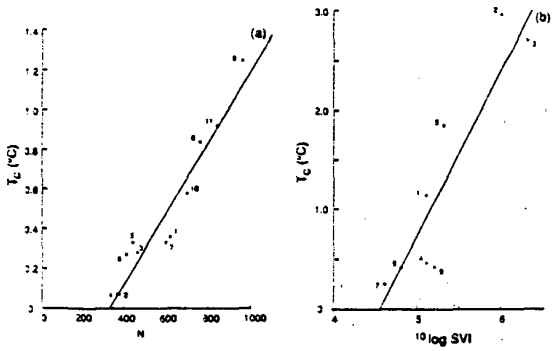


Fig. 2. Relationships between cumulative temperature deviations (T_c) and
(a) yearly means N of sunspot numbers (cumulated over the cycles) and
(b) volcanic indices SVI (cumulated over the clusters).

3. AIR POLLUTION

3.a. Introduction

The long time-scale component of the globally averaged surface temperatures should lie more or less between the two step functions in Figure 1. Fitting a straight line to the residual series (after elimination of the solar and volcanic effects) leads to a significant rise of 0.3 °C (see Figure 1). The 95% confidence limits for the regression line do not deviate more than 0.02 °C from this line. Hence, it is clear from Figure 1 that the long-term trend is not monotonic. A relative maximum is present during WW II. On the basis of a steady increase in the concentration of atmospheric greenhouse gases, a gradual rise in temperature might be expected. Hence, the presence of at least one other climate factor is required in order to explain the residual temperature behaviour.

Tropospheric aerosols are acting negatively upon surface temperatures in two ways: the net radiation to the surface is reduced directly by the particles itselves and indirectly by cloud formation, since the particles also behave like cloud condensation nuclei.

The influence of man-made aerosol particles (liquid and solid) is to a large extent determined by gaseous sulphur emissions; these emissions have caused a large increase in the concentration of aerosol sulphate in the northern hemisphere, due to fossil fuel combustion (Watson et al. 1990). Century-long histories of the emissions of sulphur (Möller 1984) and fossil fuel carbon (CDIAC 1990) show an augmentation of the rate of release after WW II. So, it is well conceivable that the negative effect of anthropogenic aerosols has become perceivable just after WW II. Moreover, the anthropogenic contribution to the total atmospheric aerosol content has proportionately extra increased because of the reducing volcanic dust amount. Hence, the residual temperature series is split up into two parts: before and after the end of WW II. It is assumed that the noticeable effect of man-made aerosol is restricted to the second subseries.

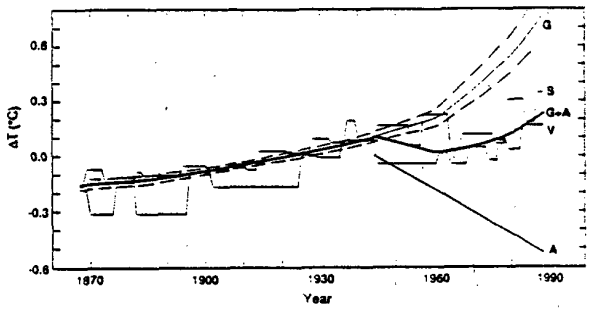


Fig. 3. As Figure 1, with the straight line replaced by an exponential curve representing the enhanced greenhouse effect (G), a linear fit representing the anthropogenic aerosol effect (A) and a line representing the combined effect (G+A) (95% confidence limits are also indicated).

3.b. The enhanced greenhouse effect

The first part of the residual series has been fitted by an exponential function by means of least squares estimation. The coefficient in the exponent expresses the rate of increase in radiative forcing and has been chosen equal to the one, used by Oerlemans (1988). In this way a linear regression problem is obtained. According to this fit (see Figure 3) the rise in temperature during the period 1867-1945 amounts to 0.28 °C. Extrapolating the first fit to the right, one would find a rise of 0.65 °C during the period 1945-1988, when the constant in the exponent has been doubled after 1960, because of the release of extra trace gases after that date. Consequently, in the absence of an anthropogenic aerosol influence one might have expected a rise of more than 0.9 °C. More than half of this temperature increase should have taken place after 1960. (The 95% confidence limits are also indicated in Figure 3.)

3.c. The anthropogenic aerosol effect

Subtracting the extrapolation of the first curve from the second part of the residual series, one finds a difference that is described adequately by a straight line (labelled A in Figure 3). This negative effect on temperature is considerable: in 1988 it amounts to more than 0.5 °C and cancels some 80 or 85% of the enhanced greenhouse effect since WW II. According to the picture in Figure 3 the enhanced greenhouse effect is dominating over the anthropogenic aerosol effect since 1960. Wigley (1989) implicitly assumes the relation between the radiative forcing (temperature) and the aerosol concentration in the troposphere to be logarithmic. Hence, a linear temperature effect points to an exponential growth in the concentration of particles; combined with a linear increase in aerosol emissions, this would be indicative of an accumulation effect in the troposphere.

Table 2. Percentages of total variance explained by different (combinations of) forcing factors.

Forcing factor(s):	ENSO	S	V	S+V	G+A
Percentage:	11	25	31	55	24

4. SUMMARY AND CONCLUSIONS

An analysis of global, yearly surface temperatures since 1867 has been presented. An attempt has been made to incorporate relevant factors influencing climate at three different time-scales. ENSO-phenomena have been mentioned as features at the short time-scale, whose influences were already estimated and eliminated from the series under consideration. The effect of volcanism has been studied by taking groups of eruptions into account. These clusters operate at the same time-scale as the sunspot cycle. The influences of these climate factors have been estimated alternately and iteratively. In this estimation procedure a two-state model for temperature (high-low) has been applied. The states have been connected to high-low states of sunspot numbers and high-low periods of volcanic indices, respectively. Although these two factors at the intermediate time-scale explain only a small part of the observed rise in temperature, they explain more than 50% of the total variance (see Table 2) and the relations with temperature are statistically significant. The residual temperature curve does show an upward trend, which most probably has to be attributed to the enhanced greenhouse effect. However, its presence can only be made acceptable by introducing the likely influence of man-made aerosol. Its negative temperature effect has been linked with fossil fuel consumption. The aerosol influence should have reduced the enhanced greenhouse effect by more than 50%. By this analysis 90% of the original variance has been explained.

References

Carbon Dioxide Information Analysis Center, 1990: Trends '90. A compendium of data on global change. Oak Ridge National Laboratory, 257 pp

De la Cruz-Reyna, S., 1991: Poisson-distributed patterns of explosive eruptive activity. Bull. Volcanol., 54, 57-67

Foukal, P. and J. Lean, 1990: An empirical model of total solar irradiance variations between 1874 and 1988. Science, 247, 556-558

Jones, P.D., 1988: The influence of ENSO on global temperatures. Clim. Mon., 17 (3), 80-89

Möller, D., 1984: Estimation of the global man-made sulphur emission. Atmos. Envir., 18, 1, 19-27

Newhall, C.G. and S. Self, 1982: The Volcanic Explosivity Index (VEI): an estimate of explosive magnitude for historical volcanism. J. Geophys. Res., 87 (C2), 1231-1238

Oerlemans, J., 1988: Simulation of historic glacier variations with a simple climate-glacier model. J. Glaciol., 34 (118), 333-341

Pettitt, A.N., 1979: A non-parametric approach to the change-point problem. Appl. Statist., 28, 126-135

Simkin, T., L. Siebert, L. Mc. Clelland, D. Bridge, C. Newhall and J.H. Latter, 1981: Volcanoes of the world. Academic Press, New York (Dowden, Stroudsburg, Pennsylvania, Hutchinson & Ross), 233 pp

Waldmeier, M., 1961: The sunspot-activity in the years 1610-1960. Zürich; Schulthess and Co. Ag., 171 pp

Watson, R., H. Rodhe, H. Oeschger and U. Siegenthaler, 1990: Greenhouse gases and aerosols. In Houghton, J.T., G.J. Jenkins and J.J. Ephraums, editors: Climate Change. The IPCC Scientific Assessment. WMO/UNEP, Cambridge University Press, 365 pp

Wigley, T.M.L., 1989: Possible climate change due to SO_2-derived cloud condensation nuclei. Nature, 339, 365-367

Hypothetical Explanation of Natural and Anthropogenic Temperature Variations with a Nonlinear Multiple Regression Model

DIETER BAYER AND CHRISTIAN-DIETRICH SCHÖNWIESE

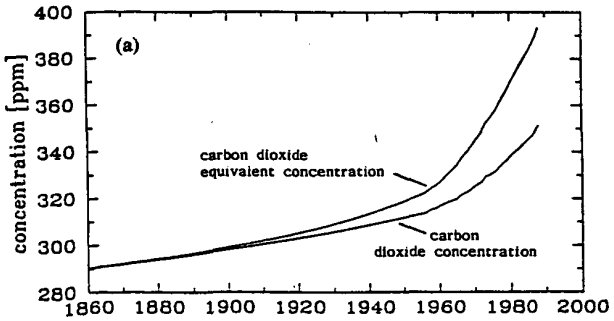
J.W. Goethe-University, Institute for Meteorology and Geophysics
Praunheimer Landstr. 70, D-6000 Frankfurt am Main 90

1. Introduction

It is a fact that increasing atmospheric greenhouse gas concentrations have considerable effects on the level of global surface air temperature. The resulting change of atmospheric circulation may affect the regional and seasonal pattern of temperature in an even more extensive way. Since any climatic change of global extent is of essential interest for mankind, the elaboration of climate models - deterministic as well as statistical - is one of the most urgent research subjects at the moment. In this paper results of a statistical model are presented, which simultaneously correlates observed climatic data time series with natural and anthropogenic forcing parameters. In this way, an assessment is enabled by which a hypothetical separation and quantification of natural and anthropogenic signals in climate becomes possible.

2. Model

The multiple regression model implemented at Frankfurt University (Schönwiese 1991, Schönwiese and Runge 1991, Schönwiese and Stähler 1991) considers natural processes such as volcanism, solar effects, and the El Niño mechanism; moreover, the anthropogenic carbon dioxide equivalent concentration increase in the atmosphere is taken into account. For all these influence parameters alternative data time series are available; a selection is shown in Figure 1 (Cress and Schönwiese 1992, Schönwiese et al. 1990,



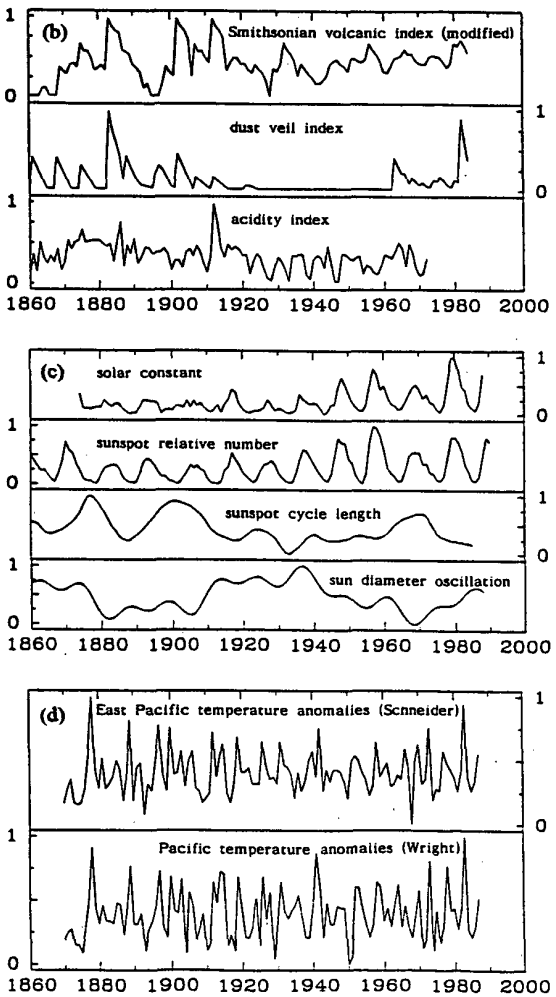


FIG. 1. Alternative time series used in the multiple regression model: (a) trace gas concentrations, (b) volcanism, (c) solar activity, and (d) El Niño mechanism. Details of forcing parameters see elsewhere (Schönwiese et al. 1990, Cress and Schönwiese 1992, Schönwiese et al. 1992).

Schönwiese et al. 1992). The computations concern both the global and hemispheric averages of the surface air temperature as well as its spatial and seasonal peculiarities.

In Figure 2a a global simulation with a nonlinear (logarithmic) trace gas concentration-temperature relation is shown. In this "best fit"-regression the multiple correlation coefficient is equal to 0.84, explaining a substantial part of the total observed temperature variance in the time interval considered. Figures 2b, 2c, 2d, and 2e decompose this simulation into time series for each of the four forcing parameters under consideration. As expected the observed increase of global temperature is driven by the equivalent carbon dioxide forcing while volcano, solar, and El Niño activities cause more or less pronounced deviations from this trend. It is worth noting that this trend result is common to all combinations of forcing parameter time series compiled in Figure 1. The attribution of the short-term temperature fluctuations, however,

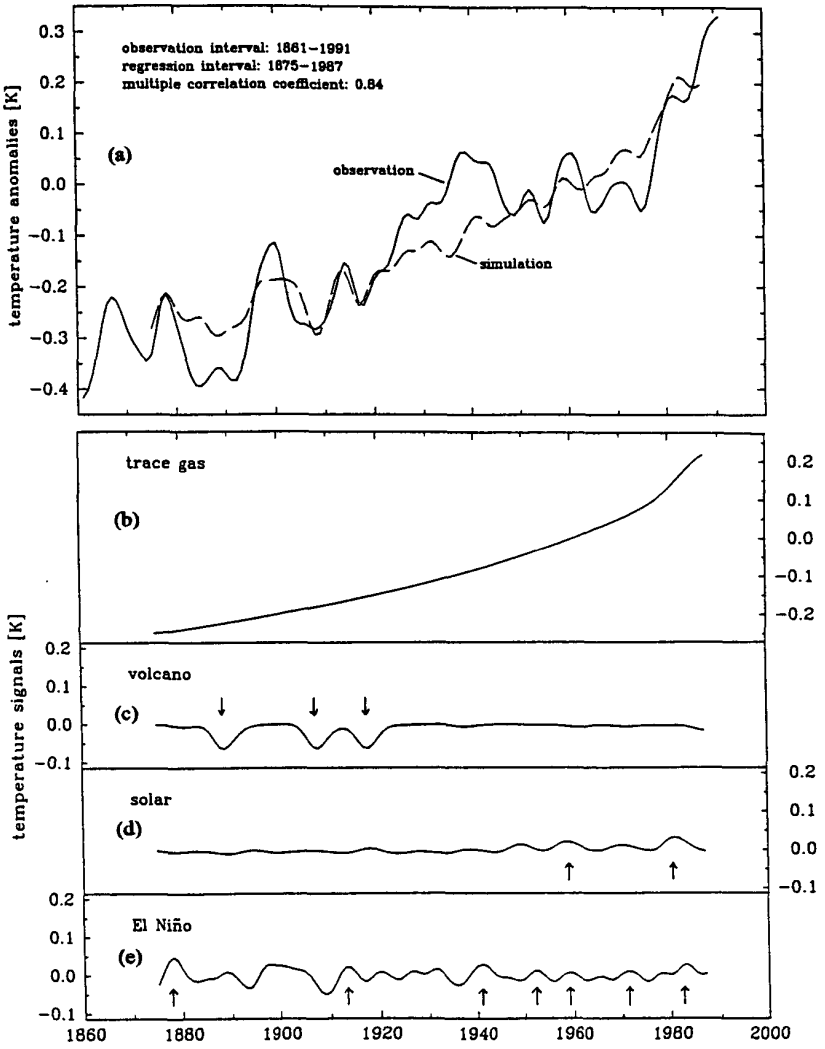


FIG. 2. (a) Plots of the observed (solid line; Jones 1988) and statistically reproduced (dashed line) time series of the global and annual averaged surface air temperature. For this simulation the forcing parameter time series of carbon dioxide equivalent concentration, modified Smithsonian volcanic index, solar constant, and winter Pacific temperature anomalies are used. For detailed description of all parameter time series see Schönwiese et al. 1990). In the lower panel the variations included in the reproduction are related to the four forcing parameters: (b) trace gas concentration (values without volcanic as well as with averaged solar and El Niño activity), (c) volcanic activity (values referring to times without volcanic activity), (d) solar activity (values referring to averaged activity), and (e) El Niño activity (values as well referring to averaged activity). The arrows indicate strong volcanic eruptions or pronounced solar and El Niño events.

depends to a high degree on the chosen combination of natural forcing parameters. The regression shown in Figure 2 is based to a time shift of twenty years in the trace gas series, of five years in the volcano series, and of one year in the solar series while the El Niño series enters unshifted into the calculations.

Reductions of the time interval used for the computations of the regression coefficients show that the relevant length of each forcing parameter time series should not fall below a certain value. Otherwise extrapolations such as shown in Figure 3 may lead to considerable misinterpretations.

3. Scenarios

Keeping in mind these limitations of the multiple regression model, scenarios of future climate trends can be carried out. Assuming a linear trace gas-temperature relation, the result of a trace gas doubling scenario is a 3 K increase of the global mean surface air temperature. When applying a logarithmic relation the increment reduces to 2.5 K, which is exactly the best estimate of the IPCC (International Panel on Climate Change; Houghton et al. 1990) using deterministic models (Schönwiese 1991, Schönwiese

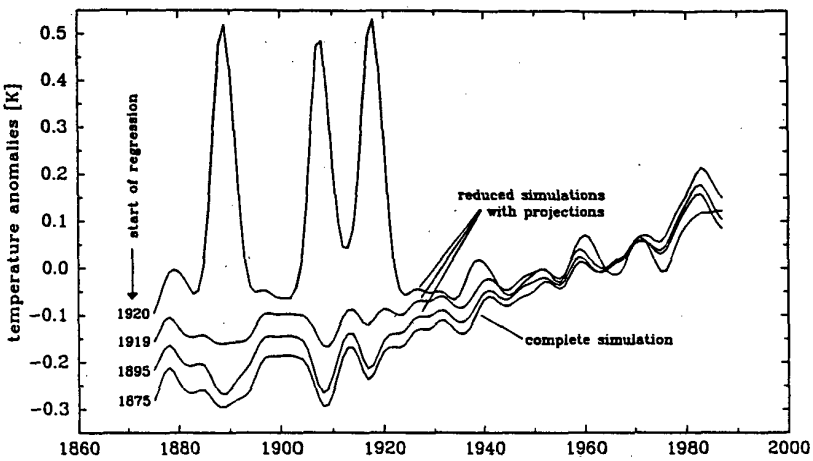


FIG. 3. Complete (1875-1987) as well as reduced regression simulations (each up to 1987) with extrapolation of the regression equation back to 1875. Regression time intervals beginning after 1919 contain only weak volcanic eruption. This may lead to some anticorrelations.

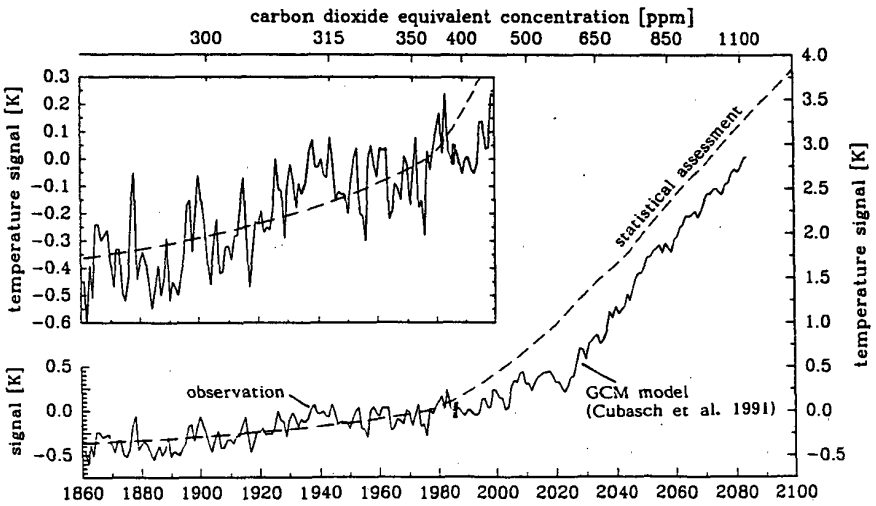


FIG. 4. Observed global mean annual surface air temperature variations 1861-1985 (Jones 1988), transient IPCC scenario A (Houghton et al. 1991) general circulation model simulation 1985-2085 (solid line; Cubasch et al. 1991), and statistical assessment of the greenhouse temperature signal with logarithmic trace gas-temperature relation (dashed line); upper plot 1860-2000 ordinate scale amplified. All temperature signals refer to the 1985 value.

1992). It also corresponds to the MPI (Max Planck Institute, Hamburg, FRG; Cubasch et al. 1991) transient climate model (GCM) results shown in Figure 4. This MPI (GCM) simulation may underestimate the climate response by approximately +0.5 K due to a systematical error (Hasselmann et al. 1992).

In the regional and seasonal domain the largest greenhouse gas-induced temperature signals can be detected in the winter of the high northern latitudes. Equivalent carbon dioxide doubling scenarios using logarithmic trace gas-temperature relations lead to a temperature increase of 10 K and more in the arctic winter while in the tropics a signal of less than 2 K is comparatively low (Figure 5). With the same distributions of the logarithmic model linear computations would clearly lead to higher signals. Because of the poor data density in the southern hemisphere no reliable results can be obtained for the area south of 60° S. Nevertheless, also in the southern winter a considerable temperature increase is indicated by the statistical calculations.

Particularly with regard to the spatial and seasonal pattern, all these assessments are in good agreement with results of some deterministic models (Schönwiese and Stähler 1991).

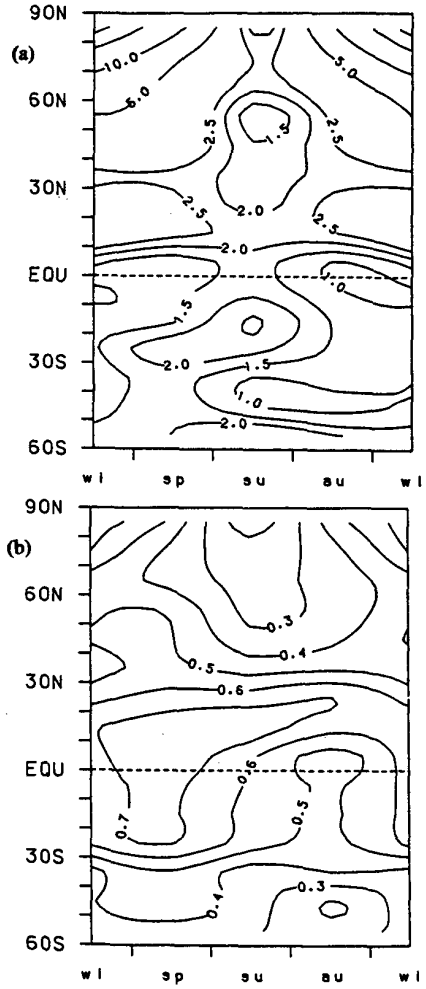


FIG. 5. Seasonal and spatial resolution of (a) trace gas doubling temperature signal, and (b) correlation coefficient of the multiple regression (1880-1985) using a logarithmic trace gas-temperature relation (underlying temperature data from Hansen and Lebedeff 1987).

ACKNOWLEDGEMENT

This study was supported by the German Climate Research Programme (BMFT, Project No. 07 KFT 16/5) which is gratefully acknowledged.

REFERENCES

Cress, A., and C.-D. Schönwiese, 1992: Statistical signal and signal-to-noise assessments of the seasonal and regional patterns of global volcanism-temperature relationships. *Atmósfera*, **5**, 31-46.

Cubasch, A., K. Hasselmann, H. Höck, E. Maier-Reimer, U. Mikolaje-wicz, B.D. Santer, and R. Sausen, 1991: *Time-dependent greenhouse warming computations with a coupled ocean-atmosphere model.* Max-Planck-Institut für Meteorologie Hamburg, Report No. 67.

Hansen, J., and S. Lebedeff, 1987: Global trends of measured surface temperature. *J. Geophys. Res.*, **92**, 13345-13372.

Hasselmann, K., R. Sausen, E. Maier-Reimer, and R. Voß, 1992: *Das Kaltstartproblem bei Klimasimulationen mit gekoppelten Atmosphäre-Ozean-Modellen.* Paper presented at the Deutsche Meteorologen-Tagung 1992, Berlin (FRG).

Houghton, J.T., G.J. Jenkins, and J.J. Ephraums (eds.), 1990: *Climate change. The IPCC scientific assessment.* Cambridge University Press.

Jones, P.D., 1988: Hemispheric surface temperature variations: recent trends and an update to 1987. *J. Climate*, **1**, 654-660.

Schlesinger, M.E. (ed.), 1991: Greenhouse-gas-induced climatic change: A critical appraisal of simulations and observations. *Developments in Atmospheric Science*, **19**, 483-503.

Schönwiese, C.-D., W. Birrong, U. Schneider, U. Stähler, and R. Ullrich, 1990: *Statistische Analyse des Zusammenhangs säkularer Klimaschwankungen mit externen Einflußgrößen und Zirkulationsparametern unter besonderer Berücksichtigung des Treibhausproblems.* Inst. Meteorol. Geophys. Univ. Frankfurt (FRG), Report No. 84.

—, 1991: Multivariate statistical assessments of greenhouse-gas-induced climatic change and comparison with results from general circulation models. In: *Schlesinger 1991.*

—, and K. Runge, 1991: Some updated statistical assessments of the surface temperature response to increased greenhouse gases. *Internat. J. Climatol.*, **11**, 237-250.

—, and U. Stähler, 1991: Multiforced statistical assessments of greenhouse-gas-induced surface air temperature change 1890-1985. *Clim. Dyn.*, **6**, 23-33.

—, 1992: *The greenhouse hypothesis: Model projections in comparison with observational statistics.* Paper to be presented at the 5th International Meeting on Statistical Climatology, Toronto.

—, R. Ullrich, and F. Beck, 1992: *Solare Einflüsse auf die Lufttemperaturvariationen der Erde in den letzten Jahrhunderten.* Inst. Meteorol. Geophys. Univ. Frankfurt (FRG), Report No. 92.

Breaking Recent Global Temperature Records

Gilbert W. Bassett Jr.
Department of Economics(m/c144)
University of Illinois at Chicago
Box 4348, Chicago, Il 60680,
E-mail:u09006@uicvm.cc.uic.edu

Prepared for the
5th International Meeting on Statistical Climatology,
Toronto June 22, 1992

ABSTRACT

Global surface temperature was a record in 1988. What is the probability that this record would be surpassed within three years? Answers are provided under a variety of simple statistical models. The answers illustrate how record breaking is influenced by alternative model specifications. Estimates for the probability of a record are shown to range widely. If annual temperature is independent and identically distributed then a new record is unlikely. But probabilities increase rapidly when there is a trend or autocorrelation. Estimates of the probability of a record using data on global temperature suggest that new records in the next few years would not be rare events. (A longer version of this paper is forthcoming in *Climatic Change*).

I. Introduction

What is the probability of setting global temperature records? This question was originally motivated by the "Hansen bet" in which,

"Climate expert James Hansen...told a group of climatologists last week that his confidence that the greenhouse effect has arrived is even higher than it was in 1988, when he testified before Congress that he believed the global warming of recent decades was driven by gases produced by human activity. So sure is he now of this conclusion that he said he'd bet even money that one of the next 3 years will be the hottest in 100 years. ...People aren't going to believe such an "incredible" and scientifically outrageous prediction, Hansen said"....Science(1990b),

Is the prediction really outrageous? Should the Hansen bet have been accepted?

The record-breaking question is considered under alternative models for temperature. Annual global temperature is taken to be the realization of a random variable whose properties will be described by a model. Attention is restricted to models in which temperature depends only on statistical parameters; there is no reference to causal mechanisms that affect global temperature. Since the models do not refer to factors that affect global climate (sun-spots, volcanic activity, atmospheric concentrations of greenhouse gases, etc.) they cannot be used to infer the impact of human activities on global climate; see Wigley et al.(1985) and Solow and Broadus(1989) for a discussion of statistical issues relating to global climate. Further, the models are all simple; they only have a few parameters. The objective is to see how the probability of recently set records varies under alternative model specifications.

II. Results

Global temperature values are denoted

by c(t) where t=1880,...,1991. The temperature values are taken to be realizations of random variables C(t); uppercase C denotes the temperature random variable with the actual realized temperature being denoted by the lower case c. The temperature values for t=1880,...,t',where t'=1988, are shown in Figure 1 along with the fitted trend. The

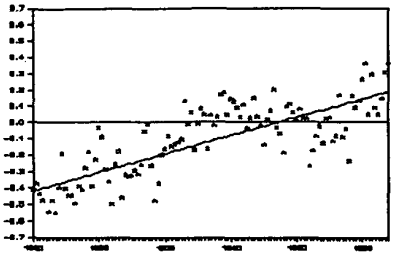


Figure 1
Global temperature 1880-1988

maximum temperature--denoted by r--occurred in the most recent year t' so that c(t')=r.

The probability of a record within three years is derived under the following specification for the temperature time series: C(t)=m(t)+e(t); t=1880,...,1991, where m(t) is the mean of C(t) and e(t) is a random variable that represents deviations from the mean. The mean is assumed to depend linearly on time, m(t)=a+bt, where a and b are parameters. The parameter b is the annual change in temperature trend in degrees C.

The random variables e(t) are assumed to be generated by the first-order autoregressive model, $e(t)=\rho e(t-1)+v(t)$, $-1<\rho<1$, where the ρ parameter is the correlation between successive e(t)'s, and the v(t)'s are independent with a common Gaussian distribution with mean zero and variance $\sigma^2(v)$; $v(t) \sim G(0,\sigma^2(v))$. The variance of C(t) is $\sigma^2(v)/[1-\rho^2]$.

Record breaking problems are often analyzed under the assumption of independent and identically distributed (iid) random variables; for an excellent review of the theory and applications of record breaking see Glick(1978). This corresponds to what will be called, Case I: m(t)=a, $\rho=0$. This says that temperature varies around a mean value that does not change over time, b=0, and that annual temperature is independent of past values. When the temperature data is used to estimate the parameters of this iid temperature model the probability of a record in the three years after 1988 is found to be only about .05.

The iid situation however need not be accurate. The time series of global temperature shows an upward drift in global temperatures. If this trend is a genuine feature of climate then we have, Case II: m(t)=a+bt; $\rho=0$. This says temperature increases as a linear function of time, but each year's temperature is independent of previous years. On making the calculations under the assumption of an upward trend the probability of a record is found to be about .37.

The iid assumption also could fail because independence is wrong. The simplest

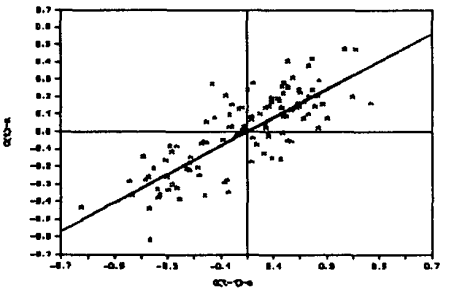


Figure 2
Current and Lagged Temperature

possibility in this regard is that successive annual temperatures are correlated. Figure 2 depicts global temperature plotted against its value in the previous year. The positive slope suggests dependence between successive annual temperatures. Serial correlation is also a plausible consequence of a variety of mechanisms that affect year-to-year changes in

Table 1
Parameter Estimates and Record Probabilities

	Case I	Case II	Case III	Case IV
$\hat{a}$	-.1135	-10.95	-.1135	-11.03
$\hat{b}$	0	.005603	0	.005647
$\hat{\rho}$	0	0	.808	.50
$\hat{\sigma}^2(v)$	$.227^2$	$.1429^2$	$.136^2$	$.124^2$
$\hat{e}(1988)$	-	-	.46	.10
Record Probability	.06	.37	.65	.75

global climate (the relation however is likely to be much more complex than the simple "one-period" serial correlation considered here). Simple reasons for serial correlation are thermal feedback between the ocean and atmosphere, sunspots, and volcanic activity. Oceans dissipate heat slowly so warm years, which heat the oceans, will likely be followed by warmer than average years; if sunspots affect climate then their correlations will begin to show up in the temperature series; and volcanic activity--though itself possibly independent--will induce correlations if volcanic ash stays in the atmosphere for long periods.

This leads to consideration of, Case III: $m(t)=a$; $-1<\rho<1$. This says there is no trend, $b=0$, but current temperature values are influenced by their values in previous years. With serial correlation the probability of a record depends not only on the current record value, but also on how long ago the record was set. When parameter values are estimated from the historical data it is shown that the probability of a record in the next three years is greater than ½. Serial correlation alone suffices to make a new record a likely event.

The combination of a trend and serial correlation is also considered. This is, Case IV: $m(t)=a+bt$; $-1<\rho<1$. When estimated values are substituted for parameters the probability of a record is also greater than ½.

Remarks: It should be emphasized that this analysis is not decisive about the question of whether increased concentrations of greenhouse gases in the atmosphere have already caused changes in global climate. Since the probability of a record is not be as small as might have been intuitively suspected and the probability can be quite high even if there is no upward temperature trend suggests caution in inferring climate change from a succession of record temperatures. But the existing temperature data leave room for views that are consistent with a wide range of beliefs regarding global warming.

Also, attention is restricted to simple models. More realistic univariate models would allow for time-varying parameters, higher order autoregressive specifications, moving average errors, first and higher order differences of C(t), and so on. Multivariate models would introduce explanatory variables, like atmospheric concentrations of CO_2, as possible determinants of m(t); for a recent example see Kuo et al.(1990) and also Barnett(1990). Conditional quantile models

also could be used to represent and estimate, not only the mean, but also the quantiles and extremes of the C(t) distribution; see Koenker and Bassett(1978) for discussion of quantile estimation and modeling.

III. Conclusion

The probabilities for a record under the alternative specifications are summarized in Table 1. It shows that a record goes from rare to likely, depending on the temperature model. It is rare when there is no trend and little correlation. It becomes more likely when there is an increasing temperature trend and when annual temperature is positively correlated with past values. Correlation matters because the recent record implies a deviation from trend that is positive and large. With positive correlation this makes future positive deviations more likely, thus increasing the probability of a record. When the probability of a record is estimated allowing for trend and serial correlation the Hansen bet becomes favorable for Hansen at the stated even odds. (After the original version of this paper was completed it was reported that a new temperature record was set in 1990; Science(1991)).

References

Barnett, T.P. Beware Greenhouse Confusion. *Nature* v.343, February 22, 743.

Glick, N.(1978). Breaking Records and Breaking Boards. *American Mathematical Monthly*, January, 2-26.

Hansen,J. and S. Lebedeff(1987). Global Trends of Measured Surface Air Temperature. *Journal of Geophysical Research*, v.92, n.d11, 13,345-13,372.

Koenker, R.W. and G.W.Bassett(1978). Regression Quantiles, *Econometrica*, Vol. 46, No. 1, January 1978, 33-50.

Kuo,C., C.Lindberg, and D.J.Thomson (1990). Coherence Established Between Atmospheric Carbon Dioxide and Global Temperature. *Nature* v.343, February 22, 709-714.

Science(1990a). February 2. p.521.

Science(1990b). May 4. p.549.

Science(1991). January 18. p.274.

Solow A. and J.M. Broadus (1989). On the Detection of Greenhouse Warming, *Climatic Change*, v. 15, 449-453.

Wigley, T.M.L., J.K. Angell, and P.D.Jones(1985). Analysis of the Temperature Record, in *Detecting The Climatic Effects of Increasing Carbon Dioxide*, edited by M.C.MacCracken and F.M.Luther, Carbon Dioxide Research Division, DOE/ER-0235, December, 57-90.

Finding Trends in Air Temperature for Northern North America

Thian Yew Gan#
Hydrometeorological Processes Division
Canada Climate Centre
National Hydrology Research Centre
Saskatoon, Saskatchewan

Now at Water Resources Division, INAC, Box 1500, Yellowknife, CANADA NT X1A 2R3

Abstract

Using univariate Kendall's test for monotonic trends from the monthly maximum temperature of the 1949-1989 period from some carefully selected stations of Canada and north-eastern United States, it is found that Western Canada for the past 40 years has experienced warming especially in January and March. However, Eastern Canada and north-eastern USA experienced virtually no warming but cooling, particularly in October.

In this study, as part of a combined effort to detecting climatic change, a robust, nonparametric, seasonal Kendall's test, based entirely on ranks, was applied for testing monotonic trend in temperature data of the 1949-1989 period for Canada and parts of north-eastern part of United States. It was first adapted by Hirsch et al. (1982) from the Mann-Kendall test for testing trends in seasonal data, in which the seasons are considered as independent variables. The Canadian stations chosen for this study were those tested homogeneous in a study for testing the homogeneity of monthly temperature series for Canada conducted by the Climate Change Detection Group of the Canadian Climate Centre (Gullet et al., 1991). This was to ensure that trends detected were caused by actual natural variations of climate and not man-induced changes often referred to as inhomogeneities by climatologists. The 40 American stations selected are among the 1036 climatological stations of the US National Oceanic and Atmospheric Administration that had been screened by Quinlan et al. (1987). In addition, stations selected were those located in towns with less than 10,000 population to avoid heat island effects.

Other than inhomogeneities, temperature data often have problems on missing values, seasonality and seasonal dependence. From past experience, filling in of missing values can be a laborious task. Hirsch et al. found that missing values and seasonality present no theoretical or computational problems to applying Kendall's test. Despite of Hirsch et at.'s claim that missing values have little effects on Kendall's test, only stations with few missing values were chosen to avoid ambiguities in the final results. Later, Hirsch and Slack (1984) made use of the work of Dietz and Killeen (1981) to extend this test to data that are serially correlated, i.e., correlation between successive seasons within the same year is allowed. The same test statistic as the independent case is used but the variance of the test statistic is now modified to account for the covariance of a multivariate vector of Mann-Kendall statistics.

The results based on 147 stations of monthly maximum temperature show that in the past 40 years there has been a distinct warming trend in Western Canada for January and March (Figure 4). In April, May and June warming trends detected are less extensive and are more concentrated in the Prairies. In contrast, for the last 40 years, there had been virtually no warming trend detected in the southern parts of Eastern Canada and north-eastern United States, including areas around the Great Lakes; instead, these regions experienced some obvious cooling in October and to a lesser degree in January and June. Apparently, warming and cooling trends are seasonal and regional, not global; and both phenomena can take place at the same locations during different seasons.

Other than finding whether there is trend or not in temperature data, the magnitude of trend was estimated by an estimator proposed by Sen (1968). Among the largest positive trend magnitudes estimated is the January data of Entrance, Alberta (Latitude 53°22'N, Longitude 117°42'W), at 0.254°C/year (Figure 1). From a Monte Carlo study on independent normally distributed variates of sample size of 40, using a two-sided significance level $\alpha = 0.025$, i.e., $2\alpha = 0.05$, it was found that the power of Kendall's test exceeds 0.9 when the trend slope per standard deviation value reaches 0.05 (Figure 2). Again, by the Monte Carlo approach, it was also found that spatial correlations between stations alter the distribution, but not the average, of the number of rejections of null hypothesis tests (Figure 3). All these additional information help to confirm the results of univariate Kendall's test on air temperature, that northern North America had experienced regional warming and cooling in the last 40 years.

References

1. Dietz, E. J. and T. J. Killeen, A nonparametric multivariate test for monotone trend with pharmaceutical applications, J. American Statistical Assoc., 76, 169-174, 1981.

2. Gullet, D., Vincent, L and Malone, L., Results of homogeneity testing of monthly temperature series - application of multiphase regression model with mathematical change points, Unpublished manuscript, Canadian Climate Centre Report 91-10, 1991.

3. Hirsch, R. M. and J. R. Slack, A nonparametric trend test for seasonal data with serial dependence, Water Resour. Res., 20(6), 727-732, 1984.

4. Hirsch, R. M., J. R. Slack, and R. A. Smith, Techniques of trend analysis for monthly water quality data, Water Resour. Res., 18(1), 107-121, 1982.

5. Quinlan, F. T., T. R. Karl, and C. N. Williams, United States historical climatology network (HCN) serial temperature and precipitation data, Carbon Dioxide Information Analysis Centre, Oak Ridge National Laboratory, June, 1987.

6. Sen, P. K., Estimates of the regression coefficient based on Kendall's tau, J. Am. Statis. Assoc., 63, 1379-1389, 1968.

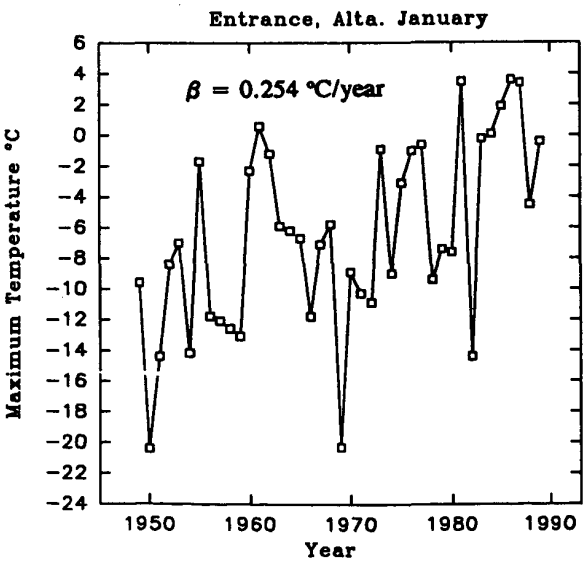


Figure 1 Selected temperature time series with large estimated warming trend.

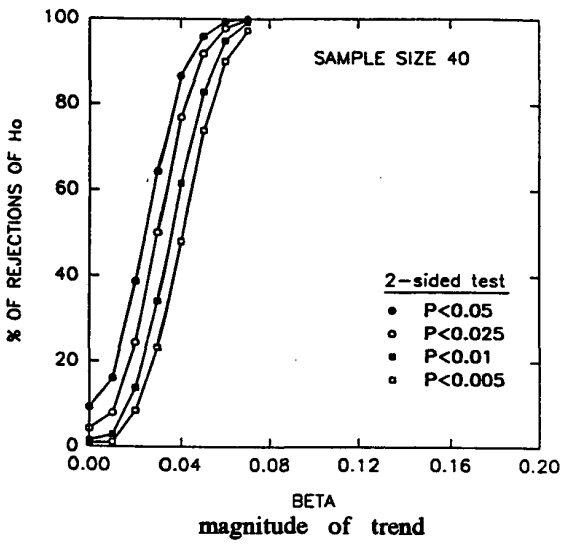


Figure 2 Empirical power curves of univariate Kendall's test.

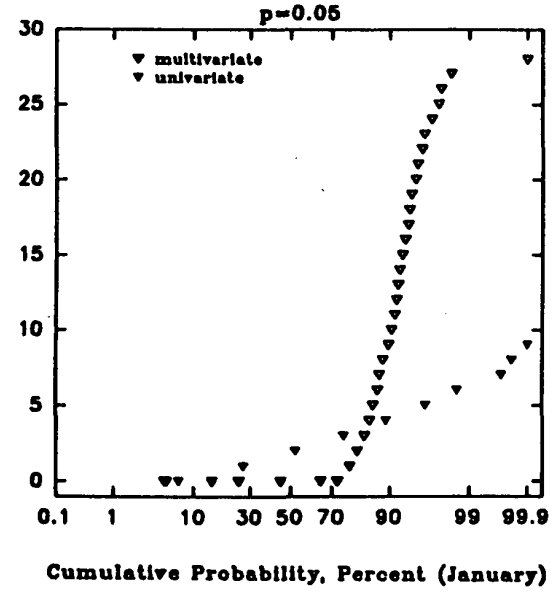


Figure 3 Comparing empirical probability curves of significant trends (rejections of H_o) for the independent and spatially-correlated temperature series of 28 Alberta & Saskatchewan stations.

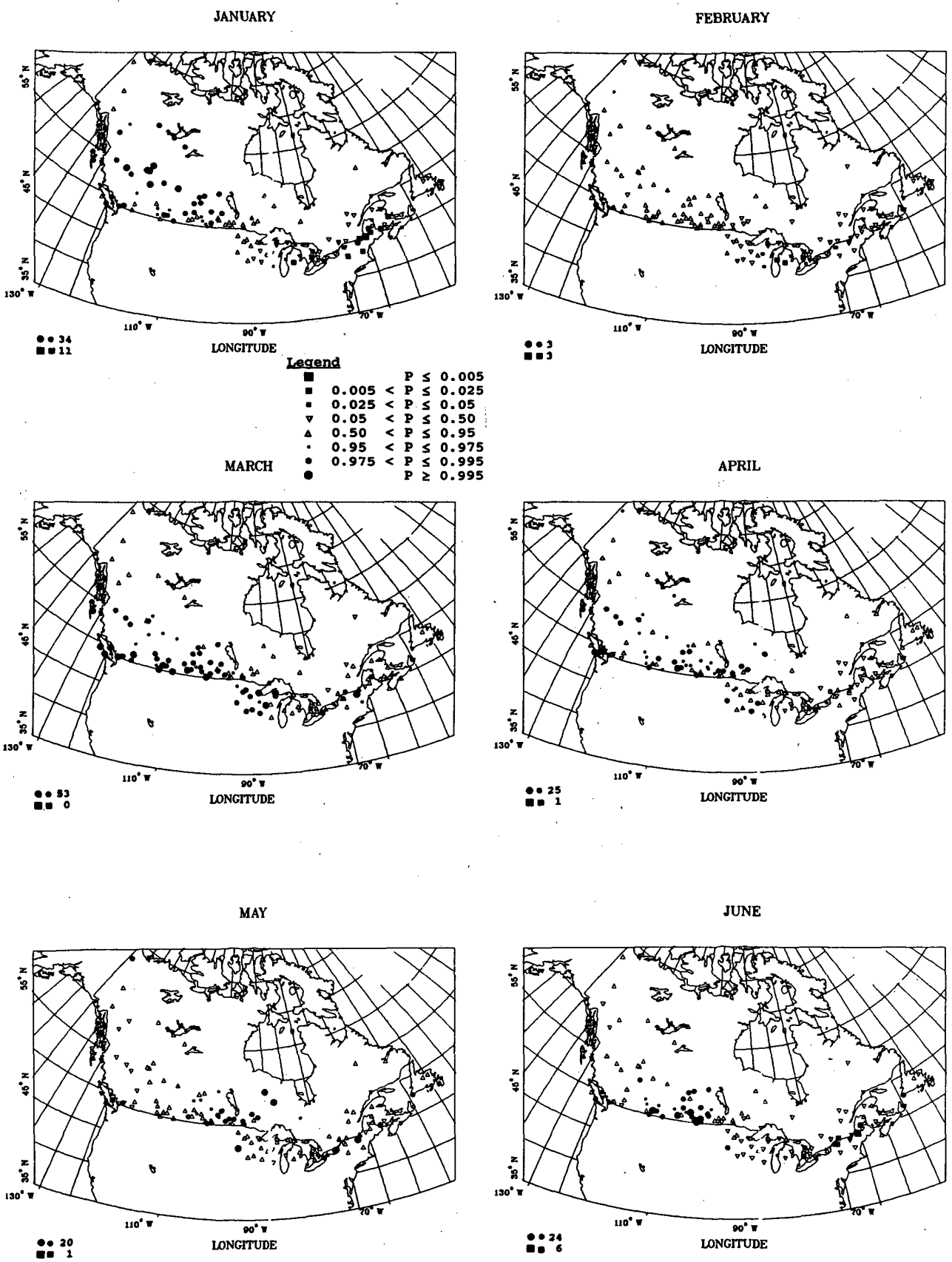
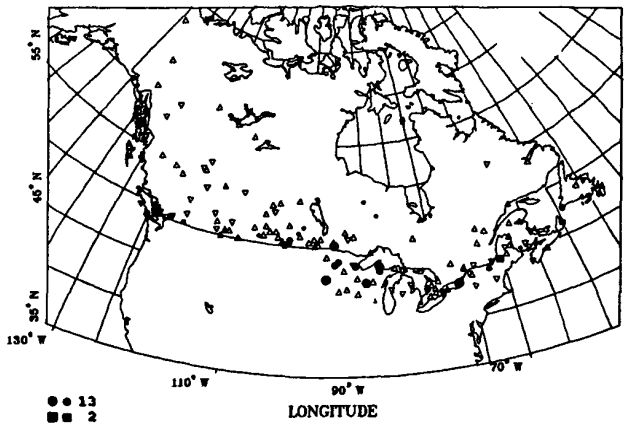
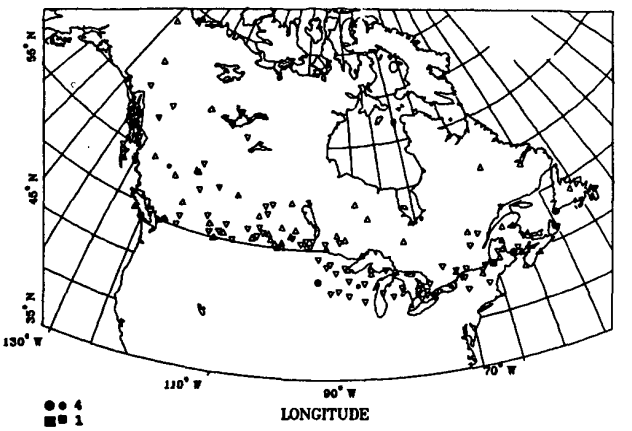


Figure 4 Distributions of maximum temperature trends based on univariate Kendall's tests by month.

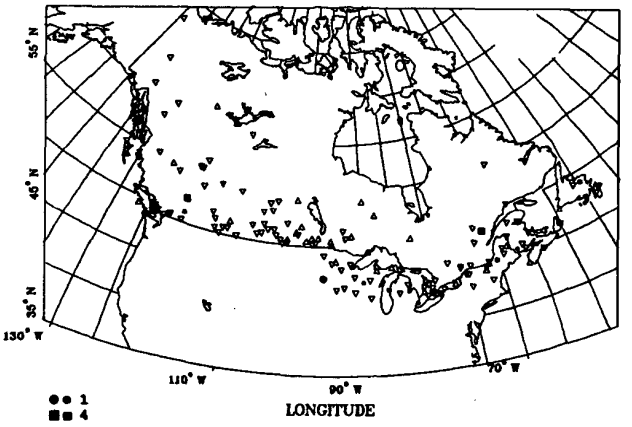
JULY



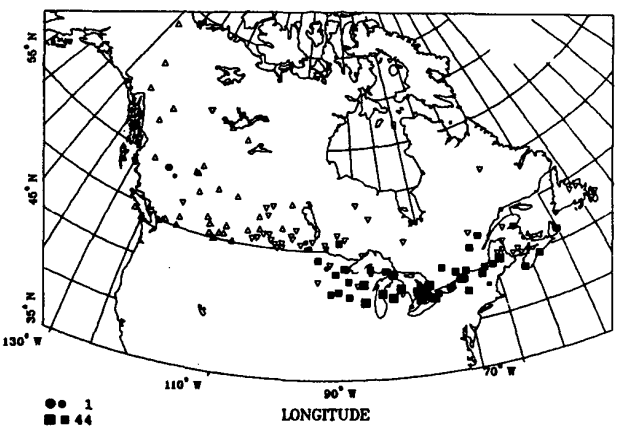
AUGUST



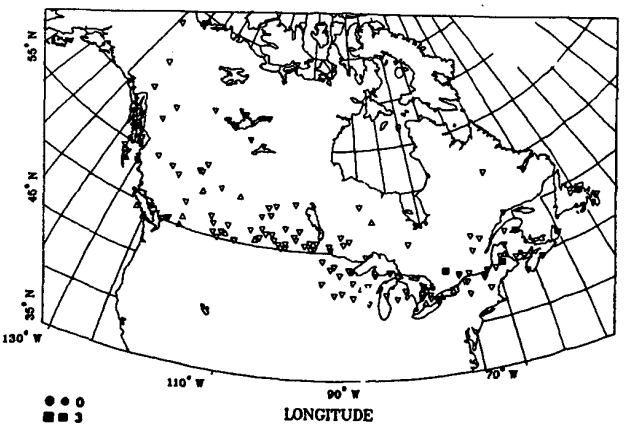
SEPTEMBER



OCTOBER



NOVEMBER



DECEMBER

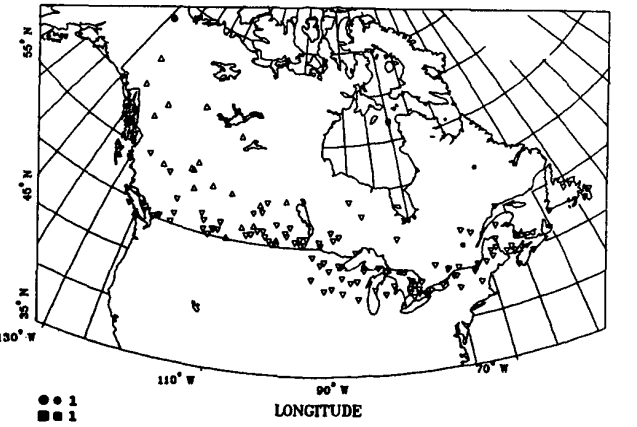


Figure 4

On the Contribution of Greenhouse Gases to Urban Heating

Dr. Oskar Essenwanger
c/o Math. Dept., MDH 204, University of Alabama in Huntsville
Huntsville, AL 35899

ABSTRACT: A possible contribution by emission gases from burning of fossil fuels to the urban heat island is investigated for daily observations of CO, NO_2, O_3, the temperture difference rural-urban, and 5 other weather parameters. The minimum temperature difference showed a high correlation with the 3 emission gases (at the 3 sigma level of confidence). Although a significant correlation does not necessarily mean a causal relationship, it may indicate that CO_2, O_3, and indirectly a higher water vapor formation could contribute about 10 to 20% to the warming on an individual day.

1. INTRODUCTION:

It is a well known fact that the average temperature in an urban area is warmer in comparison with the rural environment. This difference is caused by the different heat capacity of the stone masses in the City which raises the daily minimum temperature by releasing heat in the night, generating a "heat island". This fact is well documented (e.g. Kratzer, 1937, Landsberg, 1981 etc.). It is also known that the burning of fossil fuels has increased CO_2 in the atmosphere, which may lead to global warming (greenhouse effect). Although emission of fuel burning gases is largely found in Cities some scientists share the view that this emission does not contribute to the urban warming.

2. DATA PREPARATION:

It is unfortunate that CO_2, observations are made at places of low emission (e.g. in the US at Mona Loa, HI or Oak Ridge, TN). One has to rely on indirect methods via CO, NO_2, or O_3. Limitation of space for this article makes it impossible to discuss the chemical reactions of emission gases (see Wuebbles & Edmonds, 91). However, a reasonable assumption is that CO_2 is proportional to CO emission. Even then it is difficult to obtain daily data with simultaneous temperatures from urban and rural areas. The author located daily measurements of CO, NO_2, O_3, with City temperatures for Huntsville, AL, whose airport can still be considered to represent at least "semi-rural" conditions at the present time. Only one year of data, September 90 through August 91, were available. Therefore this study must be considered as a pilot investigation to delineate possible avenues of future research.

As a first step the monthly mean values for the chosen parameters were calculated, see Table I, in order to check whether the data were in line with present state of the art. The monthly variation of the three gases and weather elements showed the expected trend. The difference of the minimum temperatures (urban-rural) stays within limits of earlier

results by other authors (e.g. Landsberg, 1981).

They indicate that the City shows a heat island.

TABLE 1 MEAN VALUES

	CO	NO2	03	TXP	TXW	ΔTX	TMP	TMW	ΔTM	TD	CLD	CLN	P	PREC
SEP	5.8	15.5	36.3	89.3	88.4	0.9	67.9	63.9	4.0	63.0	5.1	4.9	.37	2.34
OCT	7.0	13.6	16.3	74.7	74.5	0.2	53.0	49.1	3.9	50.2	4.8	4.4	.41	3.40
NOV	12.1	17.4	11.6	68.4	68.4	0.0	46.9	42.0	4.9	43.9	5.2	4.2	.47	4.73
DEC	8.9	16.5	10.6	57.6	57.0	0.6	40.8	36.6	4.2	40.0	7.4	7.0	.48	18.68
JAN	6.3	17.3	28.9	52.2	50.9	1.3	39.2	35.0	4.2	36.5	8.2	8.0	.48	2.77
FEB	6.9	17.5	26.0	56.8	57.1	-0.3	40.6	37.8	2.8	36.6	6.6	6.8	.44	8.46
MAR	7.1	13.6	30.4	65.4	64.7	0.7	47.4	43.3	4.1	41.1	7.5	7.4	.25	9.33
APR	6.5	9.9	32.9	73.5	73.6	-0.1	58.2	55.1	3.1	54.7	8.0	7.8	.34	10.29
MAY	4.6	9.8	31.6	80.9	80.8	0.1	67.4	63.6	3.8	64.9	8.0	7.9	.37	8.68
JUN	4.7	12.2	40.6	85.6	86.1	-0.5	71.6	67.1	4.5	66.9	7.3	7.0	.34	1.72
JUL	2.9	11.9	40.3	90.6	91.7	-1.1	74.5	70.0	4.5	69.6	5.4	5.2	.34	2.59
AUG	4.5	13.2	40.0	89.4	91.0	-1.6	73.3	68.3	5.0	66.6	5.5	5.7	.37	4.53
YEAR	6.4	14.0	28.8	73.7	73.7	0.0	56.8	52.7	4.1	50.4	6.6	6.4	.39	77.52

TX=maximum, TM=minimum temperature, P=City (Pollution Office), W=airport (Weather Bureau), TD=dew point (in Fahrenheit); CL=skycover, D=daytime,N=nightime average (tenth of sky cover); P=station pressure (in $\frac{1}{100}$ of inches, deviation from 39.00); PREC=is precipitation in $\frac{1}{100}$ of inches. The unit for gases is parts per million.

3. CORRELATION ANALYSIS:

Table 2 exhibits the linear correlation coefficient of the parameters of Table 1 with the difference of the urban-rural minimum temperature. For the year, N= 365, the linear correlation coefficient is statistically significant at the 3 sigma level ≥ 0.16, for the seasons ≥ 0.30. The high correlation of clouds and the minimum temperature difference is no surprise. For clear sky the rural temperature is lower than on cloudy nights. The high correlation with CO and NO_2 implies that the temperature difference is higher when these two gases are above the annual (seasonal) mean. In other words, when the emission of these gases was high, the difference was also high. This could imply that an additional heat by CO_2 would keep the minimum in the City higher than for low emission.

While the linear correlation for the year would indicate that some relationship with the pressure exists, only the winter figure supports it. In summer the correlation with the gases is low. Ap-

TABLE 2 LINEAR CORRELATION COEFFICIENTS ΔT MIN

	YEAR	FALL	WINT	SPRI	SUM
CO	0.40	0.43	0.60	0.48	-0.08
NO2	0.31	0.32	0.48	0.47	0.05
03	-0.01	0.01	-0.24	0.03	0.18
ΔTX	0.26	0.53	0.59	0.23	-0.13
TD	0.01	-0.23	0.02	-0.38	-0.29
CLD	-0.41	-0.37	-0.46	-0.34	-0.10
CLN	-0.50	-0.45	-0.53	-0.57	-0.16
P	0.22	0.17	0.36	0.28	0.08
PREC	-0.36	-0.35	-0.43	-0.34	-0.02

parently under the higher solar angle in summer the stone masses heat up more than in wintertime. In addition, convection may contribute to a faster dispersion. Thus the influence of the fuel burning may be too small to be discovered with this limited set of data.

4. FACTOR ANALYSIS:

The multiple correlation given in Table 2 is not simple to interpret. Thus the matrix of the 10 parameters was subjected to a factor analysis (see Essenwanger, 1986b etc.). Table 3 exhibits the factor load after orthogonal rotation for the total year. In order to provide a quick look only the major numbers were included into Table 3. It is striking that the first factor comprises of weather parameters such as dewpoint, clouds (daytime), and pressure. However, the second factor depicts the emission gases and the minimum temperature difference. This would indicate a strong association between gases and the minimum temperature difference. The difference of the maximum temperature, however, discloses some independence. The fourth factor indicates that the minimum temperature difference is also influenced by cloud and precipitation but is also related to NO_2. Finally, the fifth factor delineates again an association. This time it is the cloudcover during the night. This would mean that the level of emitted gases stays high during the night of a cloudy sky.
It should be noted that the total sum of the factors is only 80%. This implies that not all factors are accounted in this scheme. It is called to the reader's attention that sunshine data were not available, and have been substituted by clouds (skycover). Thus an analysis of the heat received during the daytime, and released during the night was not possible at the present time.

TABLE 3 FACTOR LOAD YEAR, ROTATED

FACTOR	1	2	3	4	5
CO	–	.56	–	–	.76
NO2	–	.45	–	.58	.67
03	–	-.98	–	–	–
ΔTX	–	–	.95	–	–
ΔTM	–	-.83	–	-.50	–
TD	-.98	–	–	–	–
CLD	-.62	–	–	-.69	–
CLN	–	–	–	–	.86
P	-.94	–	–	–	–
PREC	–	–	–	.93	–
	22.0	17.7	10.2	14.4	15.2%
					79.4%

Table 4 discloses the results of the factor analysis for the four seasons. In every season but the summer we find a factor associating the emission gases with the minimum temperature difference. Several differences in the individual seasons are noticable. In fall and winter one factor with CO, NO_2, and the minimum temperature difference stands alone (fall, factor 3, winter, factor 5). In spring two factors with emission gases and minimum temperature appeared. In fall and winter the minimum temperature difference can be found in additional factors. Table 2 displayed no correlation (<0.1) between the minimum temperature difference and CO, NO_2, in summer. Thus no linkage is assumed during summer. This conclusion is confirmed by the factor analysis. Only O_3 stays with the minimum temperature difference in the same factor (#3). This would confirm the earlier interpretation that the much higher radiation in summer combined with the low daily average of the emission gases (except O_3) may mask any possible role of CO_2 in summer. It must also be mentioned that plants are also responsible for a low CO_2

in the summer.

TABLE 4 FACTOR LOAD FOR THE SEASONS

(rotated, only > 0.4 included)

	FALL					WINTER					SPRING					SUMMER				
FACTOR	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
CO	.68	–	.64	–	–	–	-.48	–	–	.77	.84	–	–	–	.47	–	.93	–	–	–
NO2	–	–	.99	–	–	–	–	–	–	.95	.94	–	–	–	–	.45	.83	–	–	–
03	–	.99	–	–	–	–	–	–	.99	–	–	.99	–	–	–	–	–	.88	–	–
ΔTX	–	–	–	.98	–	.57	–	.72	–	–	–	–	–	.95	–	–	–	–	.96	–
ΔTM	–	–	.43	.84	–	.63	.44	–	–	.51	.72	–	.51	–	.44	–	–	.88	–	–
TD	–	.87	–	–	–	.78	.55	–	–	–	-.90	–	–	.42	–	–	.49	.65	–	.49
CLD	-.53	–	–	–	.78	–	.47	–	–	–	–	–	.96	–	–	-.98	–	–	–	–
CLN	-.53	–	–	–	.79	–	.95	–	–	–	-.54	–	.74	–	–	-.98	–	–	–	–
P	.93	–	–	–	–	–	–	.94	–	–	–	–	–	–	.99	–	–	–	–	.99
PREC	–	–	–	–	.98	–	–	.74	–	–	-.63	–	.69	–	–	.85	–	–	.58	-.46
SUM	16.3	17.5	13.4	16.0	21.4	12.2	26.3	19.0	11.5	17.1	26.2	10.6	17.8	10.9	12.8	24.5	14.8	14.5	11.6	12.4
					84.5%					86.2%					78.3%					77.7%

5. POWERSPECTRUM ANALYSIS:

A powerspectrum analysis (see Essenwanger, 1986a, p156) was performed. All parameters of Table 1 rendered a 3 and 5 day cycle possibly weather related. Longer cycles were not uniform for all parameters. Fall and winter showed a 7 day cycle for emission gases which could be caused by the weekly traffic pattern.

6. CONCLUSION:

A statistically significant correlation was found between emission gases and the minimum temperature difference urban-rural for all seasons but the summer. This may indicate that on a daily basis a 10 to 20% contribution by emission gases from fossil fuel burning could exist. It is too early, however, to draw a general conclusion from this pilot study as the data are limited to one year and one location.

7. REFERENCES:

Essenwanger, 1986a, Elements of Statistical Analysis, World Surv of Clim.,V. 1B, pp 424, Elsevier, Amsterdam.

Essenwanger, 1986b, Comparis. of Princ. Comp. and Factor Analysis Meth. for Clim. Param. Proc. 3rd Int. Conf. Statist. Clim., Vienna, pp. 40-45.

Kratzer, 1937, Das Stadtklima Die Wissensch, pp. 184, Vieweg & Son, Braunschweig.

Landsberg, 1981, The Urban Clim., pp. 275, Acad. Pr., N. York.

8. ACKNOWLEDGMENTS:

Thanks goes to the Huntsville Pollution Office for the records on gases and City temperature. Ms. Tami Lang deserves the credit for typing the manuscript.

COOLING OVER MONSOONAL PAKISTAN

P.D. Kruss[1], K.A.Y Khan[2], F.M.Q. Malik[2], M. Muslehuddin[2] and A. Majid[2]

[1] World Meteorological Organization

[2] Pakistan Meteorological Department

The tropical and subtropical monsoon-affected regions of the world are attracting particular attention at this present time, especially with respect to links to ENSO and in the context of global change. A newly-created climatological database for Pakistan has made possible the study of the changes in temperature occurring over Pakistan during the last 60 years.

A personal computer database containing 293 million characters of meteorological information has been created by the Pakistan Meteorological Department over the past three years with the support of a World Meteorological Organization/United Nations Development Programme project. Of particular interest to this present study are the monthly values of maximum and minimum temperature, rainfall and cloudiness available for the period 1931-90 for 55 Pakistani stations.

Using the above-specified data, 30-year normals of maximum, minimum and mean temperature for the two periods 1931-60 and 1961-90 were constructed. Stations associated with marked urbanization or with less than 20 years of data in either period were then eliminated, resulting in 35 stations well-distributed over Pakistan remaining for this study.

In comparing the normals for the two periods, it was found that, in contrast to mean global trends, cooling had occurred over approximately half the area of Pakistan. The two regions of cooling identified were both without exception and cover most of the northern half and a small area in the southeast corner of Pakistan. Warming was evidenced elsewhere.

The regions of cooling are very closely associated with the area encompassed by the 100 mm monsoon rainfall isohyet. Further, initial indications are for an increase in both monsoon cloudiness and rainfall over the same regions and period, thus suggesting that a significant increase in monsoon activity over Pakistan has brought about the observed cooling. This underscores the importance of cloudiness/precipitation treatment in numerical modelling studies of greenhouse-forced temperature trends, particularly in monsoon-affected areas.

USING A STATISTICAL CONTRAST METHOD TO DETERMINE TRENDS OF MEASURED SURFACE AIR TEMPERATURE

C.Hervada
Dept. de Física i Enginyeria Nuclear, EUETIT,
Universitat Politècnica de Catalunya, Spain

V.Pawlowsky
Dept. de Matemàtica Aplicada III, ETSECCPB,
Universitat Politècnica de Catalunya, Spain

R.Soler
Dept. d'Astronomia, Astrofísica i Física de l'Aire, Fac. de Física,
Universitat de Barcelona, Spain

April 9, 1992

Abstract

The goal of this report is to show how a well known statistical method can be used to determine trends of measured surface air temperature in Catalunya. In order to do that, temperature data from 31 meteorological stations corresponding to winter periods from 1957 to 1986 have been used. In a first step, 5-year running mean has been used because it provides a simple smoothing which helps to clarify long-term change. Being so huge the data volume it is possible to consider temperature following a normal distribution and 5-year running mean too. If there are no trends, the mean of first differences will be zero. If mean is positive, it is possible to say that temperature tends to grow and if mean is negative temperature tends to be lower with time. So a t-Student test has been done for the 31 stations. The results from that test agree with graphical analysis of the data sets.

1 Introduction

Surface air temperature has been measured at a large number of meteorological stations since last century. Based on these data many analyses of trends in the average surface air temperature have been performed. Traditionally, the determination of trends of measured temperature was made by graphical techniques like in Lamb (1950), Callendar (1961), Hansen (1987,1988) and Jones (1986,1989,1991). Well known statistical techniques were applied by Díaz (1986), who worked with principal component analysis. In this paper an alternative approach, based on Student's t distribution, is proposed, so that the statistical reliability of presence or absence of trends can be established.

2 Data and Method

Daily maximum and minimum winter surface air temperature are available at 31 stations spread in Catalonia, at the north west Mediterranian bassin. Fig.1 shows Catalonia's situation in Europe and fig.2 the location of stations. All these measuring points belong to the Spanish National Meteorological Institute.

To avoid cold spells, different number of hours of sunshine and other cyclic terms, annual winter maximum and minimum mean temperatures at each of those 31 points, from 1957 to 1986, have been used. Thus, the total number of available data for each station is 30 maxima and 30 minima. Winter lasts from December to March as in Lamb (1950) and Lund (1961).

Graphical analysis of data is done. In a first step, 5-year moving averages are used because they provide a simple smoothing which helps to clarify long-term changes. That means there is assigned each year the mean of a 5 year period (from two before to two after it). The problem that arises when working with these average temperatures is, that instead of having each station characterized by 30 temperatures, it will be characterized by 26 (30 - 4) data both for maxima and minima. The advantage in doing so is that any random or cyclic movement whose period is larger than a year is smoothed and just trends will last (Hansen et al.1987,1988). So there is a small set of data, 26 points per station.

First difference series, one for maximum and the other for minimum temperatures, are defined. Each of them has 25 (26 - 1) elements. Each of its elements, d_i, is minus the difference between average maximum (or minimum) temperature of a year, x_i , and the one for the following year, x_{i+1}.

$$d_i = x_{i+1} - x_i \tag{1}$$

If there is no trend, elements of first difference series are statistically independent and, as we are dealing with mean temperatures, it can be assumed that they are realisation of a random variable wich follows a normal distribution. As only 25 data are available for each measuring point, small sampling theory has to be used to perform statistical analysis.

As it is well known, Student devised a test for small samples, that is, less than 30. This has become known as Student's t test which is widely used to perform tests of hypothesis about population mean (Spiegel 1961; Benjamin 1970). Student's t statistic used for this purpose is:

$$t = \frac{D - \mu}{s}(N-1)^{1/2}, \tag{2}$$

where D is the sample mean, μ the mathematical expectation of the considered random variable, s the standard desviation of the sample and N the number of elements in the sample.

If nul hypotesis is that there is no trend in a series of first differences, then $\mu = 0$. We work with a 90 per cent, 95 per cent and 99 per cent confidence. If t is then greater than t_o it will mean that temperature tends to grow, and if t is lower than $-t_o$ temperature tends to be lower with time.

First difference series have 25 samples, and in that case, for a 90 per cent confidence, $t_o = 1.32$, for a 95 per cent, $t_o = 1.71$ and for a 99 per cent, $t_o = 2.48$.

3 Results and Conclusions

Fig. 3 to fig. 6 are examples of the graphical analyses of data and table 1 shows the corresponding results obtained for the Student's t test.

We can conclude that the two approaches agree. Student's t test gives a 99 per cent significance if there has been any variation of $2.48°C$ or more and

a 90 per cent significance if variation has been of $1.32°C$ or more during those 30 years.
Results obtained for Begur (fig. 4) are not representative because this meteorological station has been moved: in 1957 it was in the middle of the village and in 1969 it was moved to Begur's lighting house, about 150 m up on a rock.
Results obtained for Turó de l'Home are representative of global air mass over the country. In fact, this station is on the top of a high mountain (1752 m) surrounded by plain (less than 500m). It looks like a free air measuring point. They show (fig.5) that maxima and minima have been quite the same in those 30 years. Results obtained for this station with Student's t test say that the values of - .33 for maxima and - .37 for minima are not significative, so temperatures show no significant variation in those years.
Adrall (fig. 3) shows a growing of its maxima and the test says that this growing is significative, there the two methods agree too.
For Villanova (fig.6) there is a small growing of maxima during the last years. Students' test says that this growing is significative with a 90 per cent confidence, but not in a 99 per cent.
The advantage in using Students't test to study trends is that results are given as a number, which allows anybody to know immediately if there is any trend and if it is significant or not. The disadvantage it has is that it is impossible to know the evolution of data with time for forecasting purposes.

Figure 1:

Figure 2:

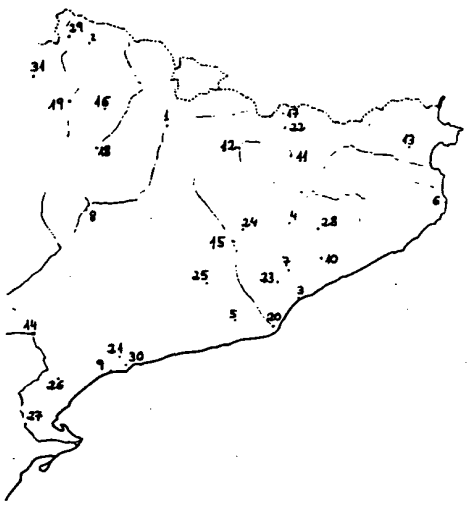


Figure captions

Fig. 1: Catalonia is at the North West of the Mediterranian bassin.

Fig. 2: Location of the 31 measuring points.

Fig. 3 to 6: 5-year moving averages for maximum and minimum temperatures at some stations.

Figure 3:

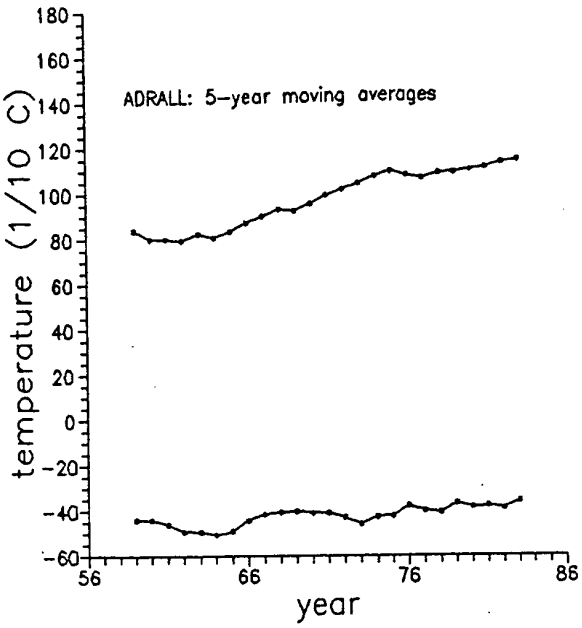


Figure 4:

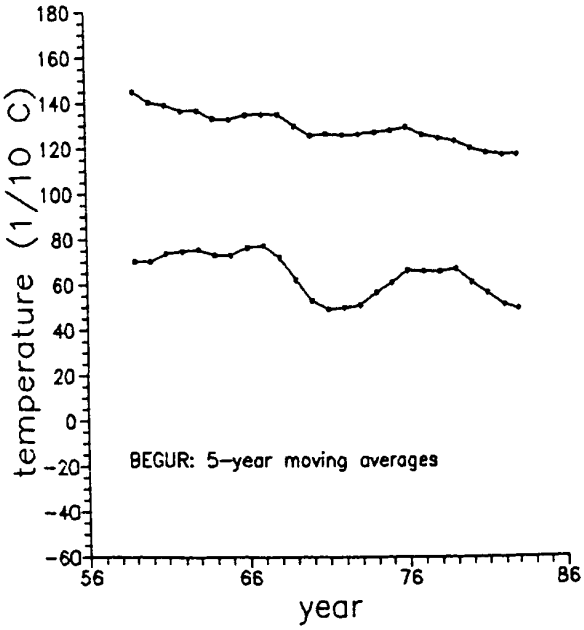


Figure 5:

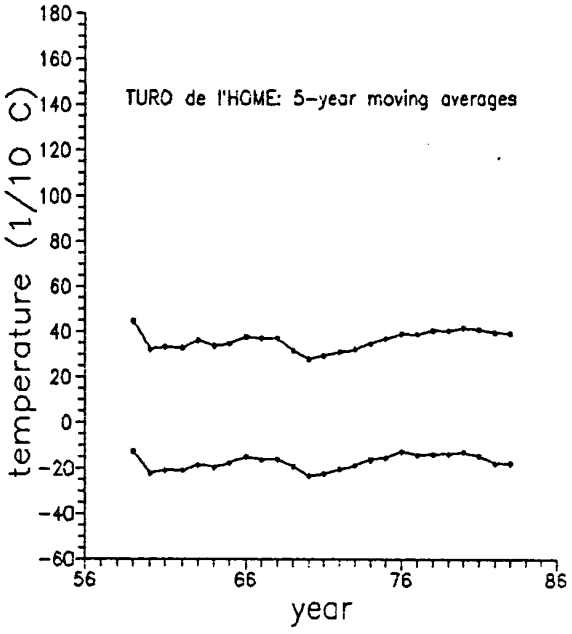
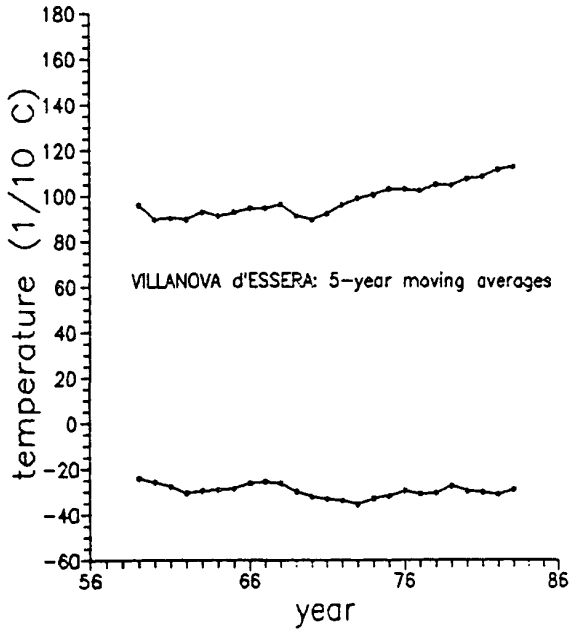


Figure 6:



Tables

Table 1: Station name and its number

Station No	Station name
1	Adrall
2	Arties
3	Badalona
4	Balenyà
5	Begues
6	Begur
7	Caldes de Montbui
8	Camarasa
9	Cambrils
10	Cardedeu
11	la Farga de Bebié
12	Fígols
13	Figueres
14	Flix
15	Manresa
16	Molinos
17	Núria
18	la Pobla de Segur
19	el Pont de Suert
20	el Prat de Llobregat
21	Reus
22	Ribes de Fresser
23	Sabadell
24	Sallent(Cabrianes)
25	Sta. Margarida de Montbui
26	Tivissa
27	Tortosa
28	el Turó de l'Home
29	Viella
30	Vilaseca
31	Villanova d'Essera

Table 2: Some results obtained in Student's t test

Stat	t	90%	95 %	99 %
1	max.3.05	yes	yes	yes
	min.0.70	no	no	no
6	max.-2.98	yes	yes	yes
	min.-1.08	no	no	no
28	max.-0.33	no	no	no
	min.-0.37	no	no	no
31	max.1.38	yes	no	no
	min.-0.57	no	no	no

References

- Benjamin J.R. and C.A. Cornell, 1970: *Probability, Statistics and decision for civil engineers.* McGraw Hill Inc., 685 pp.
- Callendar G.S., 1961: Temperature fluctuations and trends over the earth.*Quart. J. Royal Met. Soc.* **87**, 1-12.
- Díaz H., 1986: An Analysis of Twentieth Century Climate Fluctuations in Northern North America.*J. of Clim. and App Met.* **25**, 1625-1656.
- Hansen J. and S. Lebedeff, 1987: Global Trends of Measured Surface Air Temperature.*J. of Geoph. Res.* **92**, 13345-13372.
- Hansen J, I. Fung, A. Lacis, D. Rind, S. Lebedeff, R. Ruedy, G. Russell and P. Stone, 1988: Global Climate Changes as Forecast by Goddard Institute for Space Studies Three- Dimensional Model.*J. of Geoph. Res.* **93**, 9341-9364.
- Jones P.D., S.C.B. Raper, R.S. Bradley, H.F. Díaz, P.M. Kelly and T.M.L. Wigley, 1986: Northern Hemisphere Surface Air Temperature Variations: 1851-1984.*J. of Clim. and App. Met.* **25**, 161-179.
- Jones P.D., P.M. Kelly, C.M. Goodess and T. Karl, 1989: The Effect of Urban Warming on the Northern Hemisphere Temperature Average. *J. of Climate* **2**, 285-290.
- Jones P.D., 1991: Global Temperature Variations Since 1861: The Influence of the Southern Oscillation and a Look at Recent Patterns of Warmth. Personal comunication, 24 pp.
- Lamb H.H., 1950: Types and spells of weather around the year in the British Isles: Annual trends, seasonal structure of the year, singularities.*Quart. J. Roy. Met. Soc.* **76**, 393-438.

- Lund I.A., 1963: Map-pattern classification by statistical methods. *J.App.Met.* **2**, 56-65.
- Spiegel M.R., 1961: *Statistics theory and problems*, McGraw Hills Inc., 357 pp.

A DETERMINISTIC APPROACH TO THE CREATION OF A BASELINE DATASET UTILIZING THE HISTORICAL TEMPERATURE AND PRECIPITATION DATA FROM THE U.S. COOPERATIVE NETWORK

Thomas Reek, Stephen R. Doty, Timothy W. Owen

National Oceanic and Atmospheric Administration
National Climatic Data Center, Asheville, North Carolina

ABSTRACT

The National Climatic Data Center's Summary Of The Day Database (TD3200) serves as a major baseline dataset for detecting climate change. However, it is widely known that this data base contains some erroneous daily values resulting from data entry, data recording and data reformatting errors. Many of these errors are systematic and possess causal reversibility. A deterministic approach, employing empirically developed rules, has resulted in the detection and correction of most of these errors. A computer program utilizing Backus Normal Form (BNF) design has been developed as a means of modeling the human subjective process of manual data review. This paper will review the development of this program and highlight the results of the quality control process on the 25,000 station TD3200 data base.

1. INTRODUCTION

For over one hundred years volunteer weather observers, comprising what is known today as the U.S. Cooperative Observer Network (COOP), have faithfully recorded daily weather data on forms supplied by the U.S. National Weather Service (NWS) and its predecessors. Today, the COOP system consists of some 8,300 stations located throughout the contiguous U.S., Alaska, Hawaii, Pacific Islands, Puerto Rico and the Caribbean. Over the years some 25,000 stations have at one time or another participated in the COOP system. Daily records consist of maximum (TMAX) and minimum (TMIN) temperature, precipitation (PRCP), snowfall (SNOW), snow depth (SNWD) and in some cases wind, evaporation and soil temperatures. A detailed description of the COOP system has been given by the NWS (1987). Presently, the hand written forms are sent directly to the National Climatic Data Center (NCDC) for purposes of Quality Assessment (QA), Quality Control (QC), digitization, dissemination and archiving. Prior to the centralization of this function at NCDC in 1962, COOP observation forms found their way to various regional processing centers where the data were compiled for publications such as the current *Climatological Data* (CD). As punched cards became popular in the late 1940's and early 1950's, the publication process became semi-automated. Still, punch cards were viewed primarily as an interim medium for hardcopy production. Errors detected during hardcopy production were not always corrected within the punched card data medium. These collections of punched cards continued to grow until magnetic tape became the automation and storage medium of choice in the 1960's when NCDC began to acquire and archive this ever growing and increasingly valuable collection of daily weather data.

2. QUALITY CONTROL BACKGROUND

Agreements were made with various state universities to digitize (card keypunch) many of the pre-1948 data in order to expand the data base. To date, this trove of climatic data, known as TD3200 (NCDC, 1991), contains decades of data from some 25,000 stations of varying periods of record. As demand for computer access to these data grew, the need for stringent quality control became apparent. Over the years, QC has ranged from little or none to the complex spatial comparison techniques used today at the NCDC. Reek and Crowe (1991) introduced the use of Geographic Information Systems (GIS) and Expert Systems (ES) techniques as components of the operational QC process in 1989. Prior to this time, QC was primarily a manual intensive subjective review process.

Digitization methods, the various data medium transformations and data reformats (card to tape, tape to tape, etc.) sometimes introduced errors. It is unknown how much of these data were key verified prior to 1962. Verification was not standard among the numerous sites of data entry. Single key entry procedures were established at NCDC in 1962 and continued until 1989. Key verification (double keying) began thereafter as a cost and quality effective QC procedure. Prior to double keying, key entry errors were only addressed when uncovered by the QC procedures in effect at that time. The TD3200 database format was established in January 1982, at which time provisions were made to retain both the originally observed (entered) value and an NCDC edited (suggested) value for erroneous (suspect) data. Prior to the establishment of TD3200, erroneous observed values were discarded. Certain regionalized data entry (QC) procedures also resulted not only in

the loss of "original" data, but introduced errors as well.

Other problems of common occurrence include: the loss of punched cards, Electronic Data Processing (EDP) machines mangled or misread/mispunched cards, card storage procedures/transfers and environmental hazards adversely affected the integrity and usability of the cards. Many of these data were digitized by a method known as Film Optical Sensing Device for Input to Computers (FOSDIC), by which filmed images of punched cards scanned and written to magnetic tape. This process, although superior to reading the often illegible original cards, was not without error. Observers and observation practices also introduced errors, some of which are systematic. Reek et al. (1992) describes in great detail the testing and development of a QC process known as ValHiDD (Validation of Historical Daily Data) that automatically uncovers and corrects both the EDP artifact type of error and the systematic observation practice error. Space limitations within this paper prohibit detailed discussion of ValHiDD.

3. TYPES OF ERRORS

Each of the digitization and media transformations mentioned above carries its own flavor of error introduction. This being the case, the errors are therefore systematic and causal. Each possesses a distinct signature that once recognized can be reversed. Consider the examples shown in Figure 1. In all, seventeen distinctly different causal error signatures were determined by empirical review. When encountered, an Expert Systems (ES) rules base containing explicit correction modalities is invoked. Corrected data are subjected to the same scrutiny applied to the original data. Both the original and suggested correction are maintained in TD3200. Corrections that fail subsequent testing become "deletion codes". Another reason for deletions is extreme values for which no known signature exists.

In some cases, absolute causality cannot be determined due to a signature overlap. Consider, for example, the case of a maximum temperature of 127 degrees Fahrenheit for which the previous and subsequent day's maximums are 11 and 33, respectively. Each day's minimum is a single digit value. One signature (Code 19) suggests an erroneous leading ones digit (a common EDP artifact), resulting in a corrected value of 27. Another, but less prevalent, signature (Code 15) indicates data had been entered to tenths resulting in a corrected value of 13. In this instance an action rule of "Epekeia" is invoked. Taken from canon law, Epekeia simply advocates an equitable but not absolutely correct action. The 127 is, however, absolutely incorrect. The result here, based on the process of eruditional examination, would be a correction value of 27.

To facilitate the identification of errors and suspected errors, while at the same time trying to prescribe correct values, is no easy task. The method chosen was to

Signature Code	Element	Three Day Series			Correction Day N	Narrative
		Day N-1	Day N	Day N+1		
9	PREC	Missing	8.63	0	A8.63	Set Accumulation Flag
11	PREC (OR TMAX/MIN, SNOW)	Trace	89.63	0.12	DELETE	Exceeds Climatological Limits
15	TMAX OR TMIN	58	537	64	54	Data Entered To Tenths
17	TMAX OR TMIN	88	-85	92	+85	Erroneous Sign
19	TMAX OR TMIN	84	189	80	89	Erroneous Leading Digit
25	TMAX TMIN	92 78	73 55	68 52	78 55	TMIN(N-1), TMAX(N) Internal Inconsistency
23	TMAX TMIN	92 78	55 88	68 52	88 55	Switch TMAX and TMIN
29	TMAX TMIN	92 78	0 88	68 52	88 DELETE	Move TMIN To TMAX, Delete TMIN
31	[Days: TMAX OR TMIN	N-1...4 85 72	N 85 72	N+1 93 77	N, N-1...3] DELETE DELETE	TMAX and TMIN Constant for 5 or More Consecutive Days
41	TMAX TMIN	68 42	103 2	74 45	DELETE DELETE	Discontinuity In Diurnal Range
53	TMAX TMIN PREC SNOW SNWD	28 17 0 0 1	32 27 0.14 11.0 2	24 14 0 0 2	32 27 0.14 1.1 2	SNOW Decimal Error
55	TMAX TMIN PREC SNOW SNWD	14 3 0 0 0	19 0 0 T T	27 0 0 0 0	19 0 T T T	Set PREC to Trace value
57	TMAX TMIN PREC SNOW SNWD	28 17 0 0 1	32 27 1.40 1.1 2	24 14 0 0 2	32 27 0.14 1.1 2	PREC Decimal Error
59	TMAX TMIN PREC SNOW SNWD	25 14 0 0 0	28 25 0.50 0 5	27 21 0 0 5	28 25 0.50 5.0 5	Estimate Missing SNOW
61	TMAX TMIN	44 15	97 4	36 11	DELETE 4	Temporal Inconsistency
71	TMAX TMIN	92 78	88 0	68 52	88 DELETE	Delete TMIN
91	TMAX TMIN PREC SNOW SNWD	88 64 0 0 0	0 45 7235 99 140	73 33 Trace 0 0	DELETE DELETE DELETE DELETE DELETE	Multiple Unreconcilable Errors

Figure 1. ValHiDD error signature codes.

build a discrepancy file containing the station identification, date of offending value, data element in error (#); TMAX, TMIN, PRCP, SNOW, SNWD, and an error code (EC) that identifies the rule violated, severity of the infraction and suggested prescription for remedy. The EC is a two digit numerical value organized into rules groups. The first digit (1-9) signifies the rules group responsible for its production. The second digit (1-9) indicates a specific corrective action if odd, or a level of suspicion (but no action) if even. As shown in Figure 2, the

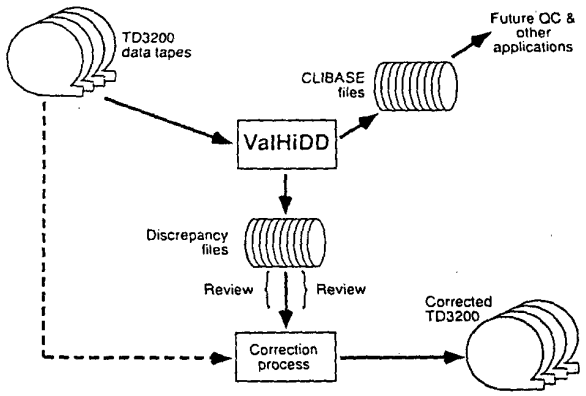


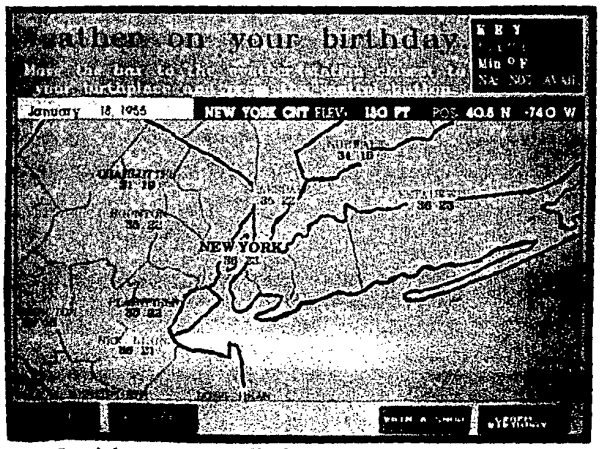
Figure 2. ValHiDD system flow diagram.

TD3200. The software was developed uses a Backus Normal Form grammar that differentiates good-data/good-data-relationships from bad ones. Data are examined three days at a time with the exception of some longer time series continuity tests.

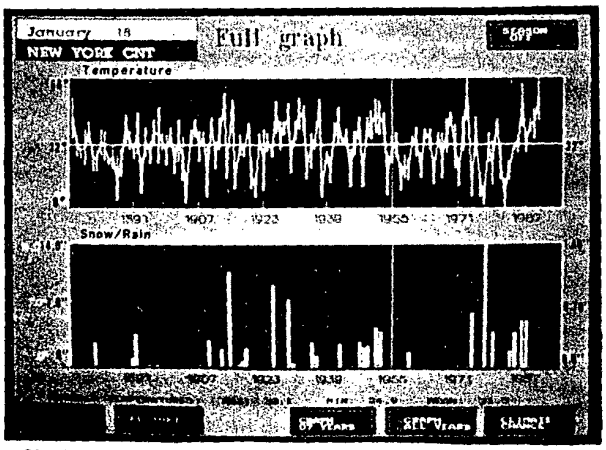
4. CLIBASE

ValHiDD has resulted in a database known as CLIBASE (Figure 2) from which all known EDP and systematic observer errors have been removed. CLIBASE is a packed binary, object oriented data base, complete with data access particulars in an encapsulated program structure. The software is written entirely in Fortran. The file structure and record formats represent a unique employment of directory driven bit packed binary encoding techniques. Highly efficient access, search, retrieval and update functions are achieved via ISAM (Indexed Sequential Access Method) input/output operations. CLIBASE is comprised of period of record entries for 2,000 stations within the contiguous U.S. Most stations have records of 70 years or greater. The parameters available are: daily maximum and minimum temperature and daily precipitation, snowfall and snow depth measurements. The first application of CLIBASE was a system known as CLIVUE for CLimate Interrogation, Visualization and User Evaluation. The system employs a human graphical interface that allows the user to view data in both temporal and spatial depictions. Figure 3 demonstrates some of the potential uses of CLIBASE via the CLIVUE system. NCDC will distribute the CLIVUE/CLIBASE system on CDROM on or about July 1, 1992. For further information, contact Tom Ross at (704) 259-0994.

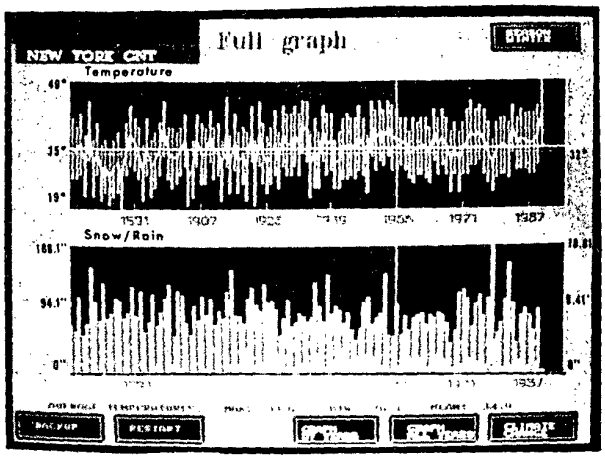
Figure 3. CLIVUE screen displays.



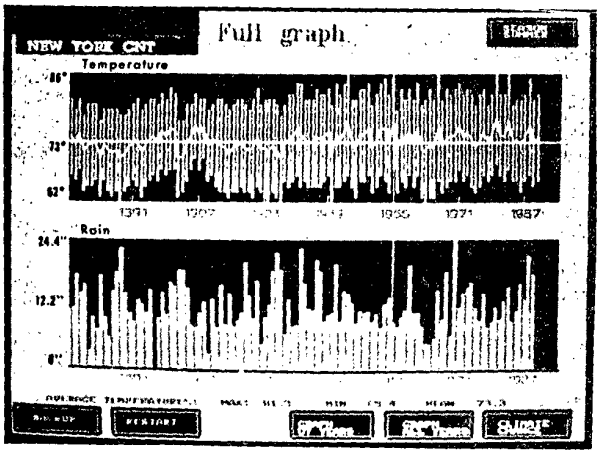
3a. Spatial temperature display for New York City area, Jan. 18, 1955.



3b. Daily period of record graph display for New York Central Park, Jan. 18 (1876 to 1990).



3c. Season mean/total graph display for New York Central Park (1876 to 1990), Winter (Dec. 21 to Mar. 20)



3d. Season mean/total graph display for New York Central Park (1876 to 1990), Summer (Jun. 21 to Sep. 20)

5. CONCLUSION

The ValHiDD/CLIBASE effort has resulted in a baseline dataset free of gross systematic errors that were often culprits of statistical discordancies in studies using TD3200. Although the actual number of errors discovered and corrected was small, their removal is significant as it is these errors that have served to discredit TD3200. Applications such as CLIVUE can proceed with greater efficiency and confidence. ValHiDD has been run on the entire TD3200 database. The result is a correction file of 232,167 entries that will eventually be applied to TD3200. Figure 4 summarizes the ValHiDD correction file. The ValHiDD software is now an operational aspect of the monthly quality controlled extensions to TD3200.

	Observations (in thousands)			Discrepancies			Percentage Discrepancies to Observations		
Decade	*TEMP	*PRCP	TOTAL	TEMP	PRCP	TOTAL	TEMP	PRCP	TOTAL
< 1900	723	1,119	1,842	2,539	533	3,072	0.3511	0.0476	0.1667
1900	2,913	4,394	7,307	4,927	1,575	6,502	0.1691	0.0358	0.0890
1910	4,183	6,830	11,013	7,460	2,851	10,311	0.1784	0.0417	0.0936
1920	5,687	9,373	15,060	6,340	3,795	10,135	0.1115	0.0405	0.0673
1930	10,348	16,849	27,197	9,886	7,227	17,113	0.0955	0.0429	0.0629
1940	16,188	34,562	50,750	11,644	13,469	25,113	0.0719	0.0390	0.0495
1950	39,011	100,923	139,934	7,302	30,882	38,184	0.0187	0.0306	0.0273
1960	43,101	98,453	141,554	3,367	34,394	37,761	0.0078	0.0349	0.0267
1970	41,902	91,012	132,914	4,703	37,701	42,404	0.0112	0.0414	0.0319
1980	38,738	81,925	120,663	2,916	35,283	38,199	0.0075	0.0431	0.0317
1990	3,820	8,116	11,936	544	2,829	3,373	0.0142	0.0349	0.0283
	206,614	453,556	660,170	61,628	170,539	232,167	0.0298	0.0376	0.0352

* TEMP: TMAX and TMIN; PRCP: PREC, SNOW and SNWD

Figure 4. Total number of ValHiDD observation reviews and corrections

6. ACKNOWLEDGEMENTS

The Authors wish to thank Pamela Hughes, Elaine Mason and Eric Gadberry of NCDC for their valuable assistance in the development of ValHiDD and the CLIBASE/CLIVUE system.

7. REFERENCES

National Weather Service, 1987: Cooperative Program Management, Weather Service Operations Manual B-17 (revised), NOAA-NWS, Silver Spring, MD.

Reek T., S.R. Doty and T. Owen, 1991: ValHiDD Documentation and Users Guide, Internal NCDC Document, 20 pp.

Reek T., S.R. Doty and T. Owen, 1992: A Deterministic Approach to the Validation of Historical Daily Temperature and Precipitation Data from the Cooperative Network. Bulletin of the American Meteorological Society, In Press, scheduled for the June 1992 edition.

Reek T. and Crowe, M., 1991: "Advances in Quality Control Technology at the National Climatic Data Center". Preprints of the Seventh International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography and Hydrology, New Orleans, American Meteorological Society, pp. 397-403.

THE VARIATION OF THE PRINCIPAL COMPONENTS OF TEMPERATURE IN CHINA AND THE NORTHERN HEMISPHERE WARMING

Tu Qipu
(Nanjing Institute of Meteorology)

Abstract

I. THE PRINCIPAL COMPONENTS OF TEMPERATURE IN CHINA

The principal component analysis of annual mean temperature series of 42 chinese stations which are scattered in various climate provinces shows that the main feature of temperature change in China during the past 100years can be characterized by the first three principal components. For most temperature series more than 70% variance contribution come from the first three principal components and 80% correlation coefficients between the temperature series and the first principal component are in the range from 0.60 to 0.90.

II. THE VARIATION OF TEMPERATURE PRINCIPAL COMPONENTS IN CHINA AND THE NORTHERN HEMISPHERE WARMING

The first principal component(PC1) of temperature in China are significantly correlated with the Northern Hemisphere mean temperature(NHMT), but there is not any significant correlation between the PC2,PC3 and the NHMT. It is shown that the major characteristics of the long-term trend of the PC1 of annual mean temperature in China are very similar to that of the NHMT by fitting a spline function to the PC1 series and the NHMT series sapratly. The distict warming trend is the common feature of the long-term variation for both the NHMT and the temperature in China. According to the observation the relationship between the interannual variation of the PC1 of temperature in China and the NHMT can be expresed by the regretion equation

$$P_1 = 0.33 + 3.25\Delta T_n$$

The PC1 loading of annual mean temperature in China has a

positive value at each stations and its spatial distribution is quite uniform. When NHMT is above normal the PC1 will be positive and positive anomalies of temperature will occur in a large areas of China.

III. THE ESTIMATED TEMPERATURE CHANGE ASSOCIATED WITH THE NORTHERN HEMISPHERE WARMING

According to the correlation between the PC1 of temperature in China and the NHMT the temperature anomaly at the respective stations can be estimated by the fomular

$$\Delta T_i = (\sum_k a_{ki} \cdot P_k) \cdot S_i$$
$$= a_{1i} \cdot (0.33 + 3.25\,\Delta T_n) \cdot S_i$$

when the NHMT is increased by ΔT_n, where S_i is the standard diviation of the temperature of the i-th station. The regretion equation indicates that the amplitudes of the annual mean temperature anomalies in the most part of China are larger than those of Northern Hemisphere, particuraly in such regions as the middle and lower recahes of Changjing River and the northern part of the Northern China Plain, where the temperature anomalies is expected to be 1.4-1.6 times as high as those of the Northern Hemisphere. Meanwhile, it is smaller in the northern part of the Northeast China Plain, and Qinghai-Xizang Plateau, being only 0.4-0.8 times the size of those of the Northern Hemisphere.

TEMPERATURE TRENDS IN BRAZIL

C.A. Sansigolo, R.C.M. Rodrigues, P.C.E. Etchichury
CPTEC, INPE, S.J. Campos, SP, Brazil

1. Introduction

Compregnsive analyses of temperature trends (eg. Jones and Wigley, 1990) provide fairly convincing evidence that on a global basis the atmosphere near the Earth's surface is warmer now than a century ago. The results of Hansen and Lebedeff (1987) indicate a global warming of about 0.5 - 0.7 ^{o}C in the past century, with warming of similar magnitude in both hemispheres. A long-term warming trend amounting to about 0.5 ^{o}C over the past 100 years is presented by Jones et al. (1986) for the Southern Hemisphere. These results for the surface air temperature data are supported by the available marine data (Folland et al., 1984).

The cause of the world's rising temperature trend is frequently attributed to a "greenhouse effect" resulting from a build up of atmospheric traces of CO_2 and other trace gases. Goodridge (1990) suggest, based on California data, that the heat trend measured in most of the large cities of the world is an expression of urban thermal pollution rather than a global atmospheric heating. The potential bias of the urban effect in global data sets has not been definitively assessed. At present, only rough estimates of the potential impacts of urbanization can be given. This includes an urban bias in the Hansen & Lebedeff (1987) data over the USA between 0.3 and 0.4 ^{o}C over the twentieth century, which is larger than the overall trend in the USA during this period.

Jones et al. (1989) estimated that earlier analyses of global trends may contain a spurious urban induced component of order 0.1 ^{o}C for the first eight decades of this century. They conclude, however, that this rise only accounts for a fraction of the observed global rise of around 0.5 ^{o}C.

Although temperature records for the southern hemisphere are less extensive and representative than those for the northern hemisphere the knowledge of their trends may will be essential in order to detect whether any CO_2 and trace gas induced global warming is occurring.

The purpose of this study is to present a preliminary analysis of annual mean surface temperature for 9 brazilian cities aiming at the detection of significant long-term trends.

2. Data and Procedures

The temperature trends were determined for 9 large cities, where long and continuous records were available : Manaus, Belém, Fortaleza, Salvador, Cuiabá, Rio de Janeiro, São Paulo, Curitiba, and Porto Alegre. Their 1991 population in millions are respectively : 1.01, 1.25, 1.76, 2.06, 0.40, 5.34, 9.48, 1.29, and 1.26. Figure 1 shows the spatial distribution of the stations, which are main climatological under the National Department of Meteorology.

The temperature trends were obtained from a simple linear regression of average annual temperature and calendar year.

Kendall rank tests (Sneyers, 1975) were used to decide on the significant bias.

3. Results and Conclusion

The annual average temperature time series and their linear trends for the 9 selected stations are shown in figure 2 (1-9).

Long-term warming trends varying from 0.007 to 0.012 ^{o}C/year were observed in Belém (0.007), Fortaleza (0.009), Salvador (0.007), and Curitiba (0.012). Higher values were observed at São Paulo (0.03), and Rio de Janeiro (0.02). There were no trends in the series for Manaus, Cuiabá, and Porto Alegre. The results for São Paulo and Rio de Janeiro are similar to those obtained for the larger cities in California (Goodridge, 1990), and Australia (Coughlan et al., 1991).

The minimum, maximum, and daily mean annual average temperature bias for São Paulo were 0.032, 0.023, and 0.030 ^{o}C/year, respectively. As expected, the urbanization affected much more the minima than the maxima.

A significant periodicity of approximately 40 years was found in the Cuiabá temperature time series.

Obviously with only 9 stations located in the large cities is impossible to get conclusive results on a global warming in Brazil. As no trends were observed in Manaus, Cuiabá, and Porto Alegre one can say that only the urbanization has been influencing the temperature at the Brazilian cities considered here.

4. References

Coughlan, M.J.; Tapp, R. & Kininmonth, W.R. 1991 Trends in Australian Temperature Records. In: D.E. Parker (ed.) Intergovernemental Panel on Climate Change. pp.III.1-III.28.

Goodridge, J. 1990. Air Temperature Trends from 1916-87 in California. Sixth Conf. Appl. Climat. (Preprints). pp.116-119.

Hansen, J. & Lebedeff, S. 1987. Global Trends of Measured Surface Air Temperature. J. Geophys. Res. 92:13345-13372.

Jones, P.D.; Raper, S.C. & Wigley, T.M.L. 1986. Southern Hemisphere Surface Air Temperature Variations: 1851-1984. J. Climat. Appl. Meteorol. 25:1213-1230.

Jones, P.D. ; Kelly, P.M.; Goodess, C.M. & Karl, T. 1989. The Effect of Urban Warming on the Northern Hemisphere Temperature Average. J. Climate 2:285-290.

Jones, P.D. & Wigley, T.M.L. 1990 Global Warming Trends. Scientific American 8:84-91.

Sneyers, R. 1975. Sur L'Analyse Statistique de Séries d'Observations. OMM Note Tech. 143. 192 pp.

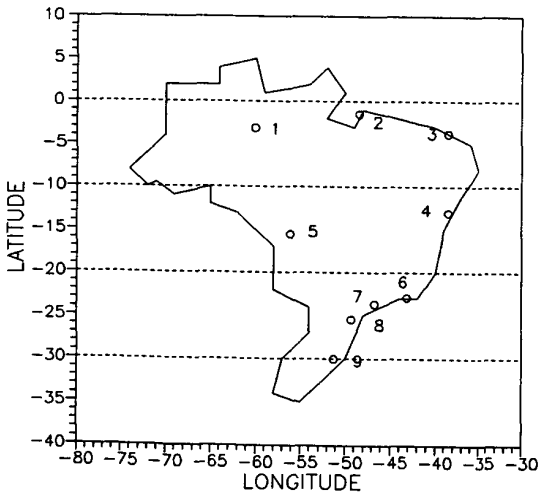


Fig. 1. Spatial Distribution of the Stations (The numbers are the same as Fig. 2)

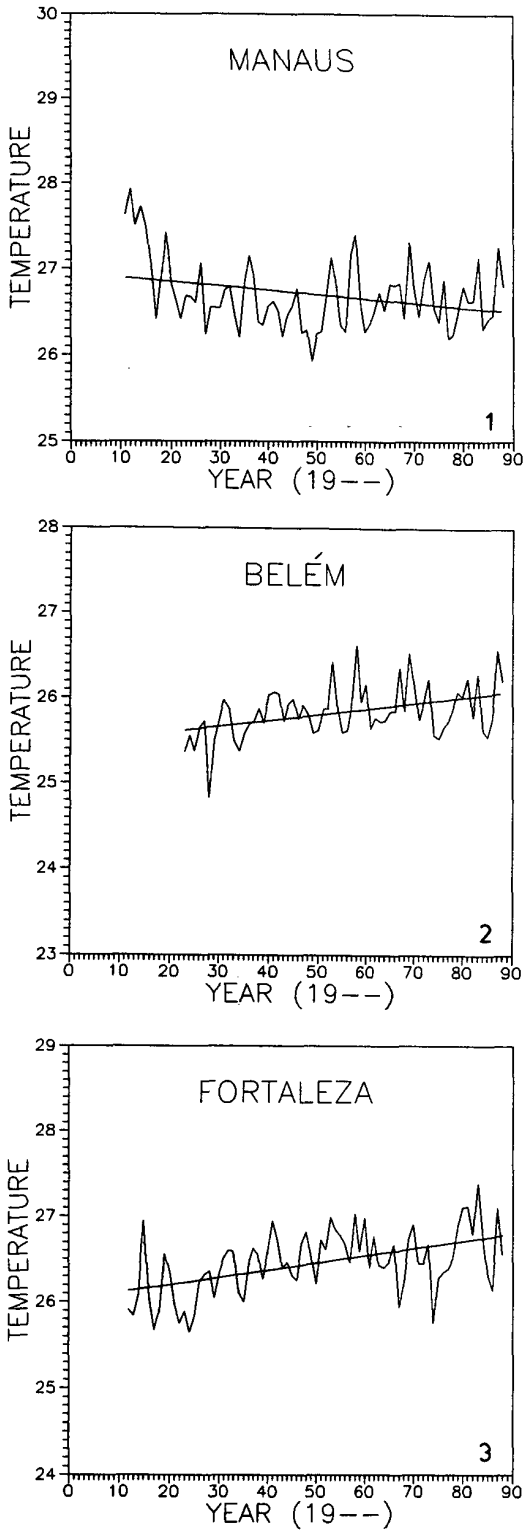
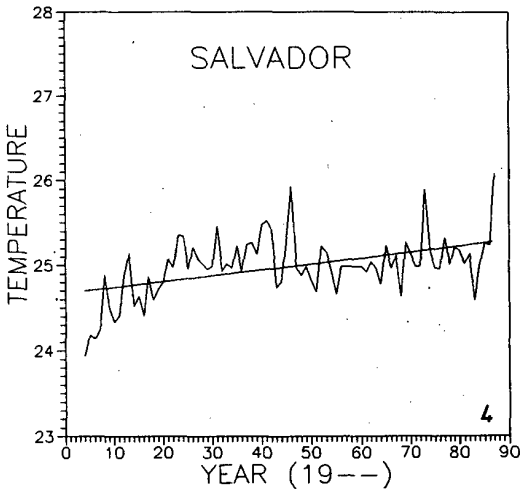
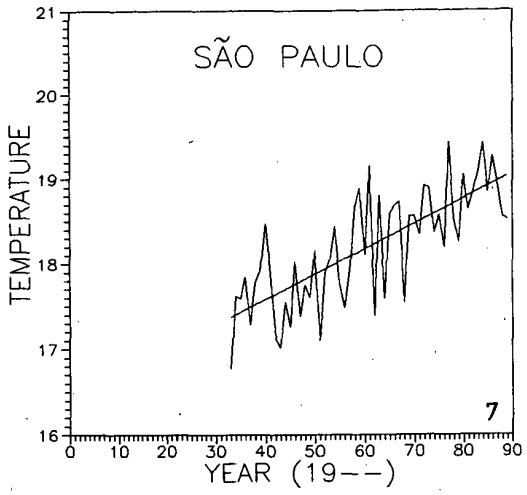
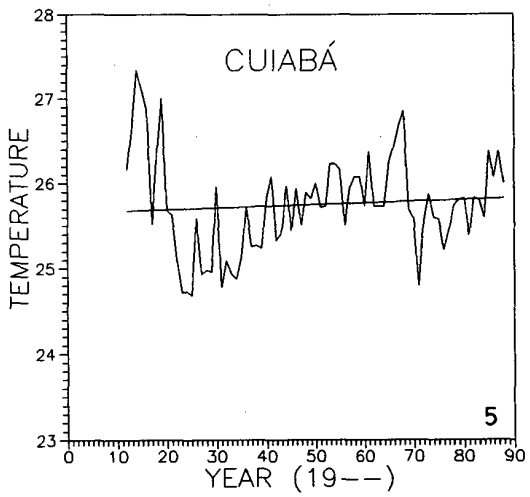
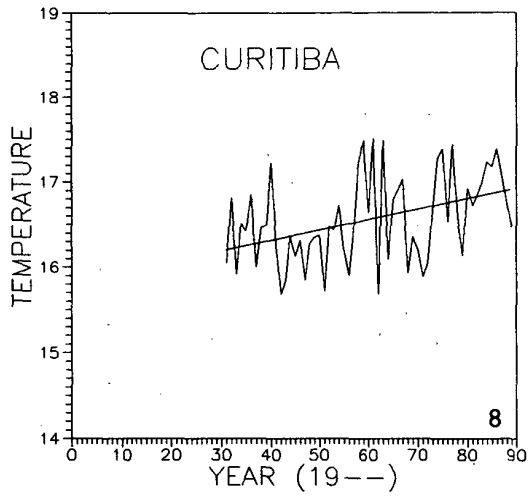


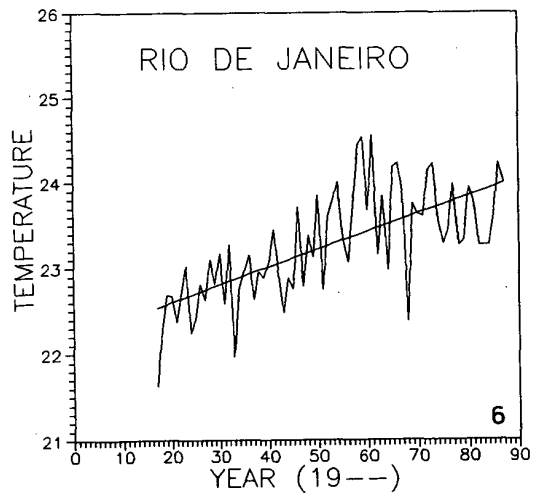
Fig. 2. Annual Average Time Series and their Linear Trends for the Selected Stations.











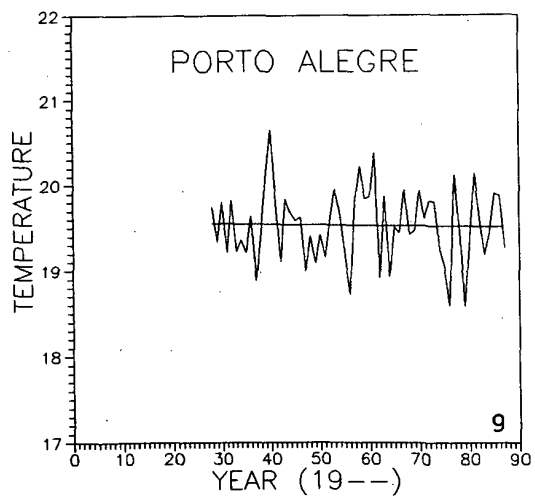


Fig. 2. Annual Average Time Series and their Linear Trends for the Selected Stations (continuation).

TRENDS OF TEMPERATURE AND ITS DAILY-RANGE IN CHINA

Yan Zhongwei
Institute of Atmospheric Physics, Academia Sinica
P.O.Box 2718, Beijing 100080 China

It is commonly held nowadays that the global warming trend during this century may be due to the enhanced atmospheric greenhouse effect (IPCC, 1990). But the widespread cooling during the 1960s and 1970s showed the reversing potential in the nature (Yan et al.1990). It is necessary to analyze the details of the regional temperature changes in order to improve the present knowledge of global warming and regional climate forecast.

Based on the analysis of some regional temperature changes, Karl (1990) found that the present warming trend was more due to the increase of daily minimum temperature. It is interesting to note that both the enhanced greenhouse effect and the metropolitanization affect also the daily minimum temperature more than the daily maximum temperature. However, the relevant regional changes need to be carefully explained. The author analyzed the series of daily temperatures of this century in different regions of China and found that the temperature change and the accomanying change of the daily temperature range in the subtropical south China was different from those in the mid-latitudinal north China, implying the possibility of two different kinds of temperature-controlling mechanisms.

Notes to the data

The primary data are the monthly mean daily maximum temperature (Tmax), daily minimum temperature (Tmin) and daily medium temperature (Tm), which were chosen from those of 60 stations of China prepared for the Sino-American cooperation research on climate (Koomanoff et al. 1988). All the stations missed few observations since 1950s. For the earlier period, only a few stations had reliable records. The author analyzed 3 regionalized temperature series back to early this century. The series of northeast China was the mean of those of 6 stations: Haerbin, Changchun, Shenyang, Dalian, Mudanjiang and Hailaer. The series of north China was formed with those of Beijing, Tianjin and Qingdao. The long-term series of Shanghai was used to represent the case of southeast China. These series would show us the differences of temperature changes since early this century along the different latitudes in the east part of China, from frigid-temperate zone to subtropics.

The author employed the daily medium temperature (Tm) rather than the daily mean temperature (T), due to the consideration that the observation system in China were often changed during this century. T was calculated as the average of as many as 24 hourly observations in a day for some years and as the average of 2 observations for some other years. Therefore, the homogeneity of T series was not good. Comparatively, Tm is calculated invariably as (Tmax + Tmin) / 2. Besides, some systematical differences between the Tm series and the corresponding T series of the periods when the data were relatively good were found. For 1950-1988, for instance, the 60-station-average series of (Tm - T) has a mean value of 0.5 °C with a standard deviation of 0.03 °C, implying the nearly-constant character of difference between T and Tm. Therefore, the trend of temperature change estimated based on Tm series will not be substantially different from that on T series.

The so-called winter and summer in this paper denote, respectively, the mean condition of December, January, Febrary and March and the mean condition of June, July, August and September. By comparing the seasonal temperature changes during the last 35 years, we will understand better the

difference of temperature change between the mid-latitudinal north and the subtropical south part of the country.

Features of temperature changes

The significant warming during this century in northeast China was demonstrated in Figure 1a. It was clear that the Tmin increased more than the Tmax did, thus leading to a decreasing daily temperature range. The linear increasing rate was estimated as 1.53 °C/century for Tm, 0.65°C/century for Tmax, 2.43°C/century for Tmin and -1.79°C/century for the daily range (Tmax - Tmin). The warming accompanied with the decreasing daily range might be explained to some extent by the theory of greenhouse effect. But the sharp decreasing rate of daily range that exceeded the warming rate in northeast China could hardly be attributed to the change of radiative force caused by the increased concentration of atmospheric carbon dioxide. The influences of the metropolitanization and some natural factors such as cloud changes should be also carefully assessed (personal communication with Zeng Zaomei, IAP, Beijing).

It was notable that all the 6 stations concerned missed records during 1941-1947. Fortunately, these few years were located in the middle of the series as shown in Fig. 1a, thus having no very large influence on the linear rate estimations which were derived by the least square method. Another kind of uncertainty might arise from the alterations of observation system. In China, a signal alteration happened in the early 1950s, when the 'standard observation system' was put into effect all around the country. Such an alteration should have induced a systematical deviation persisting all along since its beginning, if it had indeed significant influence on the series. However, as shown in Fig. 1a, there was a steady warming trend accompanied with a steady decreasing trend of daily range before the 1940s; after a few of strong yearly variations in 1950s and 1960s, an other steady but more acute trend of warming and daily range decreasing began. Such a featrue of change could hardly be due to the alteration of observation system.

The temperature change in Shanghai of southeast China was different from that in the northeast China. As shown in Fig. 1c, one of the notable features was that the daily range decreased abruptly in the early 1950s. This was considered as due to the adoption of thermometer screen to measure temperature since the early 1950s. Before that, the measurement was made in an open pavilion at Shanghai station. Here we saw the abrupt change induced by the alteration of observation system. At present, however, we have not yet excluded the possibility that the abrupt change around the early 1950s in Shanghai was partly due to the natural variability.

Let us avoid the complex history of Shanghai station for the time being. As Fig. 1c showed, the temperature and the daily range fluctuated with little trend since the middle 1950s. However, it was interesting to note the persisting warming trend accompanied by the increasing daily range before 1940s. It was clear that the earlier warming trend was more due to the increasing of Tmax. We investigated the cases of Taibei and Hongkong and found also that the Tmax behaved more variably than the Tmin did. It was implied that the temperature-controlling factor in the subtropics might be subtly different from that in the higher latitudes. The former affected more directly the daytime temperature while the latter affected more directly the nighttime temperature.

The case of north China appeared much meaningful. As Fig. 1b showed, the warming before 1940s was accompanied by the increasing daily range, which was similar to the case of Shanghai; later, especially after 1960s, the warming was accompanied with the decreasing daily range, which was similar to the case of northeast China. Such behaviors reflected the nature of the transitional zone between two regimes of climate changes.

In order to learn further the differences between the south and north part of China, we analyzed the temperature series of 1954-1988 of 60 stations, which are distributed somewhat evenly over the

country except the Tibetan region. The year of 1954 was chosen as the beginning of the series when the standard observation system began to be disseminated over the country.

Fig. 2 shows the courtrywide pattern of the trend of temperature change. It was clear that most of the north China warmed by 0.5°C during the 35 years. The Tibetan region seemed also much warmed, but this could hardly be discussed in the paper because there were few stations. The most of the south China experienced small cooling during the 35 years. The similar conclusions could also be reached in the papers of other authors (Chen,1990,for example). For comparing the detailed temperature changes between the warming and cooling regions, we averaged the data of all the warmimg stations into one data series and those of cooling stations into another data series. The linear trends of Tm,Tmax,Tmin and (Tmax - Tmin) in the warming region and the cooling region and their seasonal features were calculated respectively and the results were listed in Table 1.

Two important conclusions might be got from Table 1. The first was that the daily temperature range was decreasing in all the cases, with the decreasing rate about 0.1-0.4°C /decade. This was identical to Fig. 1. Again, the range of daily range decreasing was comparable with or even larger than that of temperature increasing, which could not be explained by the theory of greenhouse effect. Another conclusion reflected in Table 1 was that the warming in winter and the increasing of Tmin dominated the general warming feature of the warming region, while the cooling in summer and the decreasing of Tmax dominated the feature of the cooling region. It was implied that the temperature changes of north China were largely influenced by some factors more active during the cooler period (night and winter), while the temperature changes in south China were controlled by some factors more active during the warmer period (day and summer). Therefore, it was again suggested that the mechanism of temperature change in the subtropical south China might be different from that in the mid-latitudinal north China.

Remarks

The paper revealed some details of temperature changes in China during this century. Two results were worthy to be stressed.

a). The warming trend during the early half century could be traced in different latitudes of the east part of China. But the warming in the mid-latitudinal northeast China accompanied with a decreasing daily temperature range while the warming in the subtropical region accompanied with an increasing daily range.

b). The warming during 1954-1988 was found to be widespread over the most of north China, especially for winter and nighttime. Contrarily, the cooling in the same period in south China was particularly notable for summer and daytime.

The differences of temperature changes between the subtropics and the higher latitudes could hardly be explained by the popular theories such as those of greenhouse effect and metropolitanization. At any rate, it could be held that some other mechanisms had more directly controlled the temperature changes in the subtropical area. The author suggest that we should pay more attention to those phenomena that seem regional at the present and that go against the trend of the global warming.

References:

IPCC, Climate Change the IPCC Scientific Assessment, Edited by J.T. Houghton et al. Cambridge Univ.Press,1990.

Karl T., Proceedings of Beijing Intern. Symp. on Climatic Change, August 9-12 1990: A-15.

Koomanoff F., Ye T., Riches M.R., Zhao C., Wang W. and Tao S., Bull. Amer. Meteor.Soc., 61 (1988): 1301-1308.

Yan Zhongwei, Science Bulletin (in Chinese) 33(1988),19: 1487-1489.

Table 1. Linear trends of the Tm, Tmax, Tmin and (Tmax - Tmin) series 1954-1988 of the warming regions, which are located mainly in the mid-latitudes of the country, and of the cooling regions, which are located mainly in the subtropics. Unit: °C/35years.

Indice of	Warming region			Cooling region		
temperature	annual	winter	summer	annual	winter	summer
T_m	0.14	0.25	- 0.00	- 0.06	- 0.06	- 0.13
T_{max}	0.01	0.06	- 0.10	- 0.12	- 0.12	- 0.19
T_{min}	0.28	0.43	0.09	- 0.00	0.01	- 0.06
T_{max}-T_{min}	- 0.27	- 0.37	- 0.19	- 0.11	- 0.13	- 0.14

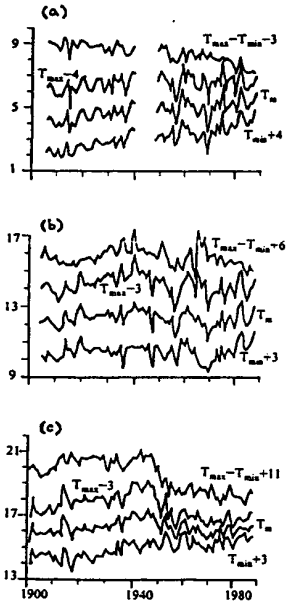


Fig. 1. Annual (Tmax -Tmin),Tmax, Tm and Tmin series, since early this century, of northeast (a), north (b) and southeast China (c). Unit: °C. The constants plus to each series are chosen for compactness of the figure.

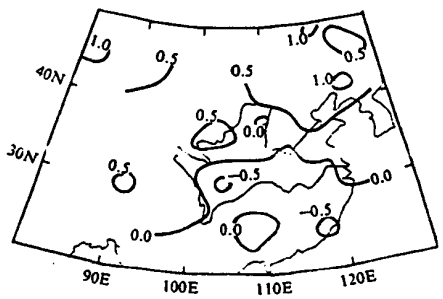


Fig.2. Geographical distribution of linear trend of the Tm series 1954-1988 over China. Unit: °C/35years.

THE CHANGES OF TEMPERATURE IN BEIJING DURING LAST 120 YEARS AND A ESTIMATION OF TEMPERATURE TREND IN 1991's

Xie Zhuang Cui Jiliang Li Hui
(Beijing Meteorological Bureau,China)

Since 1980's,it is found that the most obvious warming occours in winter and spring in North China. This warming increased water resources emergence.What would be the future temperature trend? This is what people want to know. Beijing is the capital of China. And Bejing's observation data is one of the longest period data in North China. In this paper,the features of temperatures during 1870-1989 and the influence of solar activity and occean on Beijing's temperature are studied.The temperature trend in 1990's is estimated with statistics model.

1.The features of long term temperature change:

i.The features of 10-Yr mean annual mean temperature:

Fig.1. is the 10-Yr mean annual mean temperature departures in Beijing,which shows there is a turning point in 1920's.Before it there is cold temperature trend,and then there is warming trend .And the temperature curve goes up and down except in 1870-1879 and 1880-1889.And there are three cold periods and three warm periods. 1880's and 1910's are two of stronger cold periods (minimun departure is -0.4c). The strongest warm period is in 1920's(maximun departure is 0.7c).And The second warm period is in 1920's(0.5c).

ii. The features of 10-Yr mean seasonal mean temperature:

Fig 2. gives the 10-Yr mean seasonal mean temperature departures during last 120 years. Generally the mean temperatures of spring and fall have the same amplitude as annual temperature.But in the winter mean temperature have two peaks, 1920's(0.6c)and 1980's(0.8c). Since 1950's,the departures of mean temperature in summer are below mean value and it is increasing to positive value in 1980's.

2.External influence factors:

The long term fluctuations of temperature during past 120 year is likely effected by some external factors. In this paper, we are interested in the relationship between Beijing's temperature and solar activity and occean. Cross spectrum analysis between Beijing's temperature and sunspot number (SSN) and SOI[**] shows there are high coherence in 2-5 years period and the lag time responsing lag phase of temperature against SSN or SOI is about 0.4-3.0 years.It means Beijing's temperature change could be influnced by solar activity and occean. Some results of calculation are shown in Table 1.

TABLE 1. CROSS SPECTRUM BETWEEN ANNUAl MEAN TEMPERATURE, AND SSN. and SOI.

variables	l	T	R12(1)	012(1)	L12(1)
T/SSN.y	14	4.4	0.90	-3.05	-2.08
T/SOI.w	21	2.9	0.59	-1.27	-0.58
T/SOI.s	17	2.5	0.63	-0.65	-0.36

in Tab.1. l refers to wave number; T refers to period,its unit is year; R12 refers to coherence spectrum, all R12 pass the critical value at 5% significance ; 012 refers to phase spectrum, its unit is arc, and all 012<0, which means the spectrum of Beijing's temperature is behind the spectrum of SSN and SOI; L12 refers to lag time responsing to lag phase.

The consequence of calculasion demonstrated the assumption that the temperature change is governed ,at least partly by solar activity and occean.

3. The estimation of trend_____persistency forecast:

Though the temperature in winter in 1980's is higher than in 1920's, the summer temperature in 1980's is far lower than in 1920's .The 1920's is the warmest period from 1780 through 1989.And the population and anthropogenic factors causing climate warming in 1920's is obviously less than in 1980's. A conclusion that the anomalous warm in 1980's is within the natural climate change can be deduced. And 1980's is the third one warm period.After the two former warm period ,the cold period occured. So lower temperature trend is expecting in 1990's.

4. Statistics forecasting model:

Generally a temperature series is a non-stationary time series.In this paper, "plus method model" is applied to deal with the series.A temperature series can be divided into two parts:

$Z(t)==D(t)+X(t)$ X(t) is a stationary Function

and $D(t)==f(t)+S(t)+F(t)$

D(t) is called tendency function. It consists of three parts, f(t) refers to major function,S(t) refers to periodic function and Fout(t) refers to external influence function. Usually stepwise regression is applied to estimate D(t). The factors of regression of D(t) come fromF(t),S(t),and Fout(t). The details are as following.

i.The regression estimation of tendency function D(t):

a. The estimation of major generating function f(t):

$fi(t)==ai+bit^i$ i=1,3 then f(t) ===> {fi(t),i=1,3}

ai,bi can be estimated by regression.

b. The estimation of periodic generating function S(t):

$S(t)== Sz(t)+Sp(t)$

here Sz(t) refers to integral periodic term,which is estimated by harmonic analysis.

then Sz(t) ===> {Sz,j(t),j=1,[n/2],n=120,j is integral number}.

and Sp(t) refers to non-integral periodic term, which is estimated by discrete spectrum.

then Sp(t)===>{Sp,k(t),k=2,[n/2],n=120,n/k is non-integral number}.

c.The external influence factors --- Fout(t):

In terms of 2 , there is high coherence between Beijing's temperature and SSN or SOI.Then Fout can be expressed as:

Fout(t) ===> {Fssn(t) ,Fsoi(t)}

Then f(t),S(t) and Fout(t) are seperately solved and taken as regression predictors.Further,by using of stepwise regression method,the estimation of tendency function D(t) could be derived,which is described below in some detail:

$$D(t) = Af\ fi(t) + Az\ Sz,j(t) + Ap\ Sp,k(t) + Aout\ F(t)$$

The tendency function of Beijing's annual mean temperature is calculated. The expression is as following:

$$\begin{aligned} D(t) = {} & 11.89 + 1.02Sp,2(t) + 1.00Sp,16(t) + 0.98Sp,18(t) + 0.91Sz,32(t) \\ & + 1.08Sp,10(t) + 0.99Sp,33(t) + 1.01Sp,38(t) + 0.86Sp,8(t) + 0.99Sp,50(t) \\ & + 1.02Sp,30(t) + 1.01Sp,45(t) + 1.02Sp,42(t) + 0.67Sf1 + 1.49Sp,5(t) \\ & + 1.00Sp,37(t) + 1.01Sp,47(t) + 1.00Sp,60(t) + 0.75Sp,12(t) \\ & + 0.0015Fssn(t-3) + 0.98Sp,29(t) + 0.96Sp,19(t) + 0.95Sz,19(t) \end{aligned} \quad (1)$$

here Sp,2(t) is the non-integral periodic term with wave number 2. Fssn(t-3) is the number of sunspot in three lag year. f1 is the first term of f(t) (i.e i=1). The expressions of seasonal temperature are not given here.

ii. The estimation of stationary series X(t):

In this paper AR(p) is used to simulate X(t), the maximal lag time is 15, and SRC (stepwise regression criterion) is applied. The expression of temperature of Beijing's can be written :

$$\begin{aligned} \text{Year} \quad & X(t) = 0.2013X(t-11)) \\ \text{winter} \quad & X(t) = -.2687X(t-1) - .3990X(t-4) - .1832X(t-7) - .2016X(t-13) \\ \text{summer} \quad & X(t) = 0.2192X(t-1) + .2190X(t-10) + .1841X(t-12) \end{aligned} \quad (2)$$

iii. The simulation and prediction of Beijing's temperature:

$$Z(t) = D(t) + X(t)$$

then (1) plus (2), and the simulation of temperature series is generated. The temperature of Beijing in 1986-1990 is taken to test the forecasting model, the results are in Table 2.

TABLE 2. COMPARISION OF SIMULATION AND OBSERVATION OF BEIJING'S ANNUAL TEMPERATURE IN 1986-1990

year	obs.	sim.	error
1986	12.1	12.0	0.1
1987	12.3	11.7	0.6
1988	12.7	11.9	0.8
1989	13.2	12.6	0.6
1990	12.7	12.5	0.2

From the Table 2., the consequence of simulation is satisfactory. So this model can be used to prediction temperature trend of 1990's. The forecasting values of tem. in 1990's are as following:

Table 3. THE FORECASTING VALUES OF Beijing'S TEM.

TERM	MEAN	91	92	93	94	95	96	97	98	99	2000
year	11.9	12.7	13.1	12.1	12.2	12.1	12.2	12.5	12.4	12.2	11.9
win.	-3.0	-2.9	-2.4	-1.9	-.6	-1.3	-3.0	-1.4	-3.2	-3.4	-2.3
sum.	25.2	25.6	25.0	25.0	24.9	25.6	25.8	26.2	25.8	25.5	26.1

Since then, the temperatures in 1991 are nearly the simulations normal. So it is hopeful that the prediction of temperature in 1992 would be all right!

[**] SOI come from " METEOROLOGICAL MONTHLY ", MAY 1989, Vol.15 , No.5 , Shi Wei Wang Shaowu "Southern Oscillation Index, 1857-1987".

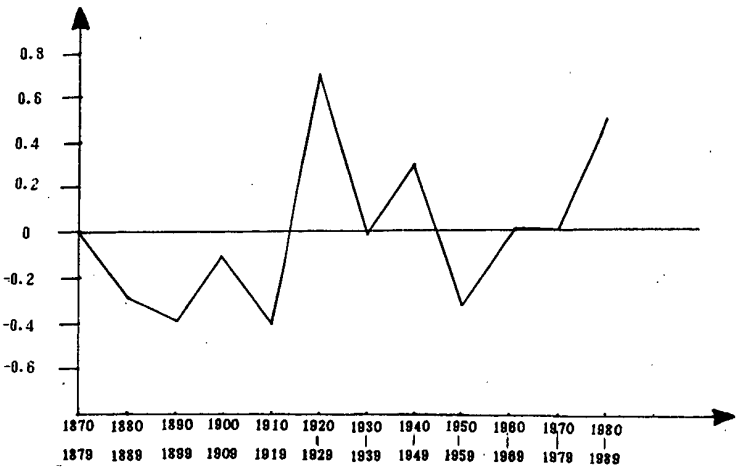


Fig.1. The 10-Yr mean annual mean temperature departures profile (1870-1989)

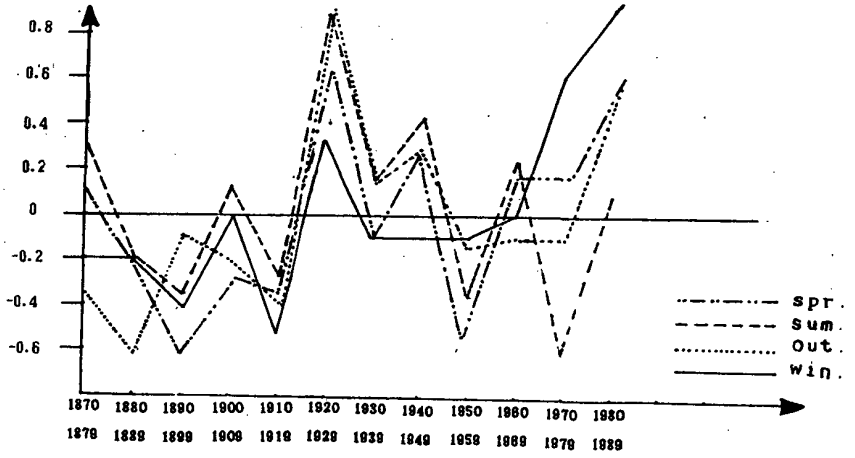
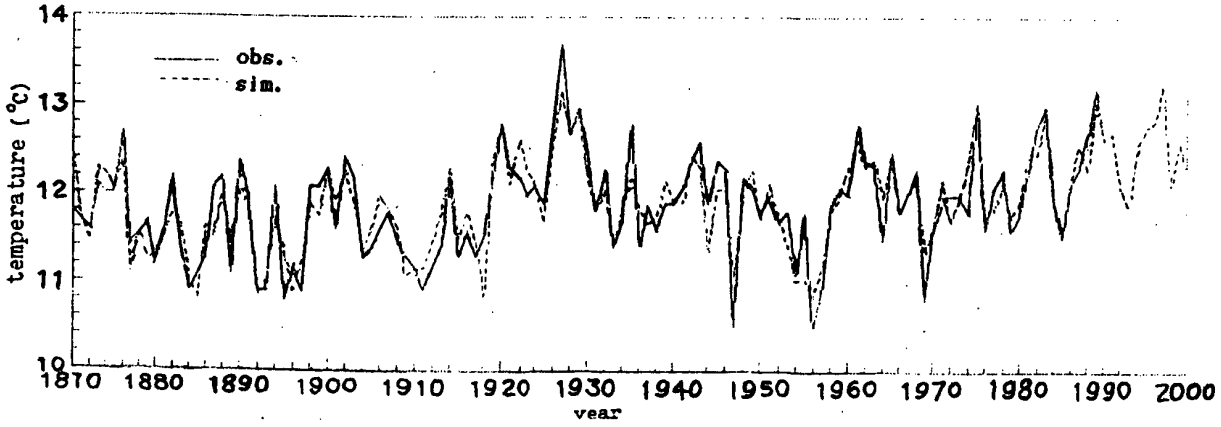


Fig.2. The 10-Yr mean seasonal mean temperature departures profiles(1870-1989)



SECULAR VARIATION OF THE AIR TEMPERATURE IN THE LONGITUDINAL AREA OF ROMANIA

Gheorghe Bâzâc, Roland Cotariu
National Institute of
Meteorology and Hydrology, Romania

The paper attempts to find the lows leading the long-term evolution of the mean temperature in Europa. The basic data contain the series of seasonal mean temperatures obtained from the monthly mean values published in "World Weather Records".

1. METHOD

It is known from the climatic processing practice that the mean values of the air temperatures show a statistical distribution which can be asimilated to the normal theoretical distribution (Gauss).

Previous studies revealed the fact that parameters of the normal distribution, the mean and the dispersion, can be estimated accurately enough if the data series are at least 40 years long. The meteorological stations used in this study were chosen according to this criterion.

As the differences in the physico-geographic conditions of the station make a comparison of the rough data impossible, it was considered as appropriate to use reduced values.

They have a close occurrence probability for the same time moment over large areas of Europe.

This sequence of the operations performed is the following.

1.1. For each values series, the mean $\bar{x}$ and the dispersion σ^2 were calculated:

$$\bar{x} = \frac{\sum x_i}{N}$$

were x_i = values of the series
N = the selection volume

$$\sigma^2 = \sum (x_i - \bar{x})^2 / N$$

1.2. The standard values were calculated:

$$y_i = \frac{x_i - \bar{x}}{\sigma}$$

thus eliminating the influence of the physico-geographic factors.

1.3. For every time moment, the reduced value field was interpolated and values in grid points were obtained with a 5 step, within the geographic space between 15 W longitude, 40 E longitude and 30 N latitude and 75 N latitude.

1.4. The values corresponding to the 25 E meridian, which crosses Romania, have been selected and then smoothed by averaging over 44 years - multiple of 11 years corresponding to the solar cycle.

1.5. The graphic representations were achieved on seasons (Fig. 1a , 1b, 2a, 2b).

2. CONCLUSIONS

2.1. The cold and warm intervals are alternated within an even variation in time, if the "small ice age" (Fernau stage) is taken into account.

The persistent warmings and cooling with amplitudes ranging between -0.6 to 0.6 are annouced by values areas disposed in train.

2.2. The general temperature tendencies increasing in the north and temperature decreasing in the south of Europe are within normal regime.

2.3. The more marked warming process during the last decades can be motivated by the contribution of the green-house gases, by the anthropic modifications upon the earth surface and last but not least, by the pollution of the planetary ocean surface, especially with oil as a film at the surface which modifies significantly the energetic exchange ocean-atmosphere.

REFERENCES

1. COTARIU, R., POPOVICI, C., GRUMAZESCU, C.

Statistic method for the computation of monthly mean air temperature and its application. Studii şi cercetări, Meteorologie 1979, Inst. de Meteorologie şi Hidrologie.

2. FRAKES, L.A.

Climate throughout geologic time, Elsevier Scientific Publishing Compani. Amsterdam - Oxford - New York 1979.

3. RUMSITSKI, L.Z.

Prelucrarea matematică a datelor experimentale (translated from russian). Ed. Tehnică Bucureşti 1974.

4. x x x

World Weather Records. U.S.Department of Commerce. National Oceanic and Atmospheric Administration (all series).

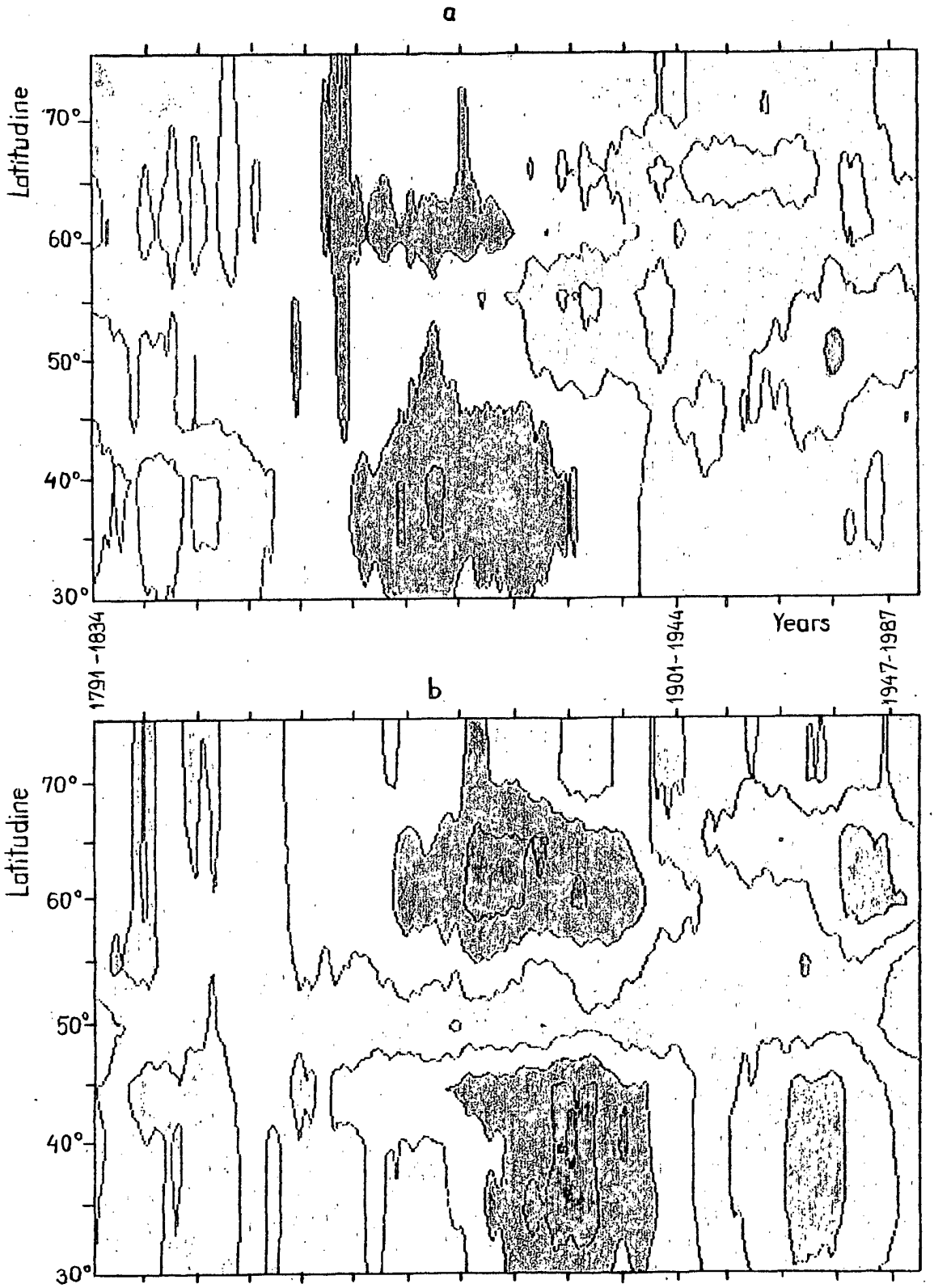
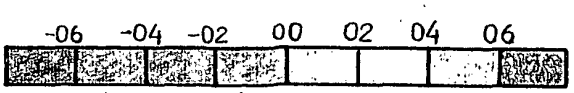


Fig. 1. Air temperature redused values variation
a. Spring b. autumn



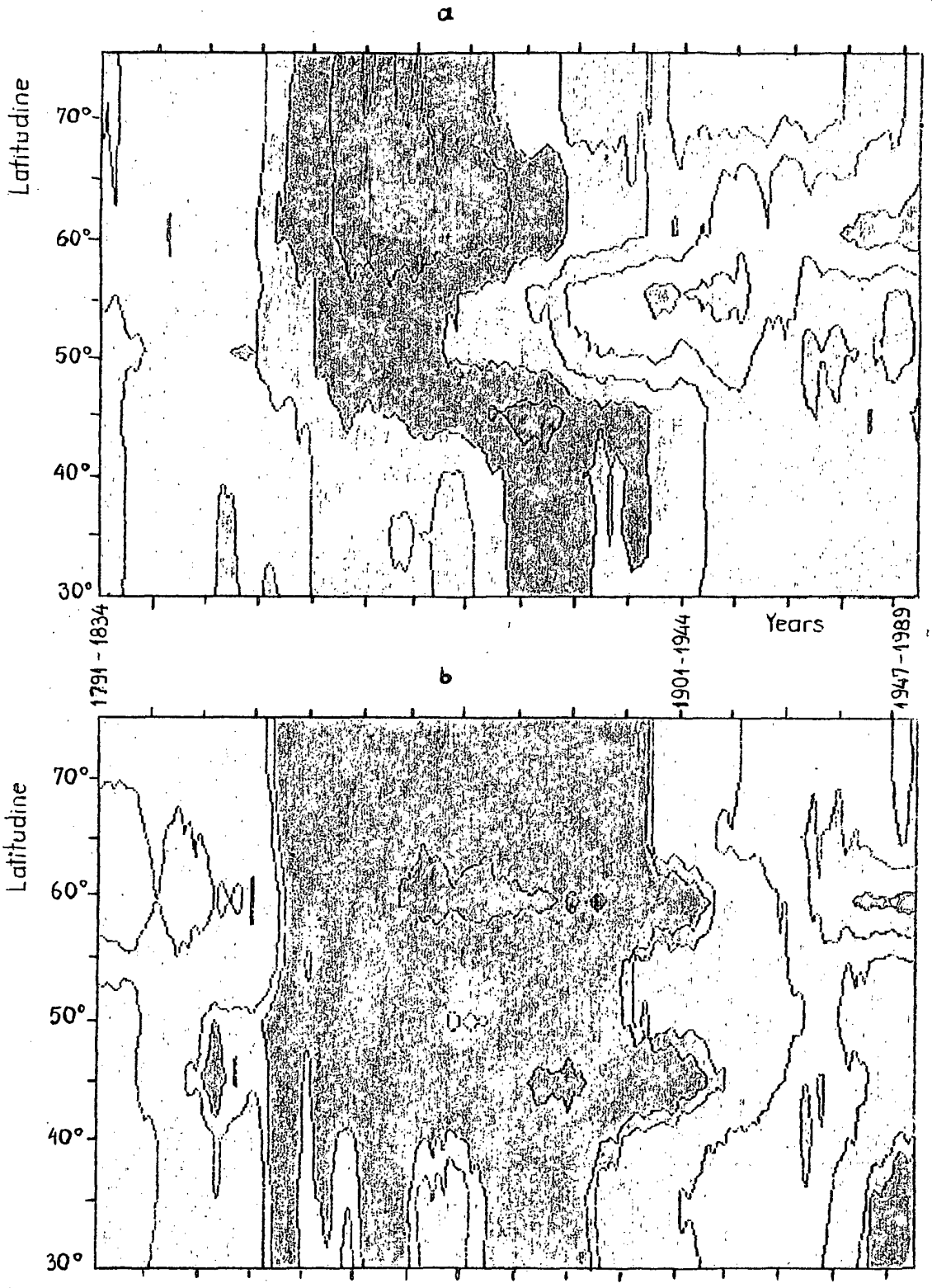


Fig. 2. Air temperature redused values variation
a. summer, b. Winter

Significant Increasing Greenhouse Effect on the Winter Temperature of the Northern China(Abstract)

Xu Qun
(Jiangsu Meteorological Institute, Nanjing 210008, China)
Yan Yiwen
(Central Meteorological Observatory, Beijing 100081, China)

The global temterature has risen significantly during 1980s, it is more markedly demonstrating in the mid- Siberia and norther China. The spatial distribution of winter temperature trend of recent 40 years (1951- 1990) indicates a general warming occurring in the northern China (north of 35°N) especially in the regions of Xinjiang, Inner- Mongolia and north-east China. This winter warming has reached a highest peak during recent 5 years (1986-1990) since the beginning of present century. The average temperatures of this 5 winters over the northern China have generally risen 1-4°C relative to the means of past 35 winters, the largest positive anomalies ($\geq$+3°C) appear in the northern parts of Xinjiang, Inner-Mongolia and north-east China locating just at the southern side of the winter warming center (mid-Siberia) of the 1980s. The time series of winter average temperature (Twn) of the whole northern China (north of 35°N) for recent 40 years has been given; through calculating the correlations between the Twn and preceding monthly SST fields of the north Pacific (50°N- 10°S), we found highly positive correlations concentrating on the larger portions of 50°N of north Pacific during preceding autumn (October/November). This bimonthly average SST of the region (50°N, 160°E-170°E, 160°W—135°W):SSTON(50°N) has a notable correlation (R=0.55, significant at 99.9% level) with the following Twn during 1955-1990 Yr., this means the abnormal rise of SSTON(50°N) will strength the zonal circulation of winter over Siberia weakening the southward invasion of arctic cold air from Siberia/North pole, hence the winter temperature of northern China will be risen.

Additionally, both the Twn and SSTON(50°N) have a significant increasing trend. The correlations between the Twn and the annual CO_2 content in Atmosphere with the latter leading of 0-25 years indicate that all

of them have larger values (0.44—0.53), but owing to the high auto-correlations of the time series of CO_2 content, the significance levels of these cross- correlations are markedly lowered by their lesser effective numbers of degrees of freedom (Chen,1982), only the correlations of the 8-9th and the 20-21th annual CO_2 content leading the Twn reach 95% level with the maximum (+0.53) occurring in the preceding 21th year of CO_2 content. This means that although the first factor of rising Twn is the direct effect from the positive anomalies of SSTON(50°N), but the Twn is still influenced by the enhanced CO_2 content of some preceding years even of about 20 years ago. We try to combine these two kinds of factors (preceding annual CO_2 content and the SSTON(50°N)) correlating with the Twn, the highest complex correlations reach to +0.60-0.61 during 1955- 1990 Yr. when the CO_2 content of the 8-10th and the 21th year leading Twn were selected separately.
It is notable that the increasing trend of CO_2 content also highly correlated to the SSTON(50°N) with the former leading of 0-25 years respectively (all significant at 95% level separately). These correlations have a trend enhancing as the year of CO_2 content leading farther except a few years only; so they reach the significance level of 99% during the preceding 15-21th year of CO_2 content with a maximum up to +0.645 when the CO_2 contnt of the 20th year leading the SSTON(50°N). Therefore the interannual variation of the Twn is influenced deeply by following two factors:

1. The preceding greenhouse effects with accumulative influences of about 20-21 years ago.

2. The direct influences of the SST anomalies of north Pacific (50°N) during preceding October/November, but a large fraction of them also reinforced by the preceding greenhouse effects which have stored in the north Pacific of about 20 years ago.

From abovementioned, it may be concluded that the most important factor causing the recent abnormal winter warming of northern China is the greenhouse effect with accumulative and increasing influences, additionally, it may be further enhanced by the warming of the north Pacific owing to the liberation of stored greenhouse effects in ocean itself of about 20 years ago. This possible climatic prospect is a matter worthy of our vigilance. Although the winter warming of northern China will

bring much benefits to agriculture, traffic and construction with the decrease of energy consume, but it may be unfavourable to Xinjiang, for her main water sources are the snow covers over high mountains; the decrease of accumulated snow on mountains by warming will be disasterous to the water supply of animal husbandry and agriculture of Xinjiang. So we should take effective measures early for this possible persistent and increasing winter warming of the northern China.

CONTRIBUTIONS TO THE STUDY OF CLIMATIC CHANGES AS OCCURRING AT SOME REPRESENTATIVE STATIONS IN ROMANIA

Constanţa Boroneanţ
National Institute of Meteorology and Hydrology, Bucharest-Romania

Norel Rîmbu
Department of Atmospherics, Faculty of Physics, Bucharest University

ABSTRACT

Statistical analysis of annual mean of air temperature time series at some representative stations on Romania's territory are performed. By using specific statistical tests such as the non-parametric Mann-Kendall and Pettitt tests, the possibilities of occurrence of a clearcut jump in the mean of this climatological variable are tested.
Their application allows to point out the abrupt character of the shift in the mean of annual air temperature time series at two stations in Romania.

1. INTRODUCTION

Several authors pointing out different aspects of regional(Sneyers, et al. 1990) or global (Demarée, 1990) climatic changes made use of statistic methods. The global warming of the Earth detected in the last decades of our century is still a controversial subject concerning the causes which have determined this effect.

This paper deals with a statistical analysis of annual mean of air temperature time series at some representative climatological stations on Romania's territory. Our aim was to detect the change in the mean and/or in the variance of this type of time series.

2. METHODS

We first calculated for each time series, the annual air temperature departure in ^{o}C, expressed as normalized variable, in standard deviations, from the overall mean.

Then, we analysed the stationarity of the records using the non-parametric Mann-Kendall test(Sneyers 1975). The statistic of the test is given by:

$t = \sum_{i=1}^{n} n_i$, where n_i represents for each element X_i, the number of elements X_j preceeding it ($j < i$) so that $X_j < X_i$.

Under the null hypothesis(no trend) the statistic is normally distributed with mean and variance given by:

$\bar{t} = E(t) = n(n-1)/4$

$var(t) = n(n-1)(2n+5)/72$

If α_o is the significant level of the test ($\alpha_o = 0.05$) the null hypothesis is accepted or rejected according to whether $\alpha_1 > or < \alpha_o$; α_1 is the calculated probability such as:

$\alpha_1 = P(| u | > | u(t) |)$ and $u(t) =$

$= (t-\bar{t}) / (var(t))^{1/2}$ is the normalized variable.

When values of u(t) leading to violation are significant, an increasing trend ($u(t) > 0$) or a decreasing trend ($u(t) < 0$) in the data is indicated.

The approximate time of occurrence of the trend is located at the intersection of the forward ($u(t)$) and backward ($u^{*}(t)$) curves of the statistic test when it occurs

within the confidence interval.

In order to localize more sharply the time of transition, the non-parametric Pettitt's change - point test (Pettitt, 1979) was applied to the same time series. For the test of H:no change against A: change, we use the statistic:

$$K_N = \max_{1 \le t < N} |U_{t,N}|$$

where the statistic $U_{t,N}$ is equivalent to a Mann-Whitney statistic for testing that the two samples $X_1,\ldots,X_t$ and $X_{t+1},\ldots,X_N$ come from the same population.

The maximum value of K_N corresponds to the change-point.

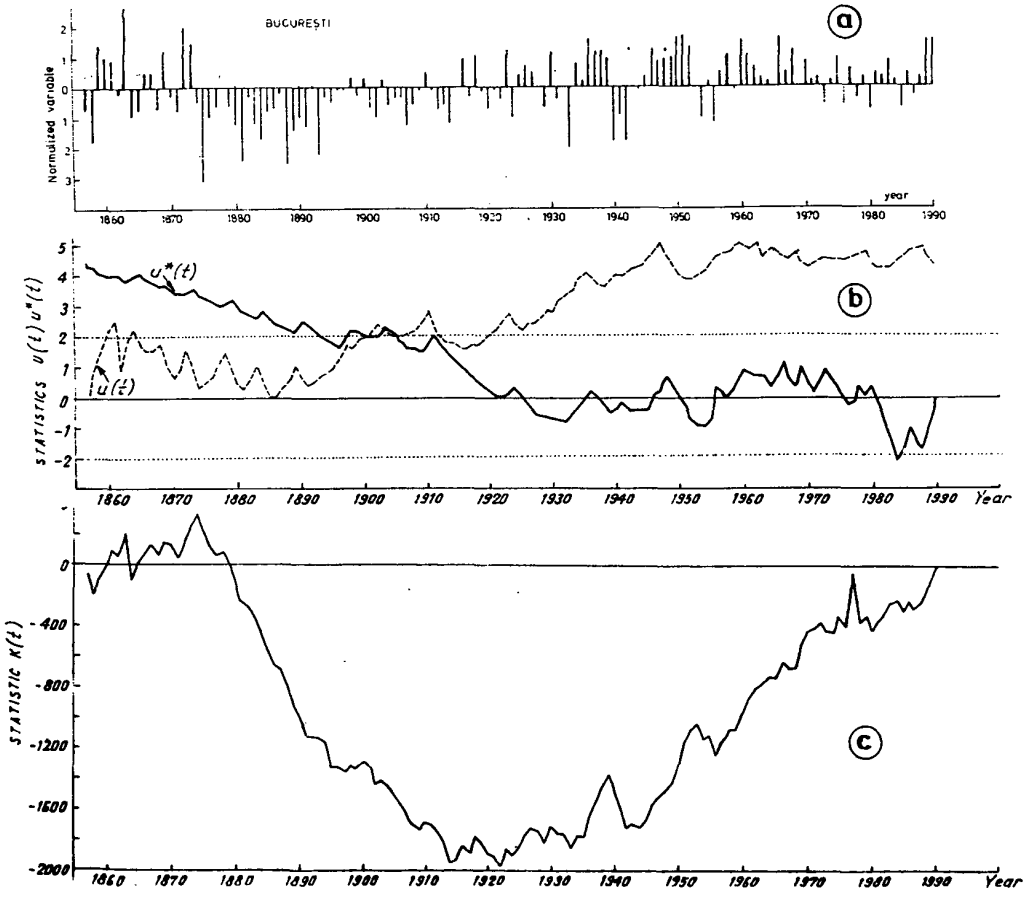


Fig.1 Bucureşti. Normalized air temperature annual departures in °C (a); Forward (u(t)) and backward (u*(t)) of the time evolution of the test statistics of the Mann-Kendall test(b); Time evolution of the test statistic of Pettitt's change-point test (c)

3. RESULTS

The techniques mentioned above have been applied to the air temperature annual means at five representative stations for different climatic regions in Romania such as, Bucureşti and Călăraşi for the south, Baia Mare and Ocna Sugatag for the north-west, and Timişoara for the west of the country.

Only Bucureşti and Baia Mare stations show temperature departures (represented as normalized variables expressed in standard deviations from the overall mean) prevalingly negative in the first

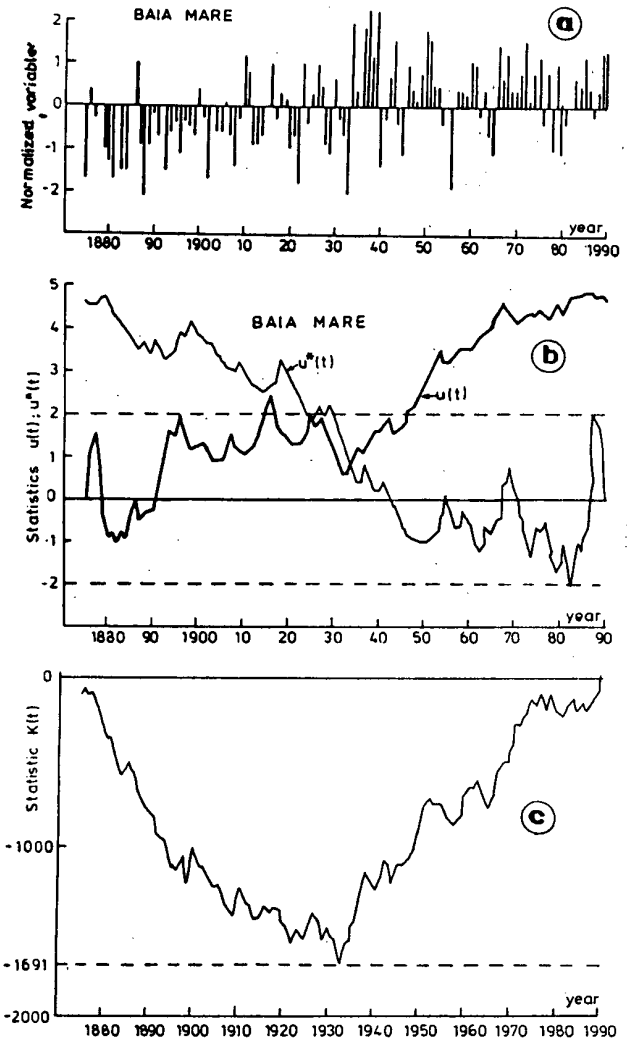


Fig.2 Baia Mare. The same significance as in Fig.1 a, b and c.

half and positive in the second half of the time series (Figs. 1 a and 2 a).

This grouping is better pointed out at Baia Mare.

The evolution of the statistics of the Mann-Kendall sequential trend test applied to the air temperature annual means indicate a highly significant upward trend, the intersection of the forward (u(t)) and backward (u*(t)) curves being within the 95% confidence interval for Bucureşti (Fig.1 b) and at the limit of this interval for Baia Mare (Fig. 2 b).

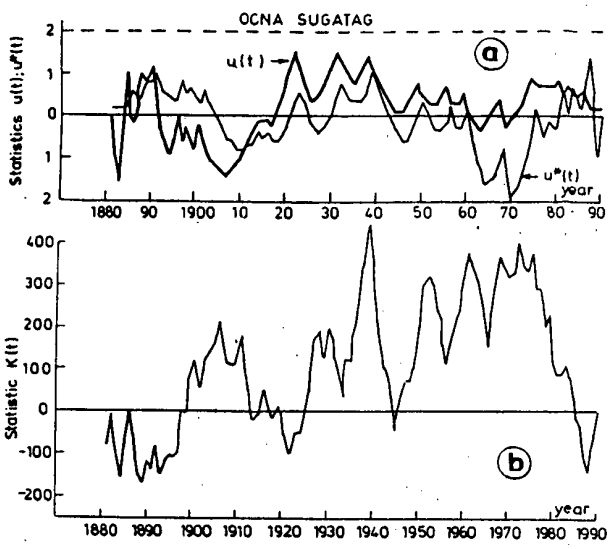


Fig.3 Ocna Sugatag. Forward (u(t)) and backward (u*(t)) of the time evolution of the test statistic of the Mann-Kendall test (a); time evolution of the test statistic of Pettitt's change-point test (b).

The statistic of Pettitt's test when applied to those data sets provides the timing of the change as well as the abruptness character of the transition.

The abrupt upward shift in the (mean (determined by the year for which the minimum of test statistic is reached) is located in 1922 - 1923 for Bucureşti (Fig.1 c) and 1932-1933 for Baia Mare (Fig. 2 c).

These techniques when applied to the air temperature annual means at other analysed stations do not emphasize changes in the time series mean, Fig.3 presents the statistics of Mann-Kendall and Pettitt's tests applied to the data set from Ocna Sugatag. This station is situated almost in the same climatic region as Baia Mare.

Analizing these statistics no trend in the data set can be emphasized.

A 0.5°C increasing in the air temperature annual mean calculated for two periods (up to and since the moment of the transition) is detected both at Bucureşti and Baia Mare.

4. CONCLUSIONS

The change in the air temperature annual mean pointed out at two stations in Romania is in accordance with the other climatic changes emphasized by other authors (Sneyers, et al. 1990 Demarée, 1990) in their paper.

Due to local causes and the insufficient knowledge on the physical mechanisms causing the warming process, this could not be revealed in the evolution of other similar time series analysed at the other stations in Romania.

REFERENCES

1. DEMAREE, G.R., (1990)
Did an abrupt global climatic warming occur in the 1920 s?
Inst.Roy.Météor. de Belgique, Pub.A., No.124, 32-37.

2. SNEYERS, R., (1975)
Sur l'analyse statistique des series d'observations, O.M.M. Note technique No.143, Genève, 192 p.

3. SNEYERS, R., VANDIEPENBEECK,M., DEMAREE, G.R., (1990)
Climatic Changes in Belgium as appearing from the homogenized series (1833-1988), Inst. Roy. Météor. de Belgique, Pub. A, No.124, 17-20.

4. PETTITT, A.N., (1979)
A non-parametric Approach to the Change-point problem, Appl. Statist. 28, No.2, 126-135.

REGIONAL CLIMATIC JUMPS AND POSSIBLE FACTORS DETERMINING THEM

Constantin Mareş, Ileana Mareş

National Institute of
Meteorology and Hydrology
Bucharest-Romania

ABSTRACT

The climatic jumps have been studied by means of the signal-to-noise ratio for the temperature field in Romania and the undulatory components (0-3) of the geopotential height at 500 hPa. Certain jumps have been emphasized at the regional level which differ from those at the hemispheric level, as well as the influence of the increase of CO2 concentration which for certain area, does not indicate a temperature increase.

1. INTRODUCTION

The work is done in the almost-intransitivity hypothesis of the climatic system (Lorenz, 1968) therefore, the transition from one regime to another is studied interannually. The climatic changes are emphasized by the signal-to-noise ratio the value of which is considered as 0.5 (Leith, 1973), as being of practical importance. The signal - to - noise ratio is defined by Yamamoto (1986), and the eventual tendencies are revealed by means of test Mann-Kendall (Sneyers 1975). The quantitative relationships between the various variables considered are attested by the use of the biserial and tetra - choric correlations.

Also the certain quasiperiodicities are emphasized by means of the power spectra based on FFT or on the maximum entropy method (Burg, 1975).

In order to emphasize certain climatic changes at the regional scale in the context of the climatic changes at the global scale the following elements were considered; the air temperature at the global and hemispheric level (1861-1988), at the regional level (1950-1991 and 1775 - 1988), the geopotential field at the 500 hPa level for the Northern Hemisphere (1947-1991), the Index of the Southern Oscillation (SOI) (1938 - 1990), CO2 at the global (1861-1988), and regional (1950-1988) level, the solar activity (1775-1988) and the ocean water temperature (the northern part of the Atlantic Ocean) (1875 - 1988).

In the final part of the work a probability change method of a somehow distribution in a Gaussian one is presented.

2. ANALYSES AT THE HEMISPHERIC AND REGIONAL LEVEL FOR EMPHASIZING THE CLIMATIC TRENDS AND JUMPS

By applying the Mann-Kendall test for the annual mean temperature in the Northern Hemisphere, it appeared that starting with 1922 the temperature tendency is increasing. At the regional level the situation is different, for example, for the Sibiu station (Romania), during the last decades the temperature tends to decrease. The regional differences are obvious also by considering the Prague station (Czekoslovakia). For the Prague station, the series being longer (since 1775) experiments could be done with different averaging lengths over intervals of a maximum of 30 years forward and backward as against the reference year, in order to emphasize the climatic jump. The results are given on Fig.1: (a) for the month of January and (b) for the annual means.

An interesting result is obtained in case of the month of January for which the 15 to 21 years averagings

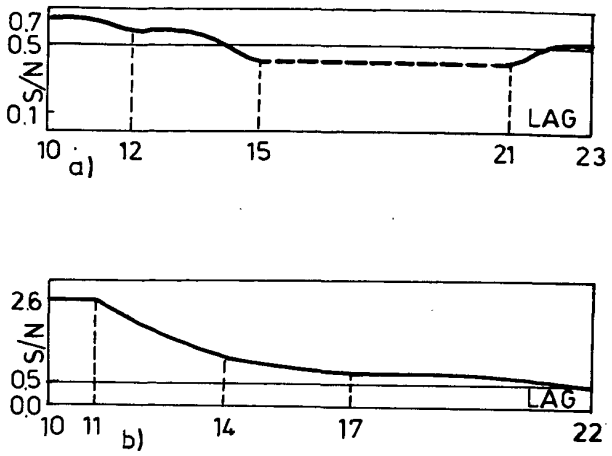


Fig.1. The signal-to-noise ratio in terms of the averaging interval for the air temperature at the Prague station: a) January; b) annual means.

lead to an unsignificant signal-to - noise ratio.

According to Madden (1976) and Leith (1973), it would mean that these averagings are not indicated for obtaining predictable signals.

On the contrary, for small lags, the annual averagings are more predictable than the monthly ones (the signal-to-noise ratio exceeds the value 2.5 (Fig.1 (b)).

For smaller lags the most significant jumps for the annual temperatures in Prague were recorded in 1789 (lag=10), the warming up being with 1.5°C and in 1976(lag=10) cooling with 3.1°C. The 1976 cooling is essential also in the temperature field of Romania represented by EOF-1. For Romania, the behaviour of the temperature field of the Sibiu station (with the longest observation series since 1851) was also analyzed.

Almost for all considered waves jumps were recorded during the years 1970-1972, with a decreasing sense in terms of the previous period at wave 2 and an increasing one at waves 1 and 3.

The analysis of the power spectra revealed various significant periods, so, at Sibiu the maximum corresponds to 26 years (Fig. 2) while in Prague to 9 years.

For the undulatory components of the geopotential field a period 2.6 is revealed for the annual means and for certain winter or summer months.

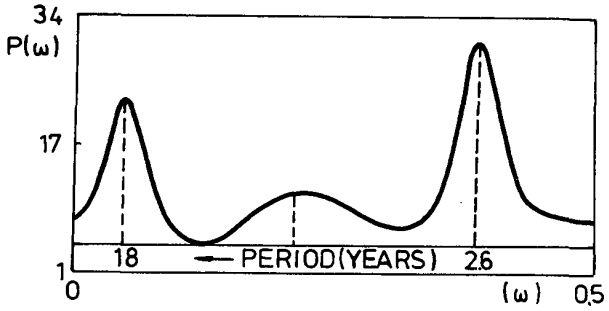


Fig. 2 The power spectrum of the annual mean temperature at Sibiu station

3. THE INFLUENCE OF CO2 CONCENTRATION ON THE REGIONAL CLIMATIC CHANGES

By using common (R),biserial(RB) (one of the variables becomes dichotomous) or tetrachoric (RT) (both variables become dichotomous) correlations, the dependent values of the temperature for the global and regional level on the CO2 concentration values were tested, taken from, Trend's 90 (Boden et al., 1990).

The results are given on Table 1.

From presented results it draws to the conclusion that the effect of the CO2 concentration increase on the temperature is different at the regional level as compared to the global one.

By testing other possible factors which might influence upon the climatic change (as the solar activity expressed by Wolf numbers) there appeared no significant results.

Out of the elements with high memory which might contribute to the increase of the predictability, Southern Oscillation Index (SOI) and the water temperature in the north of the Atlantic Ocean were considered.

These elements do not show a

Table 1. The correlation coefficients between CO2 and air temperature

Variables		Correlation coefficients			Significance level (%)
X	Y	R	RB	RT	
T-Global	CO2 Global	.68	.68	.67	0.1 (R, RB, RT)
T-E.N.	CO2 Global	.55	.55	.64	0.1 (R, RB, RT)
T-Lat. $(30^{o}-60^{o})$ E.N.	CO2 Global	-.06	-.32	-.35	5 (R, RB)
T-(EOF-1) Romania	CO2 Romania	-.33	-.31	-.08	5 (R, RB)
T-Sibiu	CO2 Global	-.44	-.33	-.25	1 (R, RB, RT)
T-Praga	CO2 Global	-.33	-.21	-.01	1 (R, RB)

4. PROBABILITY TRANSFORMATION

The meteorological phenomena have statistical distribution which deviate more or less from a Gaussian law.

Therefore, for the processing of the experimental data supposing the existence of a normal statistical distribution, the proper achievement of the probability changes is required.

A rigurous and general procedure will be presented hereafter.

Here is for study a phenomenon described by the probability law:

(f(x), x); with $x \in [a,b]$;
x - random variable;

The change to a new distribution (g(y), y) with $y \in [c,d]$, which is easily handled, (for example, a Gaussian distribution) is done by a change of variable $y = T(x)$, so that:

$I = \int_a^x f(x)\,dx = \int_c^y g(y)\,dy$ on condition that: $I = 1$ for $x=b$ and $y=d$.

The probability densities f and g can be adequate polynomial functions.

To come back to the distribution (in the physical space) characteristic to the particularly studied phenomenon is made by reverse transformation. For the monthly precipitation amounts in January for the Sibiu station the functions f and g were approximated by polynomes of order 4 and 6 respectively.

5. CONCLUSIONS

For the temperature field in Romania, climatic jump were emphasized in 1882 (warming) and in 1953 (cooling). The increase in the CO2 concentration does not lead to an increase of temperature in Romania.

For the undulatory components of the geopotential field, at 500 hPa (with wave number 0-3) during the interval 1950-1991, climatic jumps were revealed during 1970-1972, in the spring season.

REFERENCES

Boden, A.T., Kanciruk, P., and Farell, M., (1990)
Trend's 90. A Compedium of Data

on Global Change. Environmental Science Division, ORNL,Tennesse, U.S.A., August, 258 pp.

Burg, J.P., (1975)
Maximum entropy spectral analysis. Ph.D.Stanford University, 123 pp.

Leith, C.E., (1973)
The Standard Error of Time - Averaged Estimates of Climatic means. J.Appl.Meteor.12, 1066-1069.

Lorenz, F.N., (1968)
Climate Determinism. Met. Monogr. 8, no.10, 1-3.

Madden, R., (1976)
Estimates of the Natural Variability of Time-Averaged Sea-Level Pressure. Mon.Wea.Rev.104 942-952

Sneyers, Y., (1975)
Sur l'analyse statistique des series d'observations. O.M.M. Note Technique, No.143,Geneve, 192 pp.

Yamamoto, R.T., (1986)
An Analysis of Climatic Jump.J. Met.Soc., Japan, 64, 273-280.

EQUATORIAL TOTAL OZONE LINKED TO TROPOSPHERIC TEMPERATURE ANOMALIES AND ENSO

BY

L.N. NJAU,
Kenya Meteorological Department,
P.O. Box 30259,
NAIROBI.

1. INTRODUCTION

Ozone is of major importance in the study of the atmosphere and useful because of protecting life on the earth from the harmful ultra violet radiation. By virtue of its absorption of solar ultra violet radiation and resulting heating effect it largely determines the atmospheric temperature and general circulation, the rainfall and indeed the entire climate (WMO 1986). As observed by Biswas (1977), ozone is an inert tracer in the lower stratosphere and observations of its distribution provide considerable information on transport mechanisms in this region and also on transfer between the stratosphere (upper atmosphere) and troposphere (lower atmosphere).

The concentration of ozone is changing both in the stratosphere and the troposphere owing to man's activities. Ozone in the stratosphere is diminishing while ozone in the troposphere which acts as a greenhouse gas is on the increase.

The spatial pattern of ozone explains a high-latitude maxima with equatorial minima Fig 1. Studies by Muthama (1989) revealed marked seasonal ozone variation with minimum in January to February and Maximum in September to October. The seasonal fluctuation was attributed to both physical and chemical processes where the physical processes include the seasonal change of tropopause height due to revolution of the earth Fig 2. The studies concluded that the lowering of tropopause increases the depth of the lower stratosphere where we have the 90 per cent of the ozone concentration. The studies also attributed the ozone maximum to the cold seasons where less ultra violet radiation reaches the ozone layer and thus less photodissociation of ozone taking place resulting in maximum ozone amounts during cold season and vice versa during the hot season.

Chubachi (1984) carried out a study on effect of temperature on total ozone over Antarctica. However, in equatorial-tropical latitudes such studies have not been conducted. The present study is geared to a better understanding on the coupling between atmospheric total ozone and the tropospheric temperature anomalies which have been linked to climatic events such as El Nino/Southern Oscillation (ENSO) in Pacific Ocean, droughts and floods in Kenya, Njau (1988). However, because of the complexity of the stratosphere-troposphere interaction, more intensive studies will be necessary.

2. INSTRUMENTATION AND OBSERVATIONS

In Kenya, there is only one total ozone observing station located at Chiromo campus, University of Nairobi. The observatory has been making total ozone measurements since March, 1984 using the Dobson Spectrophotometer.

The Direct Sun (DS) observations were accorded the highest priority each day, followed by Zenith Blue (ZB) and then the Zenith Cloud (ZC). In general highest priority was accorded to observations closest to Local Apparent Noon (LAN). However, good ozone measurements are difficult to make and their interpretation is subject to many uncertainties (eg Long-term instrumental drift, improper observational procedures, discontinuities in data records, influence of volcanic aerosols and sulphur dioxide, as well as abrupt changes following the realignment and recalibration of instruments) and high natural variability (synoptic, seasonal, solar etc).

The Dobson Spectrophotometer is the mainstay instrument in the global total ozone observing network, upon which many other activities in the world action plan on the depletion of ozone layer depend to some degree. However, the instrument is delicate to operate and is subject to changes in its characteristics with time. The WMO is supporting the use of travelling Standard Lamps to check on the calibration of network Dobson instruments. The instrument used in the present study was calibrated twice using the travelling standard lamp. However, in order to obtain a better understanding of physical and dynamical behaviour of atmospheric ozone in the equatorial-tropical latitudes, extensive ozone observations should be carried out.

3. RESULTS

3.1 Seasonal Variation

Fig 3 shows on the average the seasonal variation of total ozone observed at Nairobi station with a minimum in January to February and maximum in September. The total ozone amounts shown in Fig.4 are representative of 1985 where the minimum occurred in February with a maximum in October. The seasonal characteristic feature therefore is a minimum in January to February and a maximum in September to October.

3.2 Year-to-year Variation

Fig.5 shows the year to year variation of the total ozone amounts for the month of September from 1984 to 1990. The year 1987 showed the highest total ozone amount while the minimum was observed in 1988.

Studies have attributed the two year cycle of total ozone to the Quasi-biennial Oscillation (QBO) which is a unique phenomenon in the tropical stratosphere. Studies have also revealed that the easterly phase of the QBO coincides with high ozone amounts and westerly phase with low ozone amounts.

3.3 Impact of Temperature Anomalies on Total Ozone and ENSO

Fig.6 shows the 300 hpa temperature anomalies in relation to the total ozone amounts for September from 1984 to 1990. The most interesting feature is the in-phase relationship of the two parameters. The sudden total ozone increase observed in 1987 and 1990 can now be attributed to the pronounced 300 hpa warm temperature anomalies. The total ozone minima observed in 1986, 1988 and 1989 can be attributed to the cold temperature anomalies observed during the three years.

The warm temperature anomalies have been associated with excessive March to May rainfall sometimes resulting in floods while the cold temperature anomalies have been associated with rainfall deficiency and droughts in the same season, Njau (1987, 1989).

The cold temperature anomalies have been associated with the onset of El Nino episode in east central Pacific Ocean in March/April. The El Nino episodes have been associated with Kenyan droughts experienced during March to May season while La Nina is usually associated with good seasonal rainfall performance. It can therefore be concluded that the El Nino events contribute to the total ozone deficiency.

In Kenya the positive Southern Oscillation indices contribute to good seasonal rainfall while negative indices are associated with poor or total failure of seasonal rainfall and high frequency of southwest Indian Ocean tropical cyclone activity, Njau et. al.(1989). The total ozone deficiency can indirectly be attributed to the negative Southern Oscillation indices.

The foregoing linkages are of great prognostic value towards the prediction of the atmosphere climate system and the total ozone variations in the equatorial-tropical latitudes.

4. CONCLUSION

Only limited studies on ozone trend over equatorial-tropical latitudes have been carried out due to insufficient data. The data base of equatorial-tropical latitudes ozone measurements with a data record of less than 20 years is not sufficient to explain natural variability in total ozone with confidence. The analysis of the available data and subsequent interpretation of the results require great care and understanding.

The present findings based on limited data have revealed that the total ozone decrease can be attributed to temperature decrease and vice versa in the upper troposphere. Further, the ENSO events can be associated with total ozone decrease over Eastern Equatorial Africa. However, significant changes in total ozone concentrations may have adverse ecological impacts which may consequently alter the climate.

It is clear from the foregoing that there is a need for further international effort to stimulate and co-ordinate extensive observational programmes to monitor future changes in total ozone in the equatorial-tropical latitudes.

REFERENCES

Biswas, A.K., 1977: The ozone layer. Proceedings of a meeting of experts designated by governments, inter-governmental and non-governmental organisations on the ozone layer, Washington D.C.

Bojkov, D., London J.B., Oltmans, S. and Kelly, J.F., 1976: Atlas of Global Distribution of Total Ozone July, 1957-July 1967. Tech. Note NCAR/TN/113 Nat. Centre for Atmos. Res. Boulder, Co.

Chubachi, S. 1984: Preliminary Result of Ozone observations at Syowa Station from February 1982 to January 1983. Proceedings of the sixth symposium on Polar Meteorology and Glaciology, Memoirs of National Institute of Polar Research Special Issue No.34, Tokyo.

Muthama. N.J. 1989: Ozone Fluctuations: The East Africa Situations. Proceedings First Annual National Research Workshop on Meteorological Applications and Services, 25-28 September, 1989, IMTR, Nairobi.

Njau, L.N. 1987: Seasonal Variability of Kenya Rainfall and its Teleconnections. Proceedings First Technical Conference on Meteorological Research in Eastern and Southern Africa, 6-9 January 1987, IMTR, Nairobi.

Njau, L.N. 1989 Towards Prediction of El Nino Kenya Droughts and Floods. Proceedings First Annual Naitonal Research Workshop on Meteorological Applications and Services, 25-28 Sept. 1989 IMTR, Nairobi.

Njau, L.N. and Odak, J., 1989: Influence of ENSO on the South West Indian Ocean Tropical Cyclone Activity and its impact on Kenya Rainfall. Proceedings First Annual National Research Workshop on Meteorological Applications and Services 25-28 Sept. 1989, IMTR, Nairobi.

WMO, 1986: Atmospheric Ozone, 1985. Assessment of our understanding of the processes controlling its present distribution and change. World Meteorological Organization Global Research and Monitoring Project Report No.16, WMO.

FIGURE 1

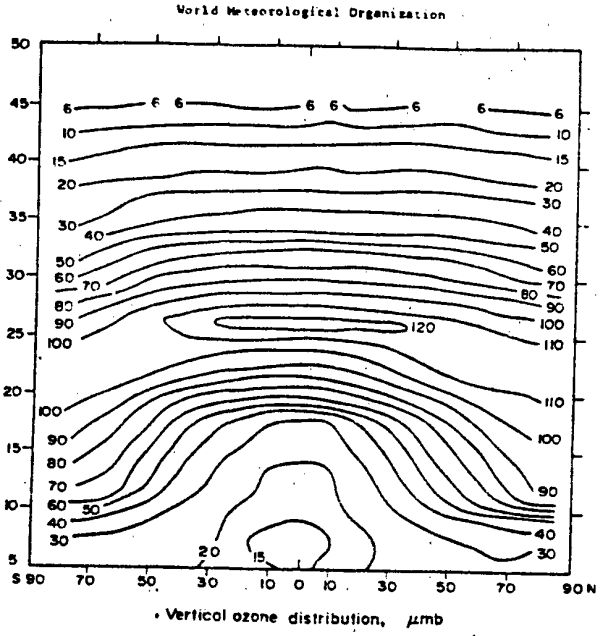


Fig. 1. Mean variation of vertical ozone distribution (partial pressure in μmb) as a function of latitude based on 8,500 Umkehr profiles and ozone soundings taken between 1956 and 1966. (Bojkov)

FIGURE 2

Fig. 2. SEASONAL VARIATION OF OZONE AND TROPOPAUSE HEIGHT

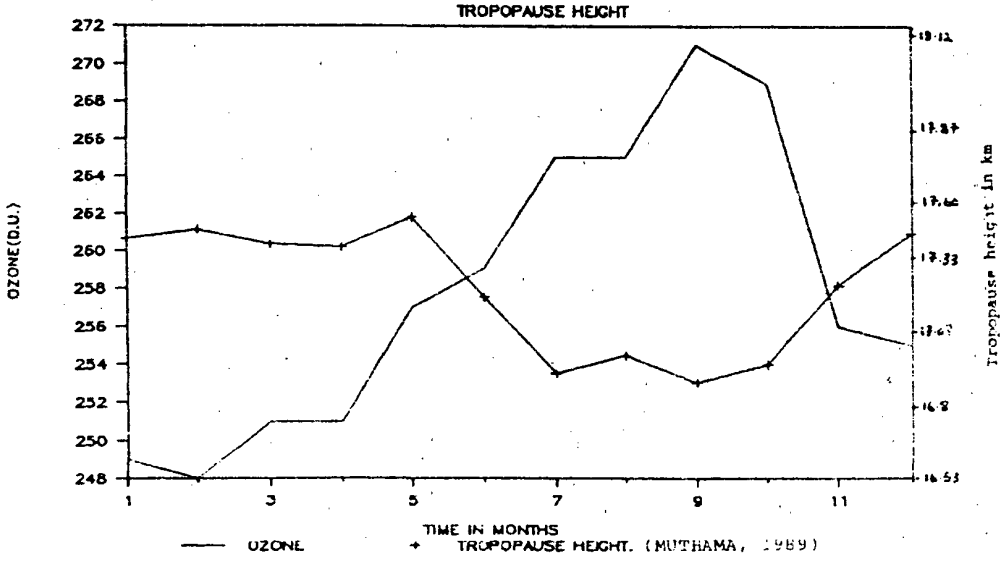


FIGURE 3

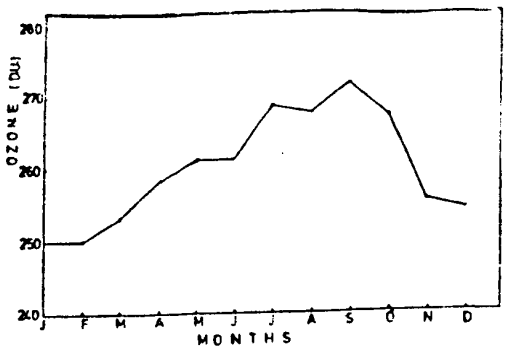


Fig.3: Seasonal Variation of Total Ozone.

FIGURE 4

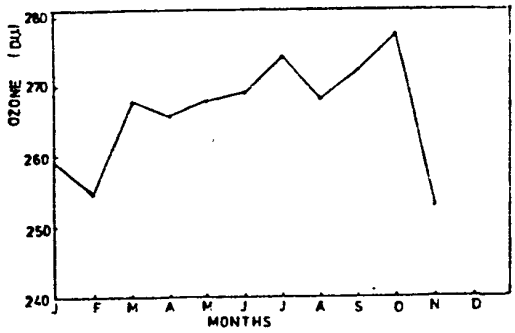
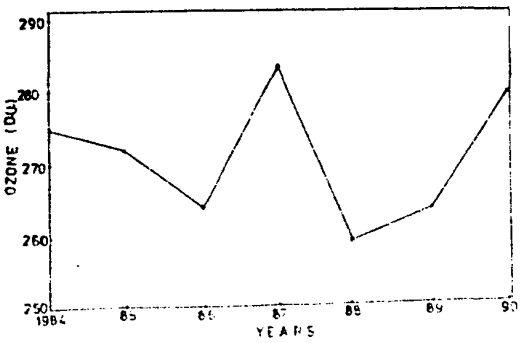


Fig. 4: Variation of Total Ozone during 1985

FIGURE 5



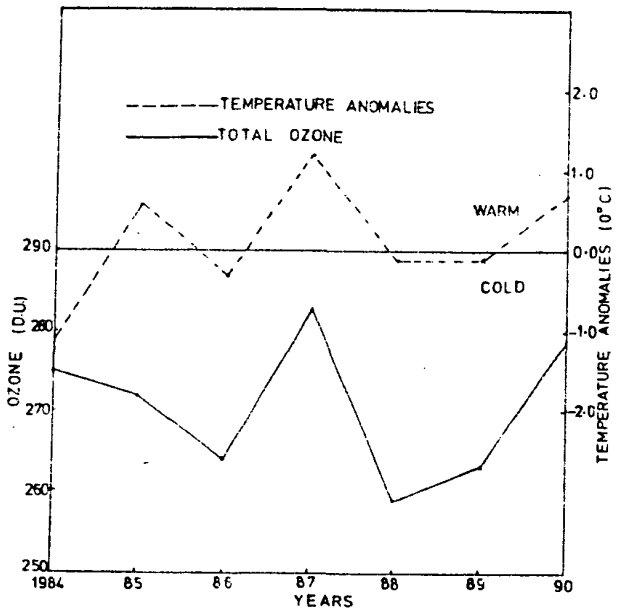


FIGURE 6

Fig. 6: Variation of Total Ozone in relation to Upper Troposphere Temperature Anomalies.

STUDYING THE NORTH AMERICAN PRECIPITATION CHANGES DURING THE LAST 100 YEARS

Pavel. Ya. Groisman

State Hydrological Institute
St.-Petersburg, Russia
Current address:
National Climatic Data Center
Federal Building, Asheville, NC 28801, USA

Introduction

During the last century an anthropogenic rise in carbon dioxide and other greenhouse gas concentrations in the atmosphere has occurred (IPCC 1990). The inevitable continuation of this rise in the nearest decades will result in global warming (Budyko and Izrael 1987; IPCC 1990). Actually, the global average surface air temperature has increased to about half a degree during the last hundred years (IPCC 1990). Related with these warming, changes in precipitation (if they can be found) become of special interest; they can be extrapolated into the future. Keeping this goal in mind, we studied the North American precipitation and snowfall data.

To find the possible relationships between precipitation and global temperature, it is necessary to filter from the time and space scales the "weather noise." The simplest such filters are provided by averaging procedures. That is why a year was selected as a time scale for precipitation and snowfall data, and the station data were spatially averaged over large regions up to 15 degree latitudinal zones. Doing so, we not only diminished the noise, but revealed the large-scale features of precipitation changes on the decadal and century time-scales.

Data problems

Precipitation measurements are very sensitive to changes in the nearest environs of the gauge, in its type, and in the methods of measurement. Sometimes, a small move of the gauge results in twofold change of values of measured precipitation (Sevruk, 1982, Karl et al. 1992a). After the spatial averaging, it can be expected that the random errors connected with such shifts will be decreased and compensate each other. However, this is not so when systematic changes in the measurement procedures occur countrywide. Several such changes have occurred during the last century in the Canadian and United States' meteorological primary networks (Fig. 1). The area-averaged precipitation time series are especially sensitive to such large-scale inhomogeneities because their noise component has been decreased by averaging. So, studying the trends in these time series, we have to be sure that the observed large-scale decadal changes represent the climatic changes and are not affected by improvements of precipitation measurements introduced into the network during the last century by national weather services.

The last data problem to be addressed is a scale problem. There are several reasons for precipitation undercatch by the standard gauges installed worldwide (WWB 1974, Sevruk 1982). Scale adjustments applied to precipitation data to obtain unbiased values for the needs of hydrology and climate modelers sometimes increase the areal totals by 50%, and winter precipitation twofold (WWB 1974). To incorporate these adjustments into climate change studies a special procedure was evaluated for annual precipitation by Groisman et al. (1991).

Data used and their processing

Two different data sets were used in our analysis. For the entire North America, the annual snowfall and total precipitation data for about 1350 U.S. and Canadian stations were used for the period 1950-1990 (for Canada up to June 1990). For the century-scale precipitation change studies, data were selected from 189 Canadian stations (mostly in the southern part of the country) and 494 United States' stations for the period 1900-1990. The stations have been chosen based on their length of records. For the U.S. they were selected from the subset of the best HCN (Historical Climatology Network) stations (Karl et al. 1990).

Century-long snowfall time series were available for analysis only for southern Canada. In the zone to the north of 55°N, the meteorological network before World War II was too sparse for any spatial

averaging and analysis. In the U.S., snowfall data exist in computer form for primary network only since 1948.

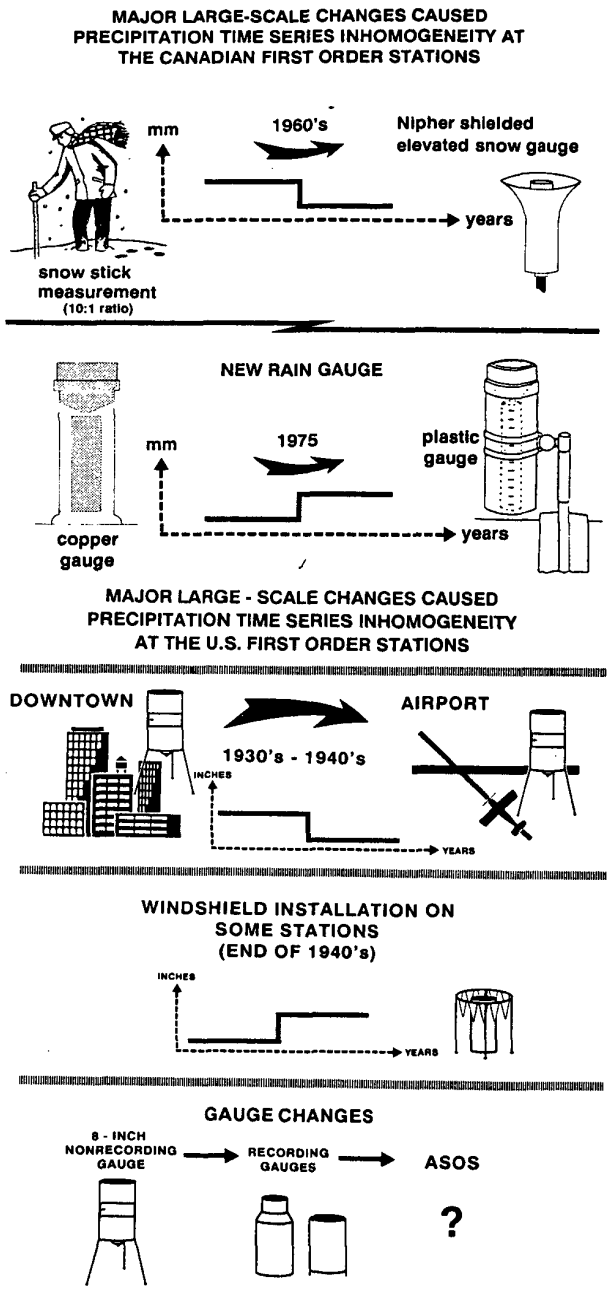


Fig.1. Sketch of the history of precipitation measurements in North America related to the precipitation time series homogeneity.

Prior to spatial averaging, the precipitation time series were homogenized (if necessary or if possible) with the help of metadata. To avoid the influence of missing data on the results of averaging, the time series were transformed into anomalies from the reference period 1951-1990 (for Canada 1971-1989). After spatial averaging, the scale corrections were applied to the area-mean annual precipitation time series to obtain the unbiased estimates of area-mean precipitation changes as described by Groisman et al. (1991).

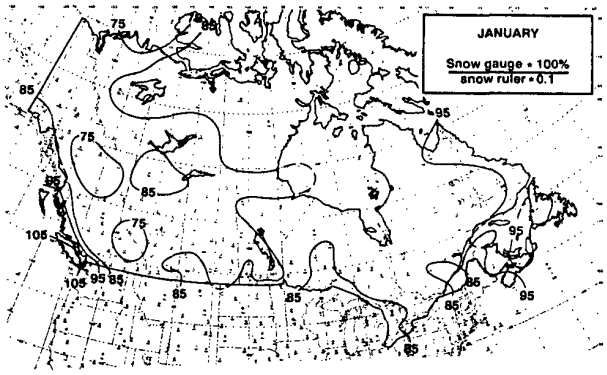


Fig.2. Mean ratio of two estimates of the water equivalent of snowfall in Canada made by Nipher shielded snow gauge and ruler. January.

In Canada the rainfall and snowfall are measured separately. Before the 1960s the snowfall was measured by snow stick only. Since then the primary network snowfall is measured by elevated snow gauge and snow stick (ruler). The presence of thirty years of parallel observations by two methods at 335 stations provided us the opportunity to evaluate and map the transfer coefficients between old and new types of snow measurements. Usually the monthly values of both snowfall measurements are closely correlated (r=0.93-0.98), but for individual storms they may differ significantly (Goodison et al. 1981). When the ruler measurements of fresh snow are transformed to water equivalent, a ratio 10:1 is used. However, these ratios differ from the water equivalent of solid precipitation measured by elevated snow gauge (Fig.2). In Karl et al. (1992b) it is shown that the above-mentioned ratio is an appropriate estimate of water content of snow measured by snow gauge only when the mean monthly surface air temperature is above 0°C on plains and above 5°C in mountains. So, we used the monthly maps (similar to January depicted in Fig.2) to infer the homogeneity of solid precipitation measurements at the primary Canadian network. To verify the accuracy of these corrections, the relationship between annual totals of water equivalent of solid precipitation (S) and mean annual maximum temperature (T) was used for southern (south of 55 °N) Canada (Fig.3). In Table 1, parameters of the regression model:

$$S(t) = a*T(t) - a*\bar{T} + \bar{S} + \text{error}(t), \quad (1)$$

where t = time, are shown as estimated using the data for two thirty-year-long periods: 1930-1959 and 1960-1989.

Obviously, this relationship was not disturbed by adjustments introduced into the data before 1960 and the difference in two $\bar{S}$-estimates is not statistically significant.

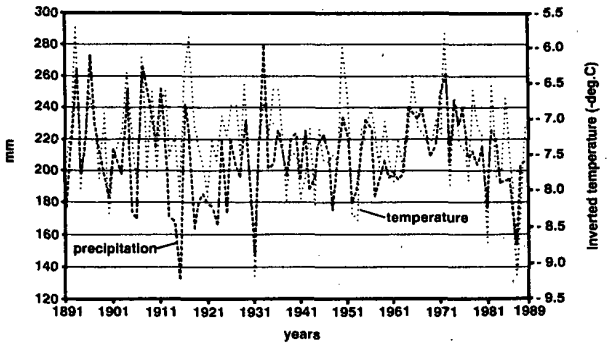


Fig.3. Water equivalent of annual snowfall adjusted to readings of Canadian Nipher shielded gauge and mean annual maximum temperature area-averaged over Canada south to 55°N.

Parameters of the equation (1)	Period	
	1930-59	1960-89
"a"±its s.d.(cm/°C)	-25 ± 5	-25 ± 5
" $a*\bar{T}$ " (cm)	-187	-187
"$\bar{S}$" (cm)	208	214
s.d. of error (cm)	18	16
R squared	0.46	0.53
Degrees of freedom	28	

Table 1. Estimates of parameters of equation (1) describing the relationship between mean annual water equivalent of snowfall and maximum temperature in southern Canada.

As another option to provide homogeneous Canadian solid precipitation time series, only snow stick measurements were used. These measurements are inaccurate; depend on temperature, humidity, time between snowfall cessation and its measurement, and so on (Goodison et al. 1981). However, this method of snowfall measurement has not changed during the century, and that is why it is of special interest for climate change studies. Both approaches do not change the conclusions about the trends in Canadian solid precipitation and the area-mean snowfall time series differ only by scale. On the contrary, if the uncorrected data were used, the slope of the trend in north Canadian solid precipitation would be twofold higher and hence misleading.

When in the 1970s the Canadian meteorological service installed a new rain gauge, only a few intercomparisons between new and old gauges were made (Goodison and Louie 1986). Their results suggest that the new gauge measures 1 to 4% more rain than the old one. The exact dates of the new rain gauge installation are not available for us. Therefore, a rough correction factor 1.025 was introduced into all Canadian rainfall data prior to 1975, to provide the proxy homogeneity of the data obtained by old and new gauges.

Even today the meteorological network in north Canada is relatively sparse. So, to analyze the changes over this zone, we have to estimate beforehand the representativeness of the network used. The spatial correlation function for north Canadian snowfall can be approximated by the relationship: $r(R) = C*\exp(-R/R_o)$, where R = distance (in km); R_o = 710 km is radius of correlation; C=0.84, and a value $(1-C)/C = 0.19$ is a measure of error of measurements together with micro-scale snowfall variability near the stations (Kagan 1979).

For the network selected the mean square error of spatial averaging of annual snowfall over northern Canada is estimated as 17% of the theoretical variance of the area-averaged snowfall time series in the first half of 1950s, and during the period 1956-1989 it was about 10%. The number and spatial distribution of snowfall-measuring stations for southern Canada and the U.S. during the last four decades is adequate, and there is no need to verify here the representativeness of the area-mean values.

Results

Fig.4 shows the zonally averaged snowfall over northern Canada (north of 55°N) and over the zone 45-55°N for the last four decades. The accuracy of area-averaging of the data in northern Canada discussed above is high enough to confirm the strong positive linear trend in annual snowfall revealed in the zone (19% per four decades). This trend jointly with trend in rainfall data (not shown) confirm the systematic increase in annual totals of precipitation during the entire period of instrumental observations. Snowfall in the zone 45-55°N is closely connected with the temperature changes on the continental and hemispheric scale. So, the correlation between this time series and the annual mean surface air temperature over the Northern Hemisphere is about -0.7 (Groisman et al. 1992). Using this relationship and taking into account the intrinsic coherence of annual snowfall totals in southern Canada (that encompassed the main part of the 45-55°N zone) with regional maximum temperature (Fig.3) we can foresee a future decrease in snowfall for this zone accompanying the global warming projected by IPCC (1990).

Fig.5 shows century-long unbiased time series of the annual precipitation area-averaged over southern Canada (to south of

55°N) and the contiguous U.S. They were calculated by area-averaging the data of 161 and 494 stations respectively by the Thiessen (1911) method. The Canadian time series were homogenized before the averaging. There were no corrections applied to the HCN stations data. Scale corrections based on WWB (1974) maps were applied after spatial averaging to obtain the unbiased precipitation time series shown in Fig.5. These corrections increased the U.S. annual totals by 20% and Canadian ones by 30%. The biases in area-mean values for the mountainous West of the continent were the main cause for these corrections. Some statistics of these two time series are presented in Table 2. The century-scale positive linear trend in annual precipitation is statistically significant for southern Canada (0.9% per decade) but not significant for the U.S. because the variability of the U.S. precipitation has a strong high-frequency component with periods below five years (above 60% of variance).

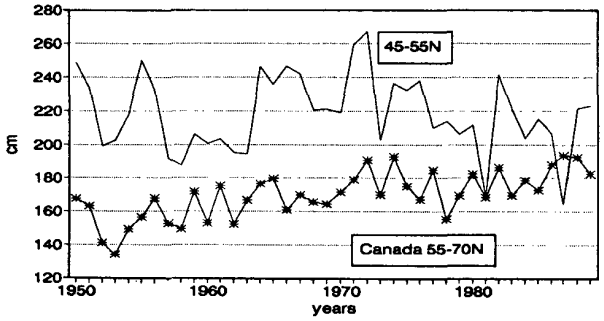


Fig.4. Annual snowfall (ruler measurements) area-averaged over two North American latitudinal zones - 45-55°N and 55-70°N (Alaska excluded).

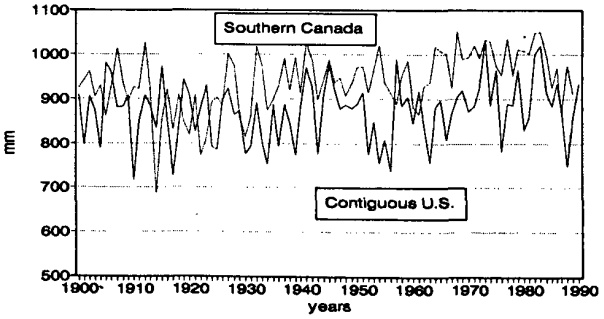


Fig.5. Unbiased estimates of the U.S. - southern Canada (south to 55°N) annual precipitation changes during 20th century.

Region	Mean and s.d., in mm		Linear trend & its error in %/decade	
S.Canada	940	66	0.95	0.20
Cont.U.S.	870	68	0.45	0.31

Table 2. Statistical characteristics of time series of annual precip. area-averaged over S. Canada (South to 55°N) and cont. U.S. 1900-1990(89).

In Fig.6 are shown the changes of ratios of solid to total annual precipitation during last four decades for three regions: Canada north and south of 55°N, and the contiguous U.S. We calculated them using stick snowfall measurements divided by ten as an estimate of precipitation fallen in solid form (data from 1107 stations in Canada and 209 primary stations in the contiguous U.S. were used for construction of this time series). Snow stick measurement is biased and inaccurate, but is a homogeneous estimate. Therefore, the ratios depicted in Fig.6 could serve as indices of redistribution between the solid and liquid parts of annual total precipitation. This is another important characteristic that is potentially sensitive to climate changes.

Fig.6 shows that there is no apparent redistribution between solid and liquid forms in northern Canada. In southern Canada, the annual total precipitation has increased during the century (Fig.5), but its solid part has decreased during last four decades, especially significantly during the last fifteen years. This decrease correlates with regional warming (about 45% of the variance of solid to total precipitation ratio could be explained by variation of regionally averaged maximum temperature). Solid to total precipitation ratio over the contiguous U.S. does not change systematically during last four decades, but its variance in the 1980s was 2.7 times higher than during the previous three decades. The probability of random occurrence of such a jump in variance is low (0.11).

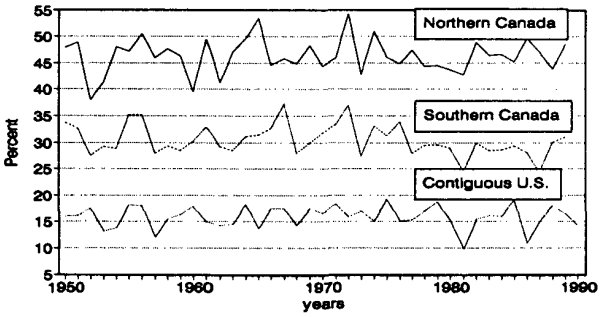


Fig.6. Ratios of solid to total precipitation area-averaged over Canada (zone 55-70°N and south to 55°N), and over the contiguous U.S.

SUMMARY

After elimination of inhomogeneities in the North American precipitation data (by adjustment and selection of the "best" network), reliable time series of annual snowfall and precipitation area-averaged over north and south Canada, and the contiguous U.S were obtained.

An increase in annual precipitation has been observed during last century over southern Canada (0.9% per decade) and the contiguous U.S. (0.5% per decade). However, for the U.S. the increase (the corresponding estimate of linear trend) is not statistically significant due to the large interannual variability of the precipitation totals.

In Canada to the north of 55°N a continuous increase of snowfall and rainfall has been observed during the entire period of instrumental observations (up to 20% during last four decades).

The zone 45-55°N was revealed as an area, where changes in annual totals of two forms of precipitation (solid and liquid) are extremely sensitive to temperature variations on continental and global scales. Also, there is *no* correlation between annual total zonal precipitation and temperature, and the redistribution between snowfall and rainfall occurs coherently with variations of the zonal maximum temperature.

Acknowledgements
This work was completed while the author was a National Research Council senior research associate at the U.S. National Climatic Data Center.

REFERENCES

Budyko, M.I. and Yu. A. Izrael (Eds). (1987) Anthropogenic climatic changes. *Leningrad, Gidrometeoizdat*. 406 pp.

Goodison, B.E., H.L. Ferguson and G.A. McKay (1981) Comparison of point snowfall measurement techniques. In: D.M.Gray and D.M. Male (Eds). *Handbook on Snow. Principles, processes, management & use*. Pergamon Press. 200-210.

Goodison, B.E. and P.Y.T. Louie (1986) Canadian methods for precipitation measurement and correction. *WMO/TD-No. 104. Instruments and observing methods. Report No. 25. Papers presented at the Workshop on the correction of precipitation measurements,* Zurich, Switzerland, 1-3 April, 1985, 141-145

Groisman, P. Ya., V.V.Koknaeva, T.A.Belokrylova, and T.R.Karl (1991). Overcoming biases of precipitation measurement: a history of the USSR experience. *Bull. Amer. Meteorol. Soc.* 72, 1725-1733.

Groisman, P.Ya., R.W.Knight, T.R.Karl, and R.R. Heim Jr. (1992) Inferences of the North American snowfall and snow cover with recent global temperature changes. *Proc. of the "Snow Watch'92" Conference. Niagara-on-the-Lake, Canada, March 29-April 1, 1992* (in press).

Intergovernmental Panel on Climate Change (IPCC) (1990) Climate Change. The IPCC Scientific Assessment. ED. J.T. Houghton, G.J. Jenkins, and J.J. Ephraums, *Cambridge University Press* N.Y., 362 pp.

Kagan, R.L., (1979). Averaging meteorological fields. *Gidrometeoizdat, Leningrad*, 212 pp. (in Russian).

Karl, T.R., C.N. Williams Jr., F.T. Quinlan, and T.A. Boden (1990) United States Historical Climatology Network (HCN) serial temperature and precipitation data. *NDP-019/R1, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee,* 83 pp. plus appendices.

Karl, T.R., R.G. Quayle, and P. Ya. Groisman (1992a) Detecting climate variations and change: New challenges for observing and data management systems. *J. of Climate* (in press).

Karl, T.R., P.Ya. Groisman, R.R. Heim Jr., R.W. Knight (1992b)
Recent variations of snow cover and snowfall in North America and their relation to precipitation and temperature variations. *J. of Climate*, (in review).

Sevruk, B. (1982) Methods of correction for systematic error in point precipitation measurement for operational use. *WMO Publication No. 589*. 91 pp.

Thiessen, A.H. (1911) Precipitation averages for large areas. *Mon. Wea. Rev.*, 39, 1082-1084.

World Water Balance and Water Resources of the Earth (WWB) (1974) *Gidrometeoizdat, Leningrad*, 638 pp. (in Russian; in English published by UNESCO Press in 1978).

CLIMATIC CHANGES NEAR THE GREAT LAKES INFERRED FROM 141 YEAR ICE RECORDS

Raymond A. Assel
NOAA/GLERL
2205 Commonwealth, 6417
Ann Arbor, MI 48105

Dale M. Robertson
USGS/WRD
Normandy Lane,
Madison, WI 53719

ABSTRACT. Freeze-up and break-up dates and duration of ice cover for lakes and rivers represent an integration of weather conditions prior to the specified event(s). Changes in mean ice conditions may be used as quantitative indicators of climatic changes if long homogenous ice records are accompanied by sufficiently homogenous air temperature records to calibrate the changes in mean ice cover in terms of climatic variables. Historical ice records dating back to 1855 are available for Lake Mendota, WI (located on the southwestern side of Lake Michigan) and back to 1851 for Grand Traverse Bay, MI (located on the northeastern side of Lake Michigan). Changes in the mean ice cover of these two systems were used to describe changes in fall, winter, and spring air temperatures in the area near the Great Lakes during the past 141 years.

INTRODUCTION. Many indicators have been used to detect and quantify historical climatic changes. The most obvious of these indicators is the weather records themselves. However, observational biases (for example, changes in observation time, measurement technique, and station location) have often caused apparent historical climatic changes (Schaal and Dale, 1977; Karl and Williams, 1987). In regions where historical weather records are very limited, other indicators have been used to detect and quantify past climatic changes or to verify or discredit the changes found in the limited weather records. One such indicator is the ice records of rivers and lakes. Annual dates of freeze-up or break-up are dependent on daily and often hourly air temperature and wind conditions; however, changes in the mean conditions (freeze-up, break-up, or total days of ice cover (ice duration)) represent changes in the local climate, primarily air temperature (McFadden, 1965; Palecki and Barry, 1986; Robertson, et al., 1992). Freeze-up and break-up events can occur over several days or weeks in some systems; therefore, these records are not free of observer biases. However, in some systems, freeze-up and break-up occur very quickly and can usually be resolved to within one or two days. Ice records for systems with rapid transition periods should have minimal observational biases and make good potential climatic-change indicators.

Changes in mean ice conditions can be used not only to detect climatic changes, but also to quantify these changes, provided the changes in ice cover can be calibrated in terms of climatic variables. Robertson, et al. (1992) used the ice records for Lake Mendota, WI, from 1855 to 1991, and the weather records for that area, for the period 1884 to 1991, to calibrate changes in mean freeze-up and break-up dates in terms of changes in air temperature. The changes in mean ice cover were then used to describe climatic changes occurring in southern Wisconsin during the ice cover record. Ice records from 1851 to 1991 are also available for Grand Traverse Bay, MI, located on the northeastern shore of Lake Michigan.

In this paper we will: 1) identify and quantify changes in air temperature near Grand Traverse Bay over the past 141 years, 2) compare and contrast climatic changes inferred for Grand Traverse Bay with those estimated for Lake Mendota, and 3) discuss changes in air temperature occurring throughout the Great Lakes region over the past 141 years.

STUDY SITES. Lake Mendota is a eutrophic lake located in the city of Madison in south-central Wisconsin (43°40' N, 89°24' W). The lake has a surface area of 39.4 km^2, a mean depth of 12.4 m, a maximum fetch of 9.8 km, and strongly stratifies during the summer. Uninterrupted annual ice records are available for Lake Mendota from 1855 to 1991. The exact criteria used to define freeze-up and break-up on Lake Mendota has not been defined; however, in general, these are rapid transitional events and can usually be described to within one or two days (Bryson and Bunge, 1956). Lake Mendota has formed a complete ice cover every year since ice records have been collected. Continuous daily weather data for Madison are available from 1869 to present. Inhomogeneities in the air temperature records associated with observer biases were removed by Robertson (1989) from 1884 to 1989. Air temperature data prior to 1884 contain several incorrectable observational biases and were therefore not used. All air temperature and ice records were obtained from the Wisconsin State Climatologist's Office.

Grand Traverse Bay, located on the northeast side of Lake Michigan (45°46'N, 85°37'W), is 51.5 km long and 16.1 km wide. The head of Grand Traverse Bay is divided into the east and west arms by a narrow peninsula extending 27.3 km northward. The west arm of the bay from the south shore to Marion Island, a distance about 9.6 km out in the bay, has a mean depth of about 46 m and an area of about 53.8 km^2. Freeze-up in this area is defined as the date that a solid ice cover forms out to Marion Island. Break-up is the date the ice breaks-up or moves out of this area. For both Lake Mendota and the western arm of Grand Traverse Bay, ice duration is the number of days with complete ice cover. In the case of multiple occurrences of freeze-up or break-up, the first freeze-up and last break-up are used in this analysis. In 25 of the 141 winters, a solid ice cover did not form out to Marion Island; ice duration during these winters was set to zero and the freeze-up and break-up dates were set to March 8. This date is the mid-point of the median dates of freeze-up and break-up for the five winters with the shortest ice duration. There were eight winters with missing data for freeze-up and four winters with

missing dates for break-up; these years were not used in our analysis. Dates of freeze-up and break-up for the west arm of Grand Traverse Bay from 1851 to 1973 were obtained from Snider (1974), and for winters 1974 to 1991 from the Chamber of Commerce, Traverse City, Michigan (personal communication). Air temperature records at Traverse City, Michigan were used for analysis of Grand Traverse Bay ice records. Only air temperature data after 1950 were used in this analysis because of inhomogeneities in the earlier record.

CHANGES IN MEAN ICE COVER. Changes in mean freeze-up, break-up, and total ice duration were determined by smoothing the original time series using a ten-year moving average and by plotting cumulative z-scores (a z-score is the standardized value calculated for each observation by subtracting the overall average and dividing by the standard deviation of the data). The ten-year moving averages for each ice-cover parameter are shown in Figure 1. A ten-year moving average was chosen because each ice cover parameter had significant periodicity between seven and ten years (found using Fourier spectral analysis). The smoothed time series for both Lake Mendota and Grand Traverse Bay demonstrate that mean freeze-up has become later, mean break-up has become earlier, and total ice duration has become shorter. These changes did not appear to be monotonic, but rather occurred as short term, rather abrupt, changes. These more abrupt changes can be more clearly seen in the cumulative z-score plots (Figure 2). The largest change occurs around 1888 when the cumulative z-score curve abruptly change slope. Prior to 1889, average freeze-up was earlier (from 1851 for Grand Traverse Bay and from 1856 for Lake Mendota), and average break-up was later (from 1865 for Grand Traverse Bay and from 1856 for Lake Mendota) (Table 1). These changes were found to be statistically significant at the 95% confidence level using a student's t-test. Taken together these changes suggest the fall, winter, and spring climate was significantly cooler prior to 1889. Unfortunately, the length of this cooler, earlier period is difficult to quantify from the present data. The data for Lake Mendota supported no earlier date to this period. The earlier (break-up) dates prior to 1865 for Grand Traverse Bay suggest warmer average late-winter and early- spring temperature (from 1851 to 1864) than the average from 1865 to 1888 (Table 1).

Since 1888, mean ice conditions for Lake Mendota were relatively stationary until after approximately 1979, when mean break-up became significantly earlier and ice duration has become significantly shorter (statistically significant at the 95% confidence level using a student's t-test). Since 1888, ice cover for Grand Traverse Bay has experienced two recent periods with earlier break-up and shorter ice durations: the late 1940's - early 1950's and the period since 1979. These changes also suggest the recent period was characterized by higher late winter and spring air temperatures. Very little change in average freeze-up dates was observed at either site, suggesting little change in fall and early-winter air temperatures. Based on these changes in ice cover, we divided the freeze-up date time series of both Grand Traverse Bay and Lake Mendota into two periods: prior to 1889 and from 1889 to present. The break-up date and ice duration time series were divided into three periods for each area: prior to 1889 (both systems), from 1889 to 1939 (Grand Traverse Bay) and from 1889 to 1979 (Lake Mendota), and from 1940 to 1991 (Grand Traverse Bay) and from 1980 to 1991 (Lake Mendota) (Table 1).

Table 1. Changes in average freeze-up (FZUP) and break-up (BKUP) dates (julian date) for Lake Mendota (LM) and Grand Traverse Bay (GTB), and the estimated changes in air temperature inferred from these changes using the sensible heat transfer model.

	Average FZUP	Days Later	Estimated Change(^{o}C)
LM			
1856-1888	347.8		
1889-1991	356.1	8.3	1.19
GTB			
1851-1888	37.0		
1889-1991	48.9	11.9	1.42
	Average BKUP	**Days Earlier**	**Estimated Change(^{o}C)**
LM			
1856-1888	101.0		
1889-1979	93.8	7.2	1.11
1980-1991	85.5	8.3	1.28
GTB			
1851-1864	87.1		
1865-1888	103.1	-16.0	-2.24
1889-1939	92.2	10.9	1.53
1940-1991	83.1	9.1	1.27

CALIBRATION OF ICE COVER RECORDS INTO CLIMATIC INDICES. Two types of models were used to relate changes in freeze-up and break-up dates with changes in air temperature: fixed period (Palecki and Barry, 1986) and sensible heat transfer (Robertson, et al., 1992). fixed period models are used to estimate the time period for which the average air temperature best correlates with either annual freeze-up or break-up dates. Annual freeze-up and break-up dates and coinciding average air temperatures of the one or two month period prior to the transition date have been found to be strongly correlated (Palecki and Barry, 1986). The slope of the regression line between transition dates and the average air temperature of a specific time interval yields an estimate of the response of that indicator per one day change in the average transition date. For Lake Mendota, the period from 1948 to 1987 was used in this analysis. For Grand Traverse Bay, the even years from 1952 to 1990 were used.

For Lake Mendota, the average November/December air temperatures were most strongly correlated with annual freeze-up dates and the average January through March air temperatures correlated best with break-up dates. The slope of

the regressions suggests that a one day change in mean freeze-up and break-up dates represents a 0.143°C and 0.146°C change in November/December and January through March air temperatures, respectively. For Grand Traverse Bay, average January air temperatures were most strongly correlated with annual freeze-up dates and average March air temperatures with break-up dates. The slope of the regressions suggests that a one day change in mean freeze-up and break-up dates represents a 0.085°C change and 0.056°C change in January and March air temperatures, respectively.

The sensible heat transfer freeze-up model estimates annual freeze-up dates by estimating when the surface water temperature on day t (W_t) becomes 0°C. Changes in water temperature, starting at 20°C on 1 October, are estimated using a sensible heat transfer algorithm (Table 2). The sensible heat transfer break-up model estimates annual break-up dates by estimating when the sub-freezing heat deficit on day t (I_t) is depleted (i.e., I_t becomes greater than 0) (Robertson, et al., 1992). Changes in I_t are estimated from the day after freezing (Table 2). Coefficients in the sensible heat transfer models were estimated using a Nelder-Mead Simplex minimization technique (O'Neill, 1971) to best approximate the freeze-up and break-up dates. The calibration period for the sensible heat transfer models was from 1948 to 1987 for Lake Mendota and the even years from 1952 to 1990 for Grand Traverse Bay. Estimated effects of changes in mean air temperature were made by comparing estimated average freeze-up dates of a 30-year period from 1948 to 1977 for Lake Mendota and from 1951 to 1980 for Grand Traverse Bay, with average predicted dates obtained from the models after daily air temperatures for the 30 years were raised or lowered by specific whole degree changes in air temperature. Daily air temperatures for all 30 years were lowered by 1 to 3°C and raised by 1 to 5°C. The slope of the line between the average transition date for each climatic condition versus the specified air temperature change yields an estimate of the response of that indicator per one day change in transition date.

For Lake Mendota, the slopes of the regressions between average transition dates and air temperature changes suggest a one day change in transition date represents a 0.143°C change in fall and early winter air temperatures and a 0.154°C change in winter and spring air temperatures. With a 4 or 5°C change in air temperature, Lake Mendota would not freeze for the first time on record although this would still be a rare event (one year in approximately 30). For Grand Traverse Bay, increases in air temperature of more than 1°C resulted in a non-linear response in transition dates because more and more years were estimated to remain ice free, and therefore, the estimated transition dates approach March 8. With increases of 3°C or more, it will be a rare event for Grand Traverse Bay to meet the freeze-up criteria. Therefore, only temperature changes between -3°C and +1°C were used in the regressions. The slopes of the regressions between average transition dates and air temperature changes suggest a one day change in transition date represents a 0.119°C change in fall and early winter air temperatures and a 0.140°C change in winter and spring air temperatures.

To determine which model provides the best estimate of the changes in air temperature, both models were used to simulate the annual transition dates for an additional period and the percent reduction in the root mean square errors was compared (Table 2). A 50 year comparison period from 1898 to 1947 was used for Lake Mendota and a 20 year period, odd years from 1951 to 1991, was used for Grand Traverse Bay. It is important to note that the fixed period models were developed to estimate changes in air temperature from changes in ice cover; therefore, additional relationships were computed to estimate annual transition events from air temperatures. For both systems, the sensible heat transfer models were significantly better at estimating freeze-up and break-up dates (Table 2); therefore, changes in air temperature estimated from these results should be more accurate and are indicated in Table 1.

Table 2. Summary of the freeze-up (Fzup) and break-up (Bkup) models for Grand Traverse Bay (GTB) and Lake Mendota (LM). The ability of the fixed period (FP) and sensible heat transfer (SHT) models to estimate annual transition events is demonstrated using the percent reduction in the Root Mean Square Errors (%RRMSE) compared to that using the mean of the entire period. The exact periods are defined in the text.

	Freeze-up Model	%RRMSE
GTB		
FP	JTa=-10.89+0.09Fzup Date	
	Fzup Date=68.35+2.65JTa	24%
SHT	W_t=0.06+0.02A_{t-1}+0.98W_{t-1}	49%
LM		
FP	NDTa=-51.01+0.14Fzup Date	
	Fzup Date=355.84+4.28NDTa	41%
SHT	W_t=0.08+0.05A_{t-1}+0.95W_{t-1}	53%
	Initial W_{t-1}=20°C on 1 October;	
	Break-up Models	**%RRMSE**
GTB		
FP	MTa=2.91-0.06Bkup Date	
	Bkup Date=76.77-3.28MTa	23%
SHT	I_t=0.04+0.08A_{t-1}+0.92I_{t-1}	73%
	Initial I_{t-1}=-0.05 on Fzup Date	
LM		
FP	JMTa=13.66-0.15Bkup Date	
	Bkup Date=93.47-3.34JMTa	22%
SHT	I_t=0.04+0.04A_{t-1}+0.96I_{t-1}	43%
	Initial It-1=-0.044 on Fzup Date	

Ta = air temperature; ND = November,December;
J =January; JM = January, February, and March;
M = March;
W_{t-1} = water temperature on day t-1;
A_{t-1} = air temperature on day t-1;
I_{t-1} = subfreezing heat deficit stored in ice on day t-1.

CLIMATE CHANGES OVER THE PAST 141 YEARS. Ice records from Lake Mendota and Grand Traverse Bay suggest fall, winter, and spring air temperatures in the Great Lakes region have changed over the past 141 years. Around 1890, a rather abrupt change in ice cover occurred in both systems: freeze-up dates became later, break-up dates became earlier, and total ice duration became shorter. Results from the sensible heat transfer model suggests these changes were caused by a regional increase in air temperature of 1.1 to 1.5°C (Table 1). Additional recent changes in mean break-up dates suggest winter and early spring air temperatures have further increased about 1.3°C; however, freeze-up dates have demonstrated little change, suggesting little change in fall air temperatures. Based on this study, the recent warming has only occurred over the past 10 to 15 years for the southwest region (Lake Mendota); however, the recent warming appears to have started in the 1940's in the northeast region (Grand Traverse Bay). Further studies are being conducted to examine the differences occurring between these sites.

ACKNOWLEDGEMENTS. GLERL Cont. No. 801

REFERENCES.

Bryson, R.A. and W.W. Bunge Jr. 1956. Ice on Wisconsin lakes. Report to the University of Wisconsin Lakes Investigation Committee. Department of Meteorology, University of Wisconsin-Madison, WI.

Karl, T.R. and C.N. Williams Jr. 1987. An approach to adjusting climatological time series for discontinuous inhomogeneities. Jol. of Climate and Applied Metero. (26):1744-1763.

McFadden, J.D. 1965. The interrelationship of lake ice and climate in central Canada. Technological report No. 20. Department of Meteorlogy. University of Wisconsin-Madison. WI.

O'Neill R. , 1971. Function minimization using a simplex procedure. Applied Statistics. (20):338-345.

Palecki, M.A., and R.G. Barry, 1986. Freeze-up and break-up of lakes as an index of temperature change during the transition season: A case study for Finland. Jol. of Climate and Applied Metro. (25):893-902.

Robertson, D.M.,1989. The use of lake water temperature and ice cover as climatic indicators. PhD Thesis, University of Wisconsin, Madison, Wisconsin.

Robertson, D.M., R.A. Ragotzkie, and J.J. Magnuson (in press - 1992) Lake ice records used to detect historical and future climatic changes. Climatic Change.

Schaal, L.A. and R.F. Dale. 1977. Time of observation temperature bias and "climate change". Jol. of Applied Metero. (16):215-222.

Snider, C.R. 1974. Great Lakes Ice Forecasting. NOAA TM NWS OSD 1. NTIS, U.S. Department of Commerce, 5285 Port Royal Rd. Springfield VA 22121.

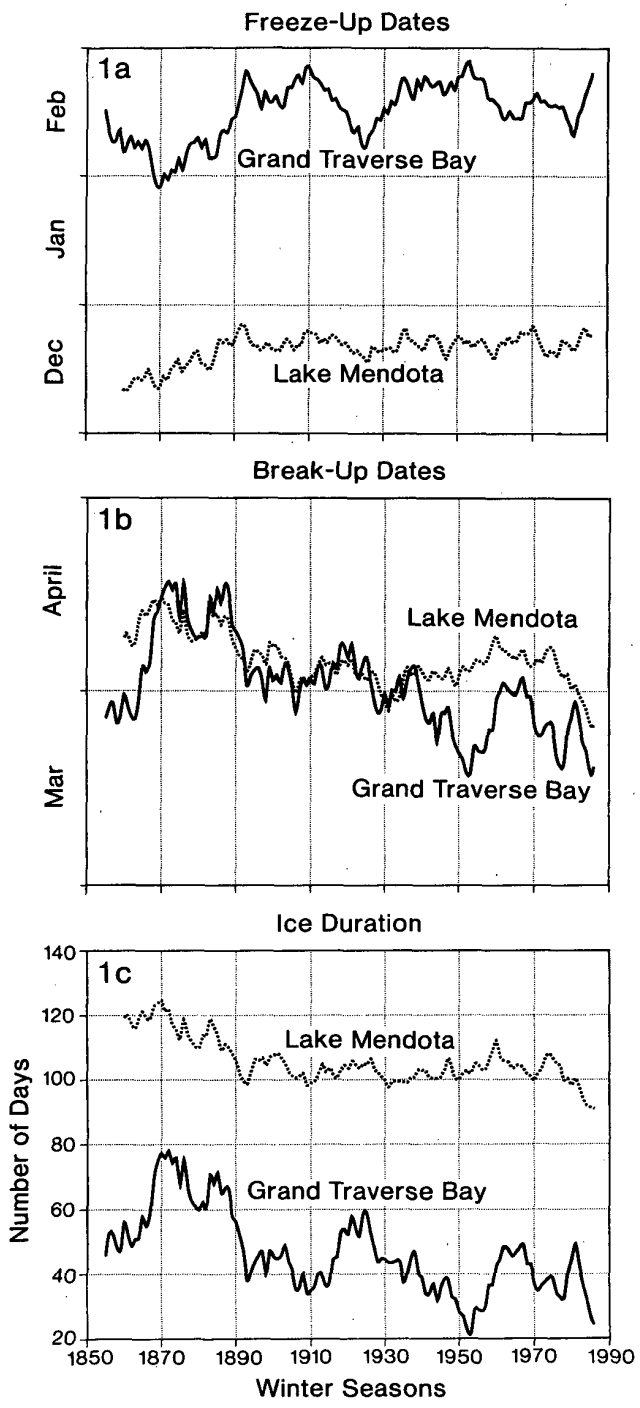


Figure 1. Decadal Running Average (1851–1991) Plotted on Year 5 of (1a) Freeze-Up, (1b) Break-Up, (1c) Ice Duration.

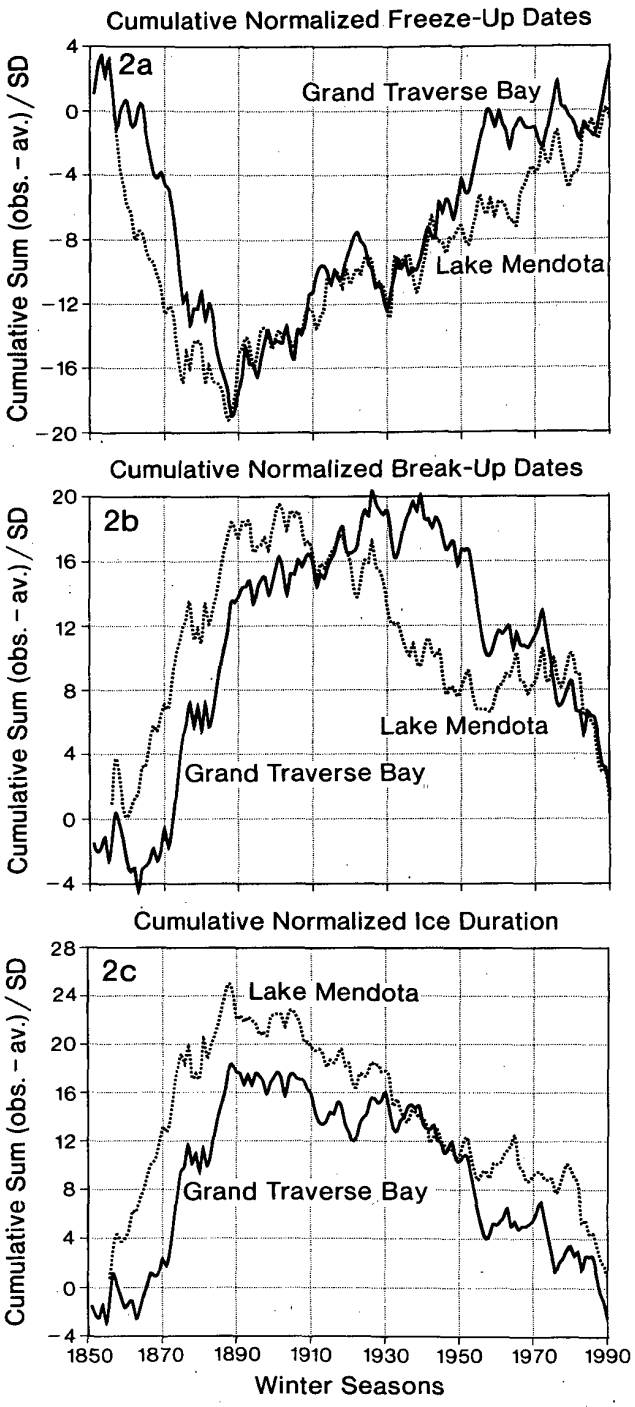


Figure 2. Cumulative Z-Score of Winters (1851–1991) of (2a) Freeze-Up, (2b) Break-Up, and (2c) Ice Duration.

Climate Change in Northeast Canada and the Northwest Atlantic: New Insights from a Long Term Sea Ice Data Set

John P. Newell

62 Sherwood Ave, Kitchener, Ont., Canada

Introduction

In the search for evidence of global warming, one region of the northern hemisphere stands out as having a pattern significantly different from the hemispheric trend. The region centered on southern Greenland, encompassing northeast Canada and the northwest Atlantic, has experienced a significant cooling trend over the past four decades (Jones and Kelly, 1983 and Jones and Wigley, 1990). Comparisons with the longer hemispheric record of air temperatures are hampered by the limited pre-1950 database in this region. A new long term sea ice record for the Labrador Sea (Newell, 1990) provides a new perspective on climate change within this region.

This ice record documents the date on which sea ice cleared from latitude 55.5°N off the coast of Labrador (Fig. 1). The ice record covers the period 1800-1990 and is continuous from 1854. Figure 2 compares the modern portion of this new data set with a record of the annual extent of sea ice coverage in the Baffin Bay/Labrador Sea area (Mysak and Manak, 1989). A visual comparison of the two data sets indicates good agreement regarding the timing of severe and mild ice years.

Ice/Climate Relationships

The use of sea ice records as proxy climate records is well established and numerous studies have demonstrated that sea ice conditions respond to atmospheric conditions in the previous months. From an analysis of modern data (1950-1989) Newell (1990) found that the date of ice clearing off Labrador responds to meteorological conditions over northeast Canada and the northwest Atlantic in the late spring and early summer, with the strongest relationship with the mean May sea level pressure pattern over this region. The relationship between mean May circulation patterns and clearing dates is one where later clearing dates (more severe ice conditions) are associated with mean May low pressure centers over Southern Greenland and earlier clearing dates (milder ice conditions) are associated with mean May low pressure centers over Hudson Strait.

Variations in Ice Conditions

The most striking aspect of the long term record of Labrador clearing dates (Fig. 3) is a marked decrease in ice severity (earlier clearing dates) following the 1920s. This corresponds with the warming trend experienced over the North Atlantic during this period (Rogers, 1985; Goossens and Berger, 1987; Yamamoto et al., 1987). The

abruptness and extent of the change in Labrador Sea ice conditions is demonstrated by a plot of the cumulative post 1854 ice index (Fig. 4). The abrupt change in ice conditions around the 1920s supports the arguments presented by Yamamoto et al. (1987) for a climate jump or discontinuity in North Atlantic atmospheric circulation around this time.

The new data set also provides new insights into events during recent decades. The data indicates that the 1960s had the earliest clearing dates of any decade examined. This is in contrast with other regions of the North Atlantic, such as the area near Iceland, which experienced relatively severe ice conditions during the 1960s. This suggest that the cooling over the northeast Atlantic during the 1960s did not represent a return to pre 1920s conditions, which had severe ice conditions off Labrador and Iceland, but represented a completely new pattern.

Conclusion

The Labrador sea ice record provides a long term perspective on the cooling trend experienced over the North Atlantic during recent decades. While ice conditions during the 1970s and 1980s are more severe than those during the 1950s and 1960s they are not nearly as severe as those during the nineteenth century. Comparison of the Labrador ice record with ice data from East Greenland and Iceland suggest that the pattern of ice conditions experienced during recent decades differs from that experienced prior to the 1920s. These differences may reflect the development of new patterns of atmospheric circulation over the North Atlantic.

REFERENCES

Goossens, C., and A. Berger, 1987. How to recognize an abrupt climatic change. In: **Abrupt Climatic Change**, W.H. Berger and L.D. Labeyrie (Eds.), D. Reidel Publishing Co., Dordrecht, pp. 31-45.

Jones, P.D., and P.M. Kelly, 1983. The spatial and temporal characteristics of Northern Hemisphere surface air temperature variations. **J. Climatology 3**, 243-252.

Jones, P.D., and T.M.L. Wigley, 1988. Recent global warmth during the 1980s and 1990. **Proceedings of the 15th Annual Climatic Diagnostics Workshop**, Ashevill N.C., U.S. Dept. of Commerce, NOAA, Climate Analysis Center, Washington, D.C., 344-349.

Mysak, L.A., and D.K. Manak, 1989. Arctic sea ice extent and anomalies, 1953-1984. **Atmosphere Ocean 27** (2), 376-405.

Newell, J.P., 1990. Spring and summer Sea ice and climate Conditions in the Labrador Sea, 1800-Present. Unpublished Ph.D. Thesis, University of Colorado, Boulder.

Rogers, J.C., 1985. Atmospheric circulation changes associated with the warming over the Northern North Atlantic in the 1920s. **Jour. of Climate and Applied Meteorology 24**, pp. 1303-1310.

Yamamoto, R., T. Iwashima and M. Hoshiai, 1987. Climatic jump in the polar region. **Proc. NIPR Symp. Polar Meteorol. Glaciol. 1**, pp. 91-102.

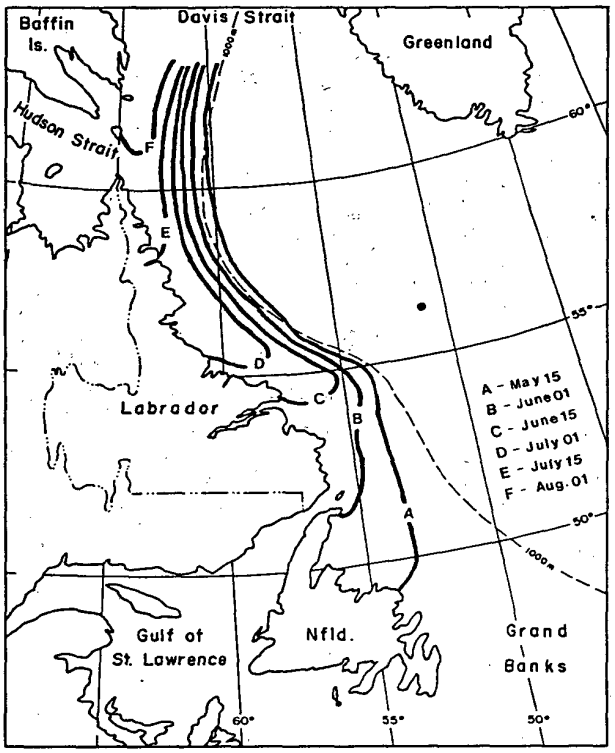


Fig. 1. Normal pattern of sea ice retreat along the coast of Labrador (1964-1973).

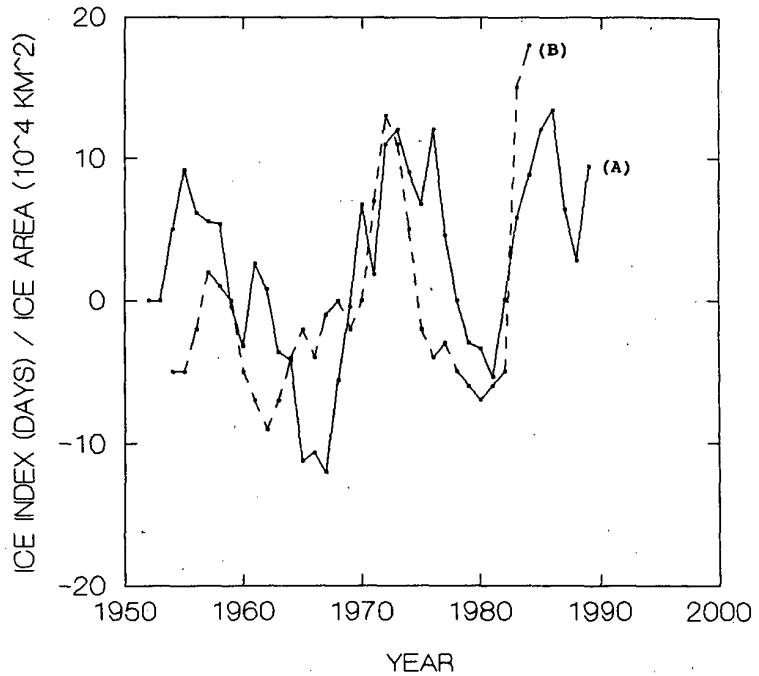


Fig. 2. Comparison of (a) the five year running mean Labrador Ice Index (Newell, 1990) and (b) a smoothed record of the annual extent of sea ice coverage in the Baffin Bay/Labrador Sea region (Mysak and Manak, 1989).

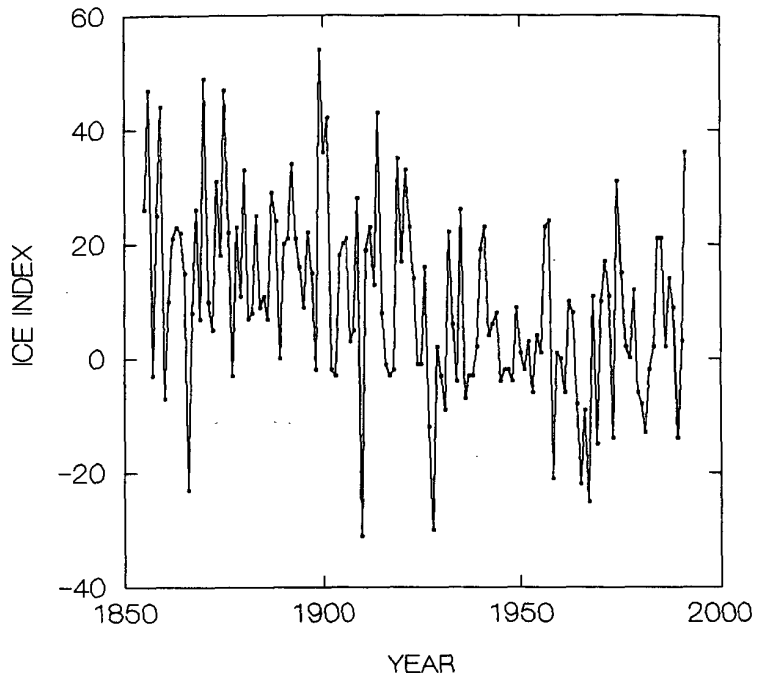


Fig. 3. The Labrador Ice Index. The departure in days of the clearing date at 55.5N from the mean clearing date for 1964-1984.

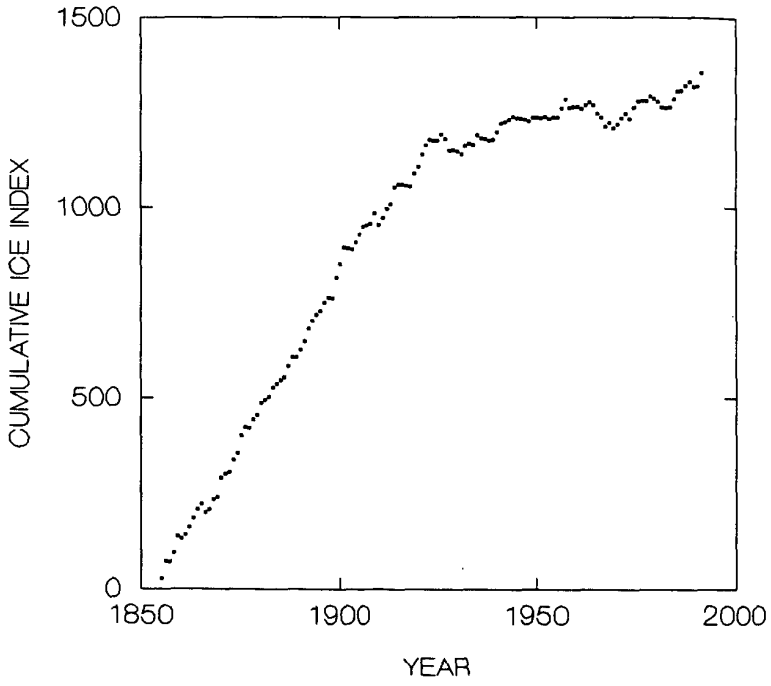


Fig. 4. The Cumulative post 1854 Labrador Ice Index.

CLOUDINESS AND DIURNAL TEMPERATURE RANGE

Povl Frich

Danish Meteorological Institute,

Lyngbyvej 100,

DK-2100 Copenhagen Ø,

Denmark.

Abstract

During the past 130 years the climate of the North Atlantic region has been observed at a number of stations. One of these stations is located in Copenhagen, Denmark. Through the years there have been gradual changes in the surroundings and the observational methods have changed. A simple time series plot of summer mean values has shown, that the diurnal temperature range apparently has been reduced by more than 4 °C and cloudiness has increased by more than 15 percent since 1861. A relocation of the thermometers caused a 1-2 °C reduction in the diurnal temperature range around 1920. The major part of this reduction can be ascribed to an abrupt increase in minimum temperatures. Using only the 1921-1991 part of the time series there is a high negative correlation ($r = -0.85$) between the mean diurnal temperature range and mean cloudiness. This result implies that variations in cloudiness is causing fluctuations in the diurnal temperature range. This result is discussed in relation to similar results from a number of rural stations in Denmark.

Introduction

In the IPCC update of observed climate variability and change, Folland et al., (1992) conclude that the diurnal temperature range has been reduced in many countries over the last few decades. This reduction is mainly due to an increasing trend in mean minimum temperatures (Karl et al., 1991). The mean minimum temperature is considerably influenced by the existence of clouds during the night, and the mean maximum temperature is affected by clouds during the day.

Despite the known problems with changes in observation code it is quite likely that observed increases in mean cloud cover (Henderson-Sellers, 1990) have contributed to the reduced diurnal temperature range. This has also been shown by Bücher and Dessens (1991), although their analysis included only one station in the Pyrenees.

This report will give additional evidence concerning temporal variations in cloudiness and diurnal temperature range from a number of stations in the North Atlantic region.

The author is coordinator of the North Atlantic Climatological Dataset (NACD) (Frich et al., 1991), in which 9 Northwest-european meteorological institutes are digitizing monthly values of more than 20 different climatic elements from more than 70 stations in the North Atlantic region.

Most of the stations in NACD have continous records which began around 1860/70, when national weather services were founded. Each country is responsible for the homogenization of their own national time series, but common methods will be applied (eg. Alexandersson, 1986).

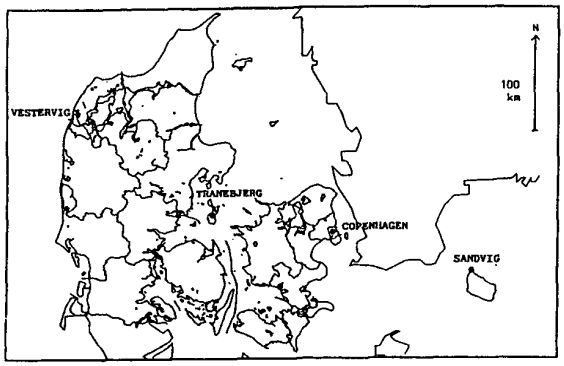


Fig.1 Locations of the 4 Danish climatological stations cited in the text.

The Danish contribution to NACD will include the urbanized Copenhagen series (1861→), and the rural series from Vestervig (1874→), Tranebjerg (1873→) and Sandvig (1874→) (c.f. figure 1). None of the rural stations are surrounded by more than a few thousand people whereas the population of Copenhagen now exceeds 1 million. A comparison of these 4 series has been made concerning diurnal temperature range and cloudiness. Some preliminary results from the summer period, June, July and August will be presented below.

Observations

Nearly all Danish climatological stations have always made observations at the same observational hours, which is 8 a.m., 2 p.m. and 9 p.m. Danish local time. Maximum and minimum temperatures have been read every morning at 8 a.m. and cloud cover has been observed at each observational hour. Until 1951 the cloud cover was observed in tenths and since 1952 in octas. All monthly means have been changed to percent of sky covered and no inhomogeneity has been detected which can be ascribed to the change in observational code.

Monthly mean of cloud cover has always been calculated as a simple mean of the 3 daily observations. Consequently there is a general bias towards daytime observations.

In Copenhagen the thermometers were prior to 1920 placed in a solid cage which was mounted on a north wall appr. 1.5 m above the ground. Around 1920 the thermometers were moved a few meters away from the wall and placed in a ventilated Stevenson screen 2 m above the ground. At Vestervig the thermometers were moved from a cage to a Stevenson screen in May 1924. At Tranebjerg the cage was replaced by a Stevenson screen in September 1919. The station was relocated i November 1972.

At Sandvig the Stevenson screen was introduced in 1914. The Sandvig station was replaced by the Hammerodde Fyr station in 1971. This new station is situated appr. 1 km to the North of Sandvig. Since 1987 we have used synoptical data from 06193 Hammerodde Fyr, but we have decided to retain "Sandvig" as the name of the time series. Some of these changes have led to problems with homogeneity which have not been solved yet.

According to Parker (1990) there is a bias in older parts of temperature series which is caused by a different design of the thermometer screens at that time.

Generally the height of the thermometer bulbs were increased from 1.3-1.5 to 2.0 m above the ground over the years 1914 to 1924 at these Danish stations. The thermometers were also moved away from buildings and ventilation improved. Both these changes may lead to a reduced diurnal temperature range (Parker, 1990).

Results

The summer (JJA) means of diurnal temperature range and cloudiness have been plotted in figure 2.

The 1921-1991 part of the Copenhagen time series and the 1925-1991 part of the Vestervig time series have been selected in order to overcome problems with homogeneity around 1920. The diurnal temperature range versus mean cloudiness is shown in figure 3. a-b.

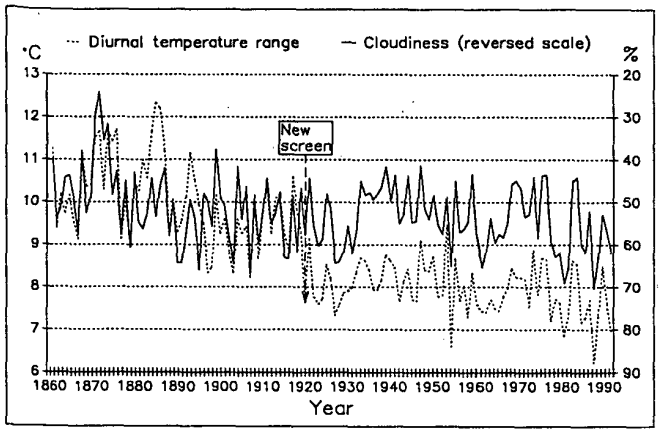


Fig. 2 Variations of diurnal temperature range and cloudiness in Copenhagen. Mean summer (JJA) values 1861-1991.

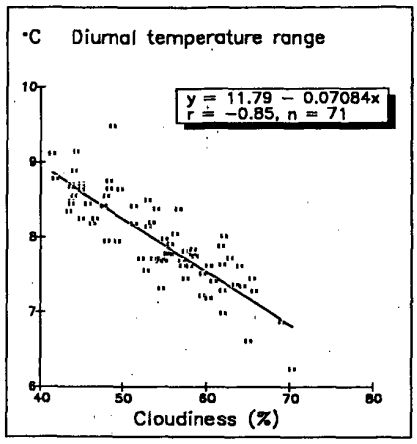


Fig. 3a Diurnal temperature range versus cloudiness in Copenhagen. Mean summer (JJA) values 1921-1991.

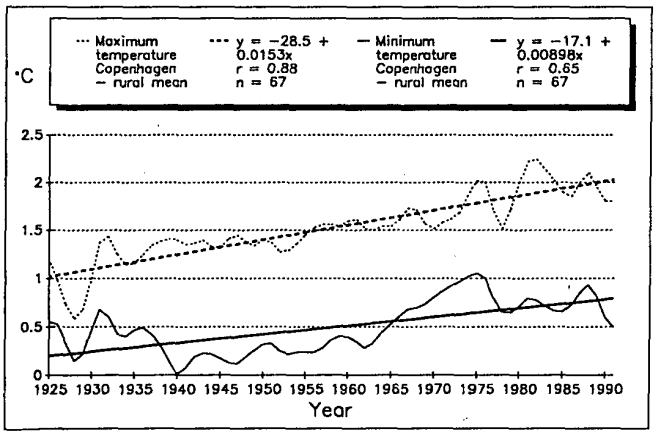


Fig. 4 Variations of urbanizational effects on maximum and minimum temperatures in Copenhagen. Mean summer (JJA) values 1925-1991.

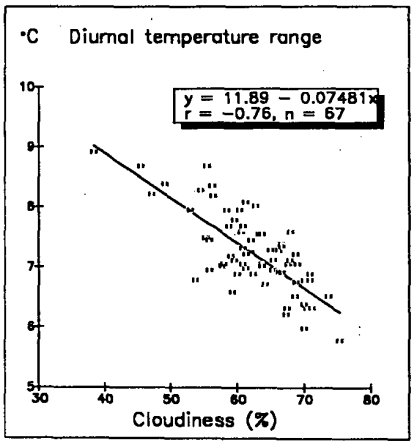


Fig. 3b Diurnal temperature range versus cloudiness at Vestervig. Mean summer (JJA) values 1925-1991.

For the period 1925-1991 all stations were equipped with Stevenson screens 2.0 m above ground. Mean values of maximum and minimum temperature series from the 3 rural stations have been subtracted from the Copenhagen series in order to depict the urbanizational effects on the Copenhagen series. The degree of urbanization is shown in figure 4 as filtered annual values and linear trends.

Discussion

The change from cages mounted on the north wall of buildings appr. 1.5 m above the ground to Stevenson screens (2.0 m) around 1920 caused a major inhomogeneity in the 4 Danish temperature series. The diurnal temperature range was artificially lowered 1-2 °C at the most protected site (i.e. Copenhagen during summer). The other Danish stations show a similar re-

sponse to the introduction of Stevenson screens.

There is a significant correlation between mean cloudiness and mean diurnal temperature range. The most urbanized and most protected site (i.e. Copenhagen) has the highest correlation (r= -0,85 c.f. figure 3), whereas a more windy rural site (i.e. Vestervig) has a correlation of -0,76. This means that variations in cloudiness can explain most of the observed variations in diurnal temperature range. This additional evidence confirms that increases in cloudiness results in reduced maximum and higher minimum temperatures (Karl et al., 1991).

The urbanization of the Copenhagen series after 1925 has led to a relative increase of both maximum and minimum temperatures compared with the surrounding country. Maximum temperature has increased 1 °C (linear trend) whereas minimum temperature has increased 0.6 °C over the 67 years. Please observe that the figures given are relative to a mean of 3 rural stations. This means that the urbanization of Copenhagen during summer amounts to 0.8 °C over the 67 years (equivalent to 1.2 °C/100 years). These figures are considerably higher than those cited by Jones et al., (1990).

Actually the mean maximum temperatures have been slightly decreasing whereas mean minimum temperatures have increased at all 4 stations. It is interesting, as noted by e.g. Folland et al., (1992) and Karl et al., (1991), that minimum temperature trends are steeper than maximum temperature trends in many industrialized areas. This assymetric behaviour can be explained by stronger atmospheric turbulence during daylight hours; meaning that the maximum temperatures are generally sampled from larger volumes of air. Minimum temperatures during nighttime inversions represent more stable conditions and are therefore more sensible to an increase in cloudiness.

References:

Alexandersson, H., 1986: A homogeneity test applied to precipitation data. Journal of Climatology, vol 6, p. 661 - 675.

Bücher, A. and J. Dessens, 1991: Secular trend of surface temperature at an elevated observatory in the Pyrenees. Journal of Climate, vol. 4. p. 859 - 868.

Folland, C.K., T.R. Karl, N. Nicholls, B.S. Nyenzi and K.Ya. Vinnikov 1992: Section C (observed climate variability and change) of the IPCC Science Report Supplement.

Frich, P., B. Brødsgaard and J. Cappelen, 1991: North Atlantic Climatological Dataset, (NACD) present status and future plans. DMI Technical Report 91-8.

Jones, P.D., P.Ya. Groisman, M. Coughlan, N. Plummer, W.-C. Wang and T.R. Karl, 1990: Assessment of urbanization effects in time series of surface air temperature over land. Nature, 347 p. 169 - 172.

Henderson - Sellers, A., 1990: Review of our current information about cloudiness changes this century. In "Observed Climate Variations and Change: Contributions in Support of Section 7 of the 1990 IPCC Scientific Assessment", D.E. Parker (ed).

Karl. T.R., G. Kukla, V.N. Razuvayer, M.J. Changery, R.G. Quayle, R.R. Heim jr., D.R. Easterling and C.B. Fu, 1991: Global Warming: Evidence for assymetric diurnal temperature change. Geophysical Research Letters, vol. 18 No. 12, P. 2253 - 2256.

Parker, D.E., 1990: Effects of changing exposure of thermometers at land stations. In "Observed Climate Variations and Change: Contributions in Support of Section 7 of the 1990 IPCC Scientific Assesment", D.E. Parker (ed).

ESTIMATING SPACE-TIME LOCAL HYDROLOGICAL QUANTITIES UNDER CLIMATE CHANGE

Istvan Bogardi[1],
Istvan Matyasovszky[2],
Andras Bardossy[3], and
Lucien Duckstein[4]

[1] Department of Civil Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588-0531, U.S.A.
[2] Department of Civil Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588-0531, U.S.A.; on leave from the Department of Meteorology, Eotvos Lorand University, Budapest, Hungary
[3] Institute fur Hydrologie und Wasserwirtschaft, University of Karlsruhe, Kaiserstr. 12 D-7500 Karlsruhe 1, Germany
[4] Systems and Industrial Engineering Department, University of Arizona, Tucson, AZ 85721, U.S.A.

1. INTRODUCTION

A modified space-time daily precipitation model (Bardossy and Plate, 1992) is presented which may be used to estimate local hydrological quantities under climate change in dry continental climate. Space-time series of precipitation are obtained by conditioning daily precipitation amount on daily circulation pattern (CP) types.

2. METHODOLOGY

We have used the National Meteorological Center (NMC) grid point analyses of the 500 hPa height fields. The analysis is based on daily values at 40 points covering the sector 25^0-60^0 N, 80^0-125^0 W for the period January 1948 - June 1989. The periods from April to September and from October to March were examined separately. A common classification method: principal component coupled with k-means clustering technique (Key and Crane, 1986) has been used to identify daily circulation pattern (CP) types. The number of CP types was chosen as a compromise between the increasing number of clusters and the decreasing sum of inner distances within each cluster. Defining nine types for both winter and summer half years seems to be a good compromise. The sequence of circulation patterns was described by a Markov chain.

Next, the time series of daily precipitation at spatially correlated locations is modeled as a multivariate autoregressive (AR) process. Bardossy and Plate (1992) presented a stochastic model to incorporate the occurrence of daily precipitation and its spatially correlated and highly skewed time series. Intermittence both in time and space, and skewness is modeled with the help of a power transformed normally distributed random vector. Let Z(t,u) the amount of daily precipitation at time t and location u. In order to relate Z(t,u) to a normally distributed random function the following power transformation is introduced:

$$Z(t,u) = \begin{cases} 0 & , W(t,u) \le 0 \\ W^{\beta}(t,u) & , W(t,u) > 0 \end{cases} \qquad (1)$$

where β is a positive exponent to be chosen in order to reproduce the skewness of the process Z(t,u), and W(t,u) is a normal random function. For locations u_k, k=1,2,..., K the vector **W** at time t is a K-dimensional normal random variable which is assumed to satisfy an AR(1) process:

$$\mathbf{W}(t) = \mathbf{B}\,(\mathbf{W}(t-1)-\mathbf{v}) + \mathbf{C}\,\Psi(t) + \mathbf{v}$$

where **v** is the expectation of **W**(t), and $\Psi(t)$ represents a K-dimensional normal white noise process with unit standard deviation. Using the properties of multivariate AR processes (Priestley, 1981) matrices **B** and **C** can be calculated as

$$B = G_1 G_0^{-1}$$

$$CC^T = G_0 - G_1 G_0^{-1} G_1^T$$

where G_l are the covariance matrices of W(t) for time lags l=0, 1.

The next task is to calculate the distribution of daily spatial precipitation given the daily CP type. As shown in Bardossy and Plate (1992) this can be done by considering each circulation type separately and calculate the multivariate normal process W(t) in the following way. If a type α_j persists at days t and t-1, that is, $A_t = A_{t-1} = \alpha_j$:

$$W(t) = B_j (W(t-1)-v_j) + C \Psi(t) + v_j . \qquad (2)$$

However, if the atmospheric circulation pattern just changes at time t, that is, $A_t = \alpha_i \neq A_{t-1}$ the precipitation process will have no "memory" at that time:

$$W(t) = D_i \Psi(t) + v_i \qquad (3)$$

where $D_i D_i^T = G_{0i}$

The main difficulty in using the above stochastic model is that process W(t,u) cannot be observed. As a consequence, its parameters, the expectation and the covariance matrices must be indirectly estimated. A moment method introduced in Bardossy and Plate (1992) has not reproduced satisfactorily the observed precipitation series in the case of dry continental climatic conditions. In consequence, the following modified approach has been used. First, for several value of β the probability distribution function of $W^{\beta}(t,u_k)$ is fitted to the empirical distribution function of precipitation by considering only large quantiles (say greater than 0.85). Values of β, v_k and σ_k that deliver the minimum of maximum absolute difference between the two distribution functions are accepted as estimates. The resulting distribution provides a good description of large precipitation amounts, but it behaves unsatisfactorily for either the precipitation probability p_k or small amounts of precipitation. This transformation serves only to estimate covariances G_0 and G_1 as will be shown next; the simulation is then based on another procedure described subsequently.

To obtain an estimation for covariances it is necessary to estimate correlations and multiply them by appropriate standard deviations that are already known (σ_k). The method developed in Bardossy and Plate (1992) is used to estimate the correlation matrices R_0, R_1.

One possible application of the above model is the simulation of time series of spatially distributed daily precipitation. First, a set of circulation patterns has to be selected. This may be a historical data set, but circulation patterns may also be generated by simulation using the Markov model mentioned earlier.

Generation of a time series of process W(t,u) and thus Z(t,u) starts with the simulation of white noise process $\Psi(t)$ consisting at a given t of K independent normal random variables, each with zero expectation and unit standard deviation. Eq.(2) or (3) is then used depending on whether there is a change in circulation pattern at the given time. Then, matrices B^*, C^*, D^* are defined using correlations instead of covariances to produce a process with unit standard deviation and zero expectation:

$$X(t) = B^*_j X(t-1) + C^*_j \Psi(t) , \qquad (4)$$

$$X(t) = D^*_i \Psi(t) \qquad (5)$$

depending on whether there is a change in circulation pattern at time t. Each variable X(t) for any t has zero expectation and unit standard deviation. Process X(t) provides a good representation of the required space-time correlation structure but is not suitable for direct precipitation simulation. As a consequence, another transformation is introduced that establishes relationships between quantiles of the empirical distribution of precipitation and quantiles of the standard normal distribution. More exactly, let x be a value at location u_k for circulation pattern α_i obtained from Eqs. (4)-(5), and $H_{ik}(z)$ be the empirical distribution function of the

precipitation at the indicated location for the given circulation pattern. Then we have:

$$z = H_{ik}^{-1}(\Phi(x)) \qquad (6)$$

Simulating a time series of X(t) and using Eq.(6) yields a simulated precipitation series.

3. APPLICATION

Daily winter precipitation data at nine stations between 1950 and 1989 are used. Daily precipitation in Eastern Nebraska is quite variable both in time and space: it is characterized by a relatively high probability of zero precipitation, a skewed distribution and spatial dependency. The hydroclimatological model can reproduce these characteristics.

To characterize two daily CP types, for example CP1, the "wettest"type, is dominated by southwesterly flows in Central and Eastern USA. Warm and wet air masses advected from the subtropics cause large amount of precipitation. CP8, the "driest" type, is a pattern with strong northwesterly flows. Cold air masses are dry due to the passage over the mountainous west. The relative frequencies of the 9 types are quite uniform. The occurrence of different CP types has a substantial influence on space-time precipitation. For instance, there is a sixteen fold difference between means of daily precipitation conditioned on the wettest and the driest CPs. Figs. 1 and 2 illustrate the probability distribution function (PDF) of daily precipitation at station 2 given type CP3 and CP4, respectively.

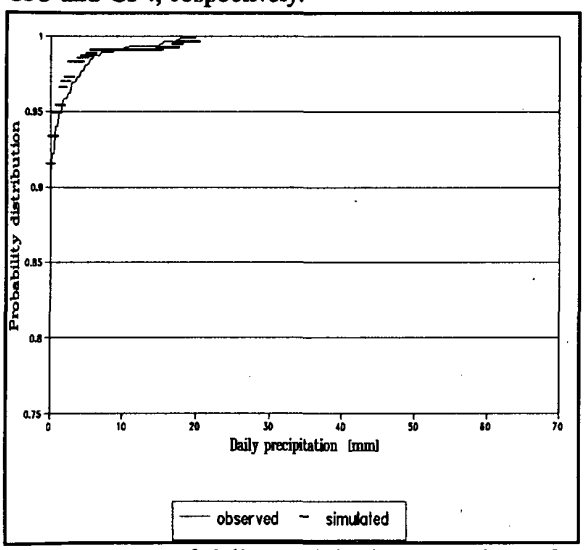


Figure 1. PDF of daily precipitation at station 2 for CP3

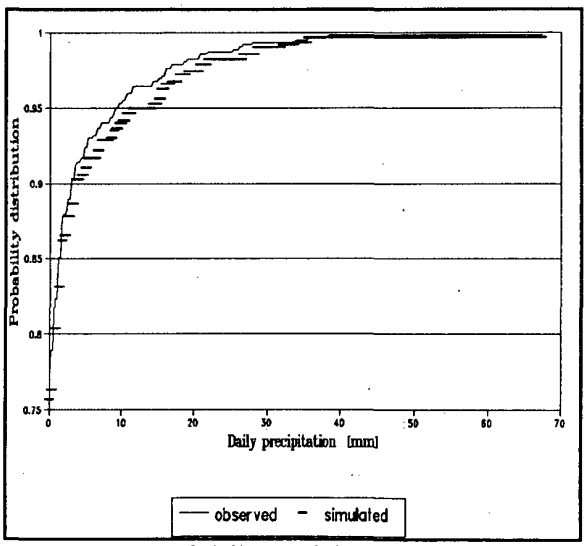


Figure 2. PDF of daily precipitation at station 2 for CP4

The effect of spatial variability can be examined in Figs. 3 and 4 where PDFs of daily precipitation at two different location for CP4 are shown.

The application considers only the winter half year; for which different CPs influence daily precipitation in a great deal. A similar method has also been applied to the summer half years (Matyasovszky et al., 1992). The approach can be used to model other local climatic variables such as temperature, wind and evaporation (Bardossy, 1991). Presumably it can be extended to model streamflow and other related hydrological variables such as floods (Duckstein et al., 1992) or lake level changes. In the latter case one of the main difficulties is to describe the lag time between daily CP and flood peaks or monthly lake water levels.

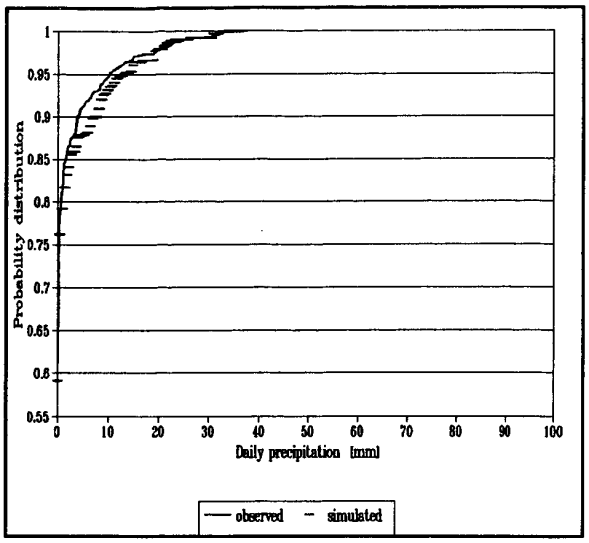


Figure 3. PDF of daily precipitation at station 1 for CP4

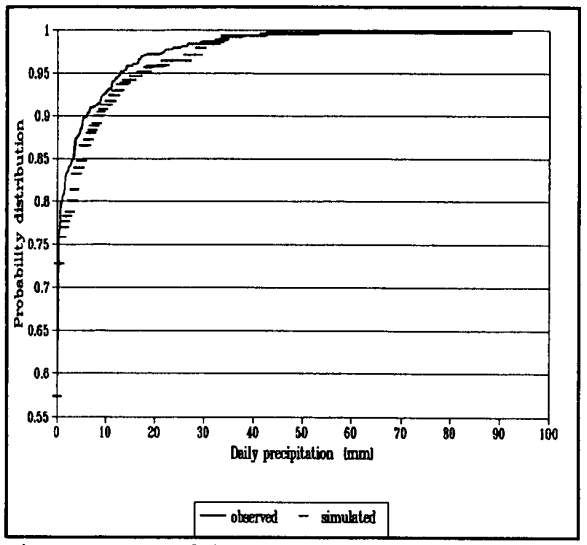


Figure 4. PDF of daily precipitation at station 8 for CP4

4. DISCUSSION AND CONCLUSIONS

Using the approach to predict the regional/local effect of climate change scenarios may be done in three parts: 1. Estimation of the CDF of the climatic variable such as daily precipitation from observation data; 2. Construction of a Markov model of CP types obtained from the output of a General Circulation Model (GCM); 3. Using the stochastic linkage model to simulate space-time series of the climatic variable reflecting climate change. In the actual realization of this idea the following problem has emerged. The difference between the relative frequencies of historical CP types and CP types of GCM produced by $1xCO_2$ case is found to be larger than the difference between $1xCO_2$ and $2xCO_2$ CP types. We have used results of the Canadian Climate Centre GCM to obtain 10-year series of daily CP, which appears to be the most accurate output of any GCMs (Crane and Hewitson, 1991). Before classifying CPs, as it can be expected, the $2xCO_2$ CPs looked quite different from the $1xCO_2$ CPs. The classification procedure used standardized values which resulted, finally, in very similar frequencies of all CP types for $1xCO_2$ and $2xCO_2$. The methodology presented in this paper may be a useful operational tool for disaggregation, but the present classification method should be modified in order to measure the difference between CPs reflecting present conditions and climate change scenarios.

The following conclusions can be made.

1. Principal component analysis and the k-means method can bee used to define nine types of daily circulation patterns in West-Central USA on the basis of 40 years of data.
2. A Markov model appears to be appropriate to describe the occurrence and duration of CP types.
3. The hydroclimatological model described in Bardossy and Plate (1992) must be modified to accommodate precipitation under dry continental climatic conditions.
4. The probability and amount of winter precipitation is strongly related to CP types.
5. The model accurately reproduces the probability and amount of both point and areal precipitation.
6. The model can be applied to simulate time series of spatial precipitation with climatic (CP) input.
7. Further research is needed to use the model for predicting the regional or local hydrological effect of climate change.

ACKNOWLEDGMENTS

Research leading to this paper has been supported by grants from the U.S. National Science Foundation, Nos. BC-9016462 and BC-9016556.

REFERENCES

Bardossy, A., 1991: Conditional local hydrological variables on circulation patterns in West Germany. Working paper (in German), Institute for Hydrology and Water Resources, University of Karlsruhe, Karlsruhe, Germany.

Bardossy, A., and Plate, E., 1992: Space-time model for daily rainfall using atmospheric circulation patterns. Water Resources Research (in press).

Duckstein, L., Bardossy, A., and Bogardi, I., 1992: Linkage between the occurrence of daily atmospheric circulation patterns and floods: an Arizona case study. Working Paper 92-3, System and Industrial Engineering Department, University of Arizona, Tucson, Arizona.

Key, J., and Crane, R.G., 1986: A comparison of synoptic classification schemes based on "objective procedures". Journal of Climatology, 6, 375-386.

Matyasovszky, I., Bogardi, I., Bardossy, A., and Duckstein, L., 1992: Hydroclimatological Modeling of Droughts Under Climate Change. Proceedings, 16 th European Regional Conference of International Commission on Irrigation and Drainage, Budapest.

Priestley, M.B., 1981: Spectral Analysis and Time Series. New York and London, Academic Press.

CHANGES OF CLIMATIC CHARACTERISTICS IN FENNOSCANDIA

Raino Heino
Finnish Meteorological Institute
Helsinki, Finland

Abstract

Climatic changes in Fennoscandian area (60-70°N , 20-30°E) are reviewed. Changes of temperature are similar to changes in North Atlantic area i.e. warming until the 1940s and subsequent cooling. Recent years have been relatively warm again. Daily temperature amplitude shows some decrease during the century. Changes of precipitation are not clear, which is due to the great temporal and spatial variability of precipitation.Moreover, precipitation records contain serious inhomogeneities. However, some increase of precipitation is seen in Fennoscandia. Preliminary analysis of other climatic parameters show decrease in relative humidity, while other parameters show no significant changes during the 1900s.

1. INTRODUCTION

The greatest potential climatic changes induced by increasing atmospheric greenhouse gases are estimated to occur in the high latitudes, although, due to the great natural variability, the signal of the change will be difficult to detect. Despite this problem, data from polar regions should be examined especially closely. The network of long-term climatological stations, however, is relatively sparse in the high latitudes.

The data of this study consist of the long-term Finnish climatological stations extending at least 100 years backwards completed with some long-term data records of the neighbouring countries (Fig. 1).

Inhomogeneities of the data records may arise from station relocations and environmental changes (e.g. urbanization) as well as changes in instruments and observing methods. These apparent changes may be in single records even greater than the real change of climate.

Urbanization and other man-made environmental changes are not, in general, a major concern in Nordic areas. In Helsinki, however, the present temperature difference between the station in the city centre and the surrounding stations is approx. 1°C and should be taken into account.

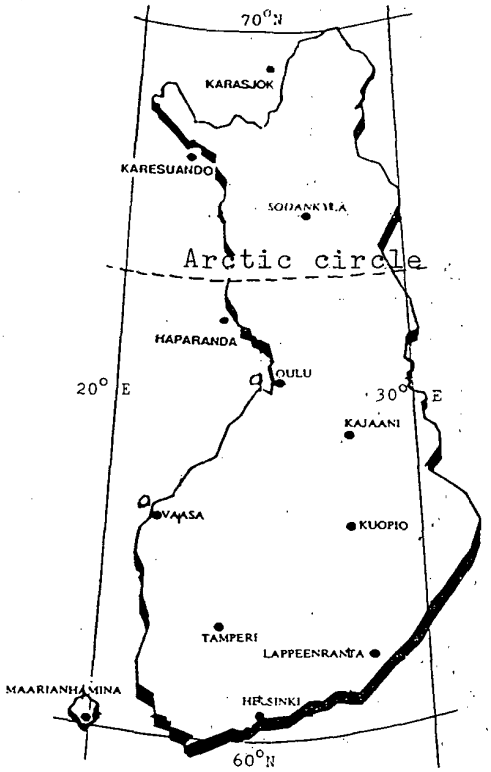


Fig. 1. Long-term climatological stations in Finland and neighbouring areas.

Station relocations may cause a major inhomogeneity source in single data records (e.g. even 1-2°C change in mean temperatures and up to 30-40% change in precipitation amounts).

Instrumental and observational changes have introduced relatively high inhomogeneities in the data. Changes in averaging methods of temperature in the 1920s have caused up to 1°C apparent cooling of the summer temperatures in Finland. The greatest systematic inhomogeity of the Nordic station records, however, is due to improvements of rain gauges. In Finland the gauge changes in 1909 and 1981 have caused all together an apparent increase in winter precipitation by approx. 30%.

Data homogenization is under way and this requires the handling of the historical characteristics of the data (metadata) based on station inspection reports and other sources of information. It is also necessary to digitize part of the metadata for the operational quality and homogeneity checking. The method to be applied by the Nordic countries in the data homogenization is developed by Alexandersson (1986).

2. CHANGES OF TEMPERATURE

According to Jones (1985) Arctic zone (65-85°N) temperatures increased strongly from the end of last century until the 1940s and turned downward during the following decades. Some recent warming is again visible.

Similar trends are true also in Fennoscandian mean temperatures (Fig. 2).

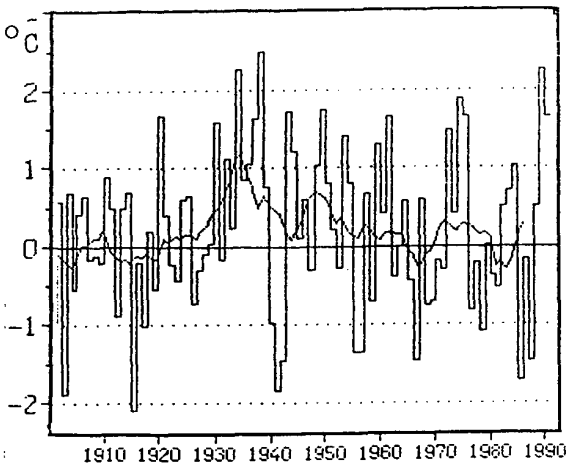


Fig. 2. Annual mean temperatures of the Fennoscandian area (60-70°N, 20-30°E) in 1901-1990 as a deviation from the 1961-1990 mean value (mean values of 30 stations of the region and 15-year moving averages).

Single station records show the similar features (Fig. 3). However, the warming until the 1940s as well as the subsequent cooling have been stronger in the northern stations than in the south.

The difference between the mean maximum and minimum temperatures (Fig. 4) shows a slight decrease in Finland similar to what has been observed elsewhere in the Northern Hemisphere. However, the minimum (night) temperatures have remained unchanged, while decreased maximum (day) temperatures explain mainly the amplitude trend.

Changes of other temperature parameters (e.g. growing degree days, number of frost, ice and hot days) show similar long-term changes as mean temperatures.

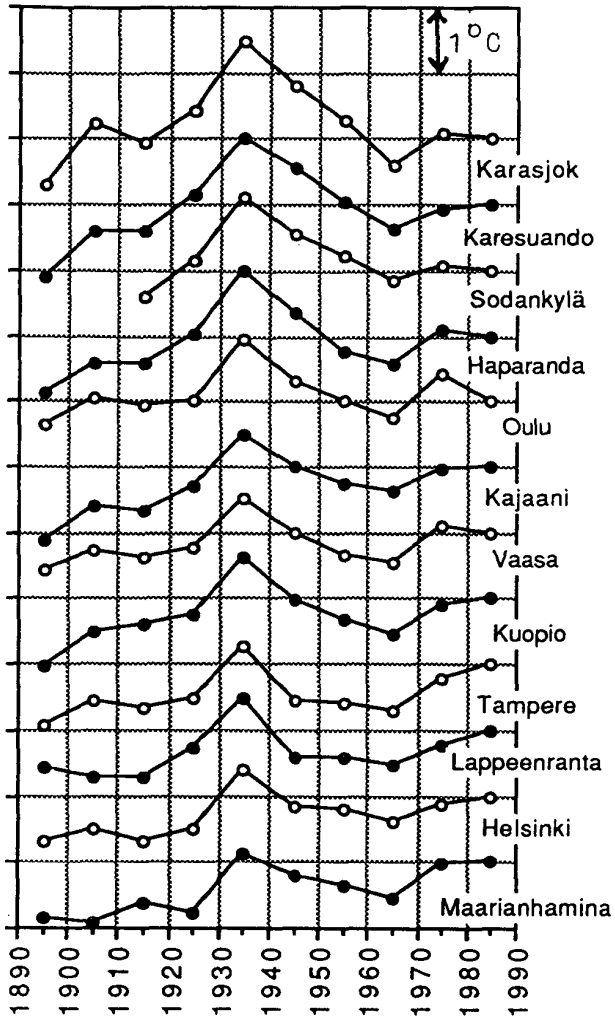


Fig. 3. Decadial mean values of annual mean temperature in Fennoscandian area since 1891 (deviations in °C from the 1981-1990 mean values of each station).

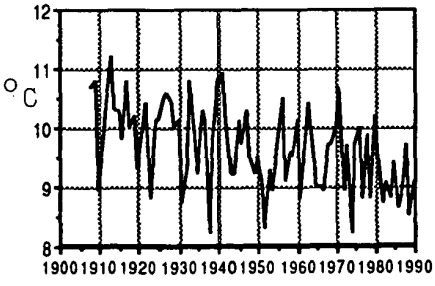


Fig. 4. Annual temperature amplitude (difference between annual mean maximum and minimum temperatures) in Sodankylä (67°N, 27°E) in 1908-1990.

3. CHANGES OF PRECIPITATION

Precipitation data is difficult to use, because variations in space and time are great. Although the network of stations measuring precipitation is normally denser than that of temperature stations, the measuring instruments and their changes as well as relocations of stations cause strong inhomogeneities in the data records.

According to Bradley et al. (1987) and Groisman (1991) precipitation has increased in the high latitudes of the Northern Hemisphere during the last decades. The long-term precipitation record in the Fennoscandian region (Fig. 5) shows also a recent increase, which is partly (2-3%) due to the change of all the gauges in Finland in 1981/82.

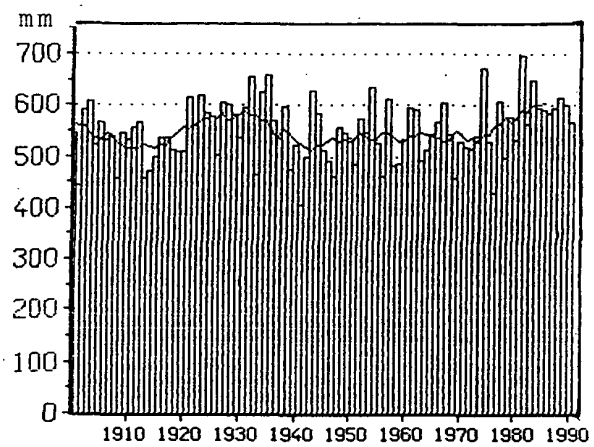


Fig. 5. Annual precipitation amounts of the Fennoscandian area (60-70°N, 20-30°E) in 1901-1990 (mean values of 30 stations of the region and 15-year moving averages).

Single station records, however, show very different features (Fig. 6), which is partly due to inhomogeneities of the data. However, some recent increase of precipitation is seen at several stations.

Days with precipitation is another measure in the study of precipitation changes. Fig. 7 shows a decreasing trend in the number of precipitation days since the 1930s.

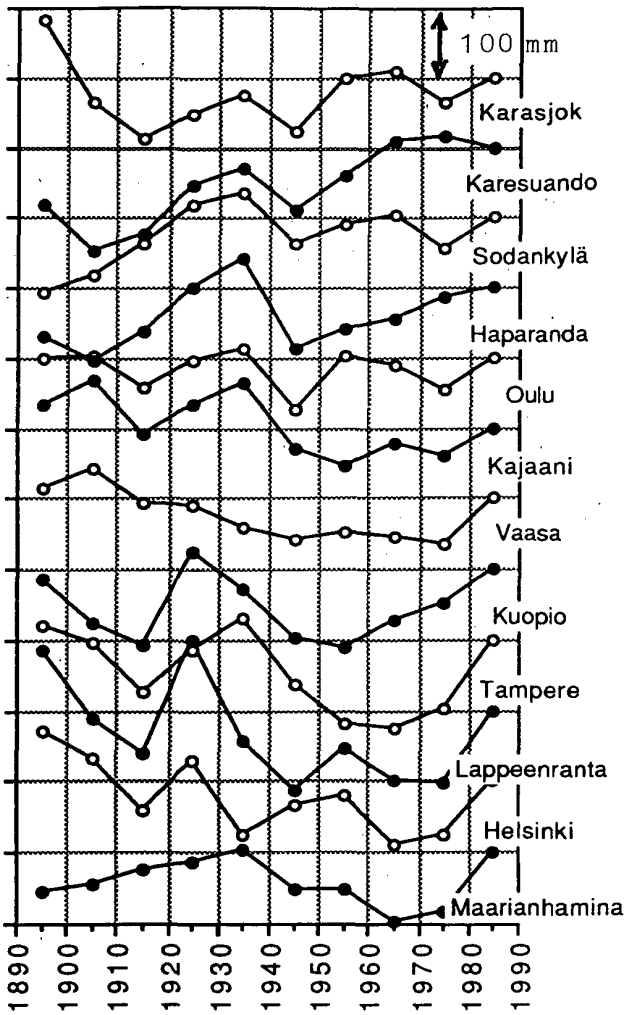


Fig. 6. Decadial mean values of annual precipitation amounts in Fennoscandian area since 1891 (deviations in mm from the 1981-1990 mean values of each station).

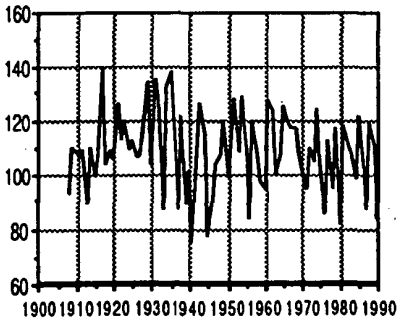


Fig. 7. Annual number of days with precipitation (at least 1 mm/day), in Sodankylä (67°N, 27°E) in 1908-1990.

4. CHANGES OF OTHER CLIMATIC PARAMETERS

In addition to temperature and precipitation several other climatic parameters are being jointly worked out by the Nordic countries (cf. Heino 1990). A few preliminary examples of them are presented in Fig. 8 for one northern station (Sodankylä observatory).

Annual mean cloudiness (Fig. 8a) does not show the same increasing trend that has been seen in many other North Atlantic regions (e.g. Karl-Steurer 1990). On the other hand, shorter records of sunshine duration show lower values during the 1980s.

Relative humidity (Fig. 8b) shows a clear decrease since the 1940s. On the other hand, this may be due to some unknown inhomogeneity in the humidity data. Days with snow cover (Fig. 8c) were more frequent in the 1960s than before and after that decade.

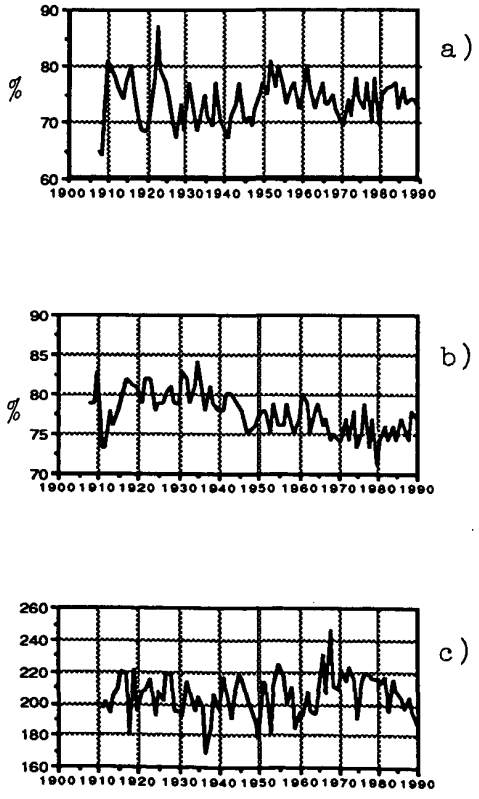


Fig. 8. Annual mean cloudiness (a), relative humidity (b) and number of snow cover days (c) in Sodankylä observatory (67°N, 27°E) in 1908-1990.

5. CONCLUDING REMARKS

Changes of temperature and precipitation in Northern Europe are consistent with the corresponding changes of the whole northern sector of the Atlantic and to some degree with the whole Northern Hemisphere.

Changes of the other climatic parameters need still more research in Finland as well as elsewhere. Further work in Finland directs to better knowledge of these records and to the use of various statistical parameters. This will be done jointly with the five Nordic countries and some West-European countries (cf. Frich et al. 1991).

REFERENCES

Alexandersson, H., 1986: A homogeneity test applied to precipitation data. Journal of Climatology 6.

Bradley, R.S. et al., 1987: Precipitation fluctuations over Northern Hemisphere land areas since the mid 19th century. Science 10 July 1987.

Frich, P. et al., 1991: North Atlantic Climatological Dataset (NACD); present status and future plans. Danish Meteorological Institute, Technical Report 91-8.

Groisman, P., 1991: Unbiased estimates of precipitation change in the Northern Hemisphere extratropics. In: Fifth Conference on Climate Variations, Oct 14-18,1991, Denver, Colorado, American Meteorological Society.

Heino, R., 1990: Nordic Climatological Data Set. In: Report of meeting of experts on climate change detection project, WMO-WCDP-No.13.

Jones, P.D., 1985: Arctic temperatures 1851-1984. Climate Monitor 14:2.

Karl, T.R. & P.M.Steurer, 1990: Increased cloudiness in the United States during the first half of the twentieth century: Fact or fiction? Geophysical Research Letters 17:11.

CHANGES IN ISRAEL'S PRECIPITATION PATTERNS DURING THE PERIOD 1960-1990

N. Gazit-Yaari and E.H. Steinberger

Department of Atmospheric Sciences,
The Hebrew University of Jerusalem, Jerusalem 91904, Israel

ABSTRACT

Quantitative estimates of global precipitation amounts are necessary to verify general circulation models (GCM's). GCM's are our main tools of simulating climate, and climate change. Temporal and spatial analysis of global precipitation is not yet available. Nevertheless our best resolution can not yet describe regional changes. Precipitation is a strong climatic factor, which is very sensitive to changes in the general circulation. It is, also, an important indicator of climate change.

The importance of precipitation to the water resources in Israel need not be elaborated. Israel's water reserves are in a very poor condition, and Lake Kinneret, the only relatively large body of water, situated in the northern part of the country, is at risk of salinification, due to its decreasing water level.

Israel, although small in size, consists of more than five different climatological sub-regions. Regional analysis is, therefore, the only tool in this case, of detecting climatological changes.

Yearly precipitation amounts of 44 different stations were analyzed, using moving averages technique and straightforward linear regression so that trends could be identified and isolated from the "year to year" variation. The data set consisted of thirty years data sequence (1960-1990),

with a seven year moving average. The apparent yearly fluctuations were smoothed, and the results generally suggest a decrease in annual rainfall amounts in the north, and an increase in the southern and central mountain regions during the period of study.
The adjacent diagram shows the results for all 44 chosen stations.

The clear negative trend throughout the northern region amounts in some stations to a decrease of more than 1% per year. This is of special concern to the water management authorities because Lake Kinneret's drainage valley also recieves smaller amounts of precipitation, causing its water level to continue descending.
When looking at the seven year moving average of a typical northern station, at a visual, non-statistical level, an obvious trend towards a drier period clearly emerges. When using statistical methods, we find a decrease in the order of around 1% per year with values of R^2 being about ≈ 0.7 at the stations with a pronounced decrease.

In the southern and central mountain regions we find an area of precipitation increase that is also statisticaly significant. The results for a number of stations indicate an increase of almost 1% per year.

We presume that these results are due to changes in the character of the synoptic systems producing Israel's rainfall. These changes could be because of a shift in cyclone tracks that might have been caused by changes in general circulation. These changes produced a situation in which more than before, an increasing amount of Israel's precipitation is associated with Southern lows, coming from the South in a northward direction. This is very much different from the past when most of Israel's precipitation was associated with lows developing over the Mediterranean Sea coming from the North-West and moving inland. These lows now move in a more northward direction producing rain over Turkey, and at some occasions, a small amount of precipitation in the Northern part of Israel.

At this time, we do not yet have the full statistics that is needed for this assumption to be validated. We hope in the near future to complete the analysis by checking the nature of the systems producing most of the annual rainfall in Israel.

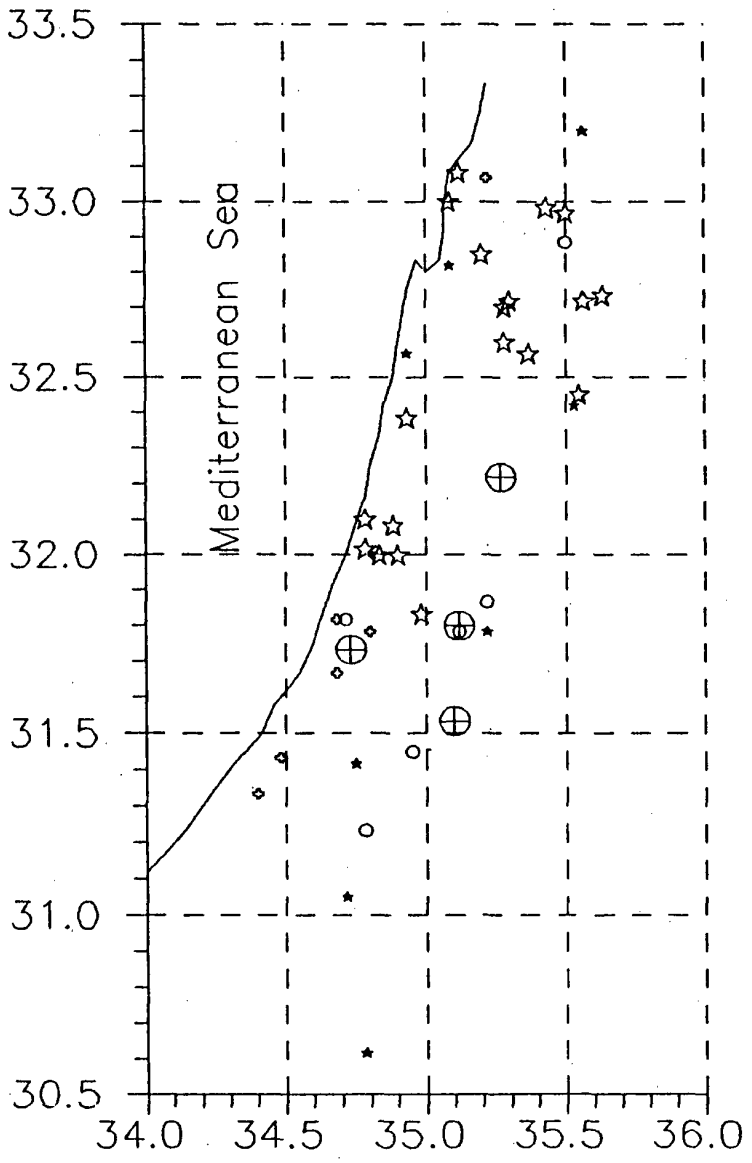


Fig. 1 : Spatial distribution of raifall trends in Israel.

big star - large decrease , small star - small decrease
circles - no significant change
small plus - small increase , big plus - big increase

The Climate Variability in Fujian Province of China During the Last 1000 Years

Lin Zheng—yun

Department of Geography, Fujian Normal University, Fuzhou, Fujian

ABSTRACT

This paper uses the documentary data (continuous rain, flood, drought, snow, freezing , and frost) and instrumental records to analyse climatic variability in Fujian Province of China during the last 1000 years.

In this paper, first, quantitative level or standard of documentary data are determined. Second, periods of wet, drought, cold, and warm of Fujian climate in documentary period and this century are divided. After that, period of climate variability, correlation coeffcient between wet/drought and cold/warm, and trend of urban heat island intensity of Fuzhou are calculated. Finally, feature and trend of climate variability in Fujian are discussed.

1. The dry and wet periods

a. period of documentary data

For quantitative analysis of documentary data, the documentary wet and dry events are divided into four levels of drought and four levels of wet in Table 1.

Table 1. Level of drought and wet

Level	Documentary date
−4	Continuous drought during summer and spring (or autumn) in the east and the west
−3	Local continuous drought during summer and spring (or autumn) or seasonal drought during spring (or summer or autumn) or severe drought in the east and the west.
−2	Local severe seasonal drought during spring (or summer or autumn) or severe drought during winter in the east and the west.
−1	Local drought during spring (or summer or autumn) or severe drought during winter.
0	No drought or flood
1	Local flood
2	Flood in the east and the west or local continuous rain (rain period is longer than one month) during late autumn or early spring or local contiinuous rain during summer
3	Local continuous rain (rain period is longer than one month) from late spring to early autumn or local flood during two months
4	Continuous rain (rain period is longer than one month) from late spring to early autumn or flood during two months in the east and the west

Using standards of Table 1, ten-year mean levels of wet/drought are calculated. Figure 1 shows the ten-year mean levels of wet/drought in the west (to the west of centre mountain range) and the east (to the east of centre mountain rang) of Fujian.

From 980 to 1369, there were several common wet or dry periods in the west and the east of Fujian. The wet periods were 980—1039, 1130—1149 , and 1170—1209, The dry periods were 1090—1109, 1160—1169, 1210—1259, and 1330—1359.

The climate have become drier than before in the east of Fujian since 1410. The frequency of drought in the east is more than that in the west. The continuous period of drought in the east in longer than that in the west. Therefore their conditions of wet/ drought are different. In the east of Fujian, the dry periods were 1410—1419, 1450—1489, 1500—1549, 1580—1599, 1650—1709, 1730—1749, 1780—1789,1810—1839,1860—1879,and 1893—1899 and the wet periods were 1600—

Corresponding author address: Prof. Lin Zheng—yun, Department of Geography , Fujian Normal University, Fuzhou 350007, *Fujian, China.*

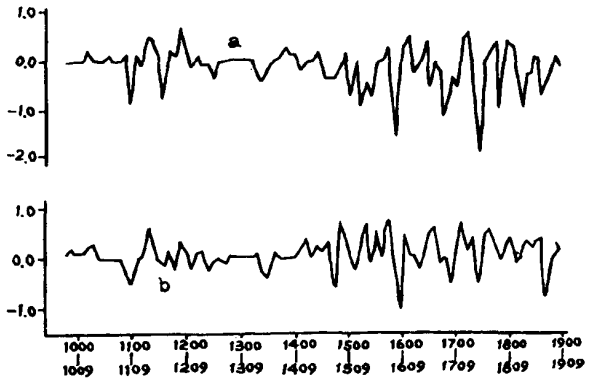


FIG. 1 The ten-year mean levels of wet/drought in the west (curve b) and the east (curve a) of Fujina.

1649, and 1790—1809. In the west of Fujian , the dry periods were 1470—1497 1580—1599, 1630—1639, 1670—1699, 1740—1749, and 1870—1889 and the wet periods were 1410—1469, 1480—1499, 1520—1579, 1600—1609 1640—1669, 1700—1739, 1750—1869, and 1890—1897

b. This century

Instrumental measurements of precipitation are available for the last 100 years in the east of Fujian. In the west of Fujian, instrumental measurement began in the 1930s. So we can use the records of annual precipitation to analyse the variability of condition of wet/drought. The results of analysis show that the dry periods are 1900—1902, 1925—1930, 1954—1957, 1963—1971, and 1976—1987 and the wet periods are 1903—1924, 1931—1953, 1958—1962, and 1972—1975 in the east of Fujian. In the west of Fujian, the dry periods are 1900—1910 (according to documentary data), 1938—1940, and 1955—1985 and the wet periods are 1911—1922 (according to documentary data), 1931—1937, and 1941—1954.

Figure 2 shows the ten-year running means of annual precipitation in Fuzhou (the east of Fujian) and Pucheng (the west of Fujian) .

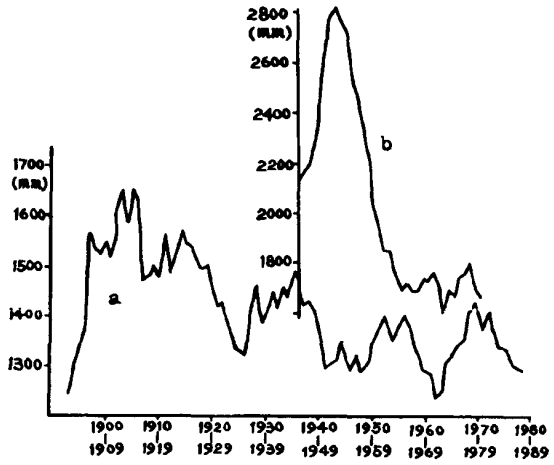


FIG. 2. The ten-year running means of annual precipitation in Fuzhou (curve a) and Pucheng (curve b) .

2. The cold and warm periods

a. Period of documentary data

Many cold events and a few warm events were put down in writing in documentary data. For quantitative analysis of documentary data, annual mean temperatures that correspond with cold or warm events are determined. These quantitative standards are given in Table 2. The corrections of annual mean temperatures that are affected by continuous rain or drought are also given in Table 2.

Using standards and corrections of Table 2, anomalies of ten-year mean temperature in Fujian from 1110 to 1899 are calculated. Then cold periods and warm periods are divided. The cold periods were 1110—1199, 1350—1359, 1500—1519, 1560—1569, 1640—1679, 1720—1739, 1750—1769, 1790—1819, 1830—1849, and 1890—1899. The coldest period is 1640—1679 in Fujian during the last 1000 years. Two long warm periods were 1200—1349 and 1360—1499. The temperatures in the two periods are higher than now. Figure 3 shows the anomalies of ten—year mean temperature in Fu-

jian from 1110 to 1900. As the documentary data of cold events had been rare before 1501, the curve is dashed line .

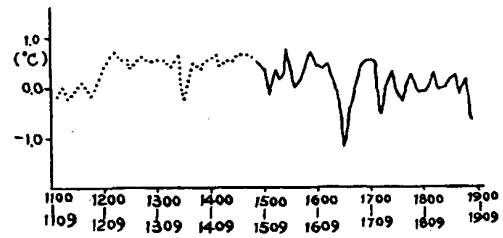


FIG. 3 The anomaly of ten-year mean temperature in Fujian.

Table 2. Quantitative standard and correction of annual mean temperature.

Annual mean temperature ℃	Documentary data
14.0	River is frozen
15.0	Stream is frozen or continuous frost (40—50 days) during winter or frost in July or August
16.0	Frost in June
17.0	Frost in May or birds and fish are frozen to death during autumn or glaze during winter
17.5	Crop damaged by frost during autumn
18.0	Snow depth≥1m or days of snow cover ≥4 days
18.5	3 days heavy snow or 0.3m＜snow depth＜1m or 3 days snow cover
19.0	1 day heavy snow or 2—3 days snow
19.5	North line of planting Litchi in large area
20.0	1 day snow
Correction ℃	Continuous rain during summer
−0.5	
+0.5	Severe drought during summer or continuous drought from spring to summer (or from summer to autumn) or thunderstorm during winter in the south Fujian except south of Quanzhou
+1.0	Thunderstorm during winter in the north of Fujian

b. This century

The east of Fujian in temperature trend is similar to the west of Fujian. Therefore the temperature variation in Fuzhou (the east of Fujian) represents the temperature variation in Fujian.

Figure 4 shows the ten-year running mean temperature in Fuzhou from 1903 to 1989.

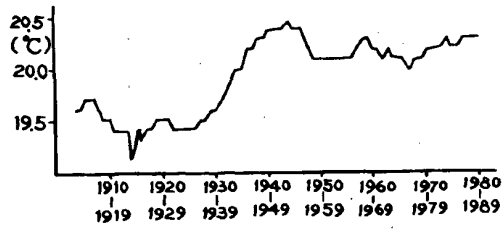


FIG. 4 The ten—year running mean temperature in Fuzhou from 1903 to 1989

From the end of last century to 1915 , tenperature had risen about 1℃ in Fujian. 1917and 1918 were cold years. From 1918 to 1946 , temperature had risen 2— 3℃ in Fujian. Up to 1989, the warmest year is 1946 and the warmest decade is the 40s in Fuzhou in the present century. 1976 was also a cold year. In general, temperature had risen from 1976 to 1991 in Fujian. But temperture in 1992 will be lower than that in 1991 in Fujian.

c. Spectral analysis

Anomalies of ten-year mean temperature in Fujian are studied by spectral analysis. The results show that there are periodic variations of 87—year and 24 —year (weak) .

d. Urban effect

The development of Fuzhou have affected urban temperature. We calculated trend of urban heat island intensity of Fuzhou. The calculation shows that mean heat island intensity of July-September have risen about 0.1℃ per ten-year.

e. Correlation coefficient between cold/ warm and wet/

drought

Using the documentary data from 1499 to 1899, the correlation coefficient between cold/warm and wet/ drought is calculated. The correlation coefficient is −0. 3.

3. Discussion and conclusions

a. From the 12th century to the 19th century, climate of Fujian is divided into warm period (from the 13th century to the 15th century) and cold periods (the 12th century and from the 16th century to the 19th century) . There is a periodic variation of 700 —year in Fujian climate. This century is the beginning of next warm period. According to this conclusion, the 21th century and the 22th century will be warm periods.

b. Besides the periodic variation of 700—year, there are periodic variations of 87-year and 24-year in Fujian climate. In Fujian, climate have become warm since the 1970s. It is caused by two factors. The first factor is the natural variations (different time scale) of climate. The second factor is man's action.

c. There is negative correlation between rainfall and temperature during the last 500 years in Fujian.

REFERENCES

Wang Risheng , Wang Shaown, 1990: Reconstrution of winter temperature in east China during the last 500 years using historical documents. *Acta meteorological Sinica* , *Vol* 48, *No* 2.

Lin Zheng-yun , 1992: An analysis of the urban heat island effect in southeastern coastal cities of Fujian Province. *Journal of Fujian Normal Vniversity*, 8 (1), 109—114.

CLIMATE CHANGE AND ENVIRONMENTAL IMPACTS: ON THE DEVELOPMENT OF A STATISTICAL METHODOLOGY FOR THE PRODUCTION OF A 500 MB HEIGHT FIELD SCENARIO

Arthur N. Samel
The University of North Carolina at Chapel Hill
Department of Geography
302 Saunders Hall
CB# 3220
Chapel Hill, NC 27599-3220
USA

Over the past decade there has been increasing evidence that links increased atmospheric greenhouse gas concentrations to a projected increase in global temperatures. A major concern associated with a greenhouse gas induced climate change is the potential for environmental impacts. The purpose of this research is to develop a statistical methodology that allows for the production of a scenario that projects changes in the relative frequency distribution of specific 500 mb height field patterns that are associated with the occurrence of regional ozone (O_3) events over the northeast United States.

Ozone concentration data from the Environmental Protection Agency (EPA) Office of Air Quality and Planning Standards (OAQPS) are used to identify events during the period 1980 to 1989 for a 21-state region of the northeast. An event is defined as a period of at least two consecutive days when O_3 concentration values at a minimum of 15 sites exceed EPA defined threshold values. National Meteorological Center (NMC) 500 mb height data for the period 1980 to 1989 are used to identify patterns that correspond to event days. A subjective identification scheme is then used to identify a set of observed 500 mb "characteristic patterns" that are found to be strongly linked to the occurrence of northeast O_3 events.

A modified Kirchhofer sum of squares analysis is used to develop a climatology for the observed 500 mb height field. The standard Kirchhofer analysis identifies a set of common or "key day" patterns. The standard analysis is "modified" such that the key day patterns are prescribed to be the characteristic patterns developed earlier. The modified Kirchhofer analysis performs a series of map pair comparisons whereby significant similarity between each pattern in the data set and the characteristic patterns is identified. For each map pair comparison, Kirchhofer grid, row and column scores are calculated. Each score represents the sum of squares differences between the heights in the map pair. The map pair is considered to be significantly similar if all scores are less than user defined threshold values. Each

significantly similar pattern in the data set is assigned to that characteristic pattern with which the similarity exists. If a pattern is found to be significantly similar to more than one characteristic pattern then it is assigned to that characteristic pattern with which it has the lowest grid score. After the analysis is completed, the climatology provides the relative frequency distribution of patterns that are significantly similar to the characteristic patterns.

The modified Kirchhofer analysis is then used to develop two additional climatologies for the atmosphere of a Geophysical Fluid Dynamics Laboratory (GFDL) general circulation model (GCM). Both GFDL GCM output data sets contain 10 years of daily averaged 500 mb height values. The GCM climatologies are developed for the set of characteristic patterns discussed above. The first climatology is produced for a control atmosphere where the prescribed concentration of CO_2 is roughly equal to current observed values. The observed and control climatologies are compared to determine the ability of the GFDL GCM to reproduce the observed frequency distribution of the characteristic patterns. The second climatology is developed for the GFDL GCM atmosphere where the concentration of CO_2 is double that prescribed in the control atmosphere.

The results of the observed, control and doubled CO_2 climatologies are compared in order to determine greenhouse gas induced changes in the frequency distribution of patterns that are associated with the occurrence of northeast O_3 events.

ON TRENDS IN THE ZAMBIAN PRECIPITATION MAIZE CROP INDEX: 1950-1990

P.D. Kruss[1], E. Mukhala[2] and M.R. Muchinda[2]

[1] World Meteorological Organization

[2] Meteorological Department of Zambia

The tropical regions of the world are home to a vast population which is often dependent on local agriculture for basic food supply. Short meteorological events and longer term climate change can and have drastically affected this food supply. However, in terms of greenhouse change, mean temperature increases as speculated in the literature are not in themselves sufficient to significantly affect the growing capability of a tropical country. Hence, when discussing the possible effects of greenhouse forcing, it is relevant to investigate the climatic record from the viewpoint of trends in agriculture-critical indices. This present paper addresses this issue with respect to trends during the last 40 years in a precipitation-based agroclimatic index of the Zambian maize crop.

A personal computer database of daily agroclimatic information is being created by the Zambian Meteorological Department with the support of a World Meteorological Organization/United Nations Development Programme project. Of particular interest to this present study are daily rainfall values available for the period 1950-90 at 35 Zambian stations.

The varieties of maize being grown in Zambia (e.g., MM501/2/4, MM600/1/3/4, MM752) possess various characteristics in terms of number of days to silking and maturity, yield levels, etc., with both precipitation and temperature parameters being important to final yields. In this present paper only precipitation-related parameters are considered. The critical precipitation variables include such as: length of wet season (e.g., 100 to 160 days for the above varieties); days of moisture stress during growth, particularly at the time of silking (e.g., silking occurs 60 to 75 days after planting for the above varieties and yield reduction may exceed 6% per day of stress.); etc. Based on such considerations, separate precipitation-based maize crop indices have been constructed for each of the three primary agroecological zones of Zambia.

It was found that, over the last forty years, significant change in the ability of the country to support maize production has occurred in the southern-most agroecological zone. While the two northern zones remain basically unaffected, a conclusion difficult to reach based on conventional analyses of precipitation amounts alone. The southern zone of Zambia (the driest zone) is now significantly less well able to support the varieties of maize grown there, with the most rapid change occurring prior to the mid 1970's. This study highlights the importance from the agricultural and human-impact perspective of reaching beyond standard meteorological analyses when considering the significance of any greenhouse-induced climatic change.

MONITORING AND ESTIMATION OF THE TRENDS OF TEMPERATURE-MOISTURE CHARACTERISTICS OF THE BASIC GRAIN-PRODUCING ZONE IN THE USSR FOR 100 YEARS

Meshcherskaya A.V., Blazhevich V.G.

There are analyzed series of precipitation, temperature, aridity indices, and snow cover depth for the period of 1891 to 1990. The analysis is based on the extensive initial information published in four handbooks - monographs. The distinctive feature of the published data is their averaging over the area of 104 administrative regions, which fact allows us to decrease the space-time variability and to smooth random measurement errors.

To average meteorological values over the area of administrative regions, the results of precipitation observations were used from 1400 stations and snow cover observations from 287 meteorological stations.

The air temperature was not averaged over the area due to its sufficiently high representativity related to the low space variability. The air temperature is presented by observation results from 104 meteorological stations.

When preparing the initial information much attention is given to eliminating the non-uniformity in the data series used.

Monitoring of the above meteorological characteristics has been made separately for European (ETU) and Asian (ATU) Territories of the USSR within the basic grain-producing zone. This is explained by the difference of meteorological processes in ETU and ATU which are frequently in antiphase (excessive moisture in ETU corresponds to droughts in ATU and vice versa). For the same reason meteorological series for the cold (XI-III) and warm (IV-X) seasons of the year were examined separately.

The means ($\bar{X}$) for ETU and ATU values of meteorological quan-

tities for the cold and warm season in each particular year have been obtained by averaging the mean region monthlies of meteorological quantities with due regard for the area of regions using the following formula:

$$\bar{X}_n = \frac{1}{l} \sum_{j=1}^{l} \sum_{i=1}^{m} \frac{P_i}{P} X_{ij} \qquad (1)$$

where X_{ij} is the value of meteorological quantity X in the i-th region in the j-th month; m is the number of administrative regions; l is the number of calendar months; P is the total area of regions included in the territory of ETU or ATU.

a. Analysis of temperature series and their linear trends has shown that in ETU the temperature increased by 1°C in the cold season and by 0.3°C in the warm season of the year. The temperature growth in ATU was 1.4°C and 0.5°C respectively. The linear temperature trend in ATU in the cold season is statistically significant with 5% level of significance.

The contribution of anthropogenic temperature growth due to urbanization was estimated from literature data and the authors' calculations.

b. Trends in precipitation series, unlike temperature, are expressed weakly. The precipitation series in ETU in the warm season are practically stationary, and in the cold season precipitation increases slowly with the rate of 0.18 mm/yr. In ATU for 100 years precipitation decreases with the rate of 0.12 mm/yr in the warm season and 0.16 mm/yr in the cold season of the year.

c. The series of snow cover depth in late February in ETU and ATU have much in common, which is reflected in the nature of trends obtained by second-degree polynomials approximation of the initial series.

The period of winters with much snow observed in the late XIX-

early XX century changed in the 30es to the period of winters with little snow. Since the late 80es of the XX-th century the snowfall in winters has increased again.

There are discussed tendencies in the change of snow cover depth due to the change of atmospheric circulation type.

d. Monitoring has been made of summer aridity indices which are the first factor (a_1) of expansion in empirical orthogonal functions of value S:

$$S = \frac{\Delta T}{\sigma_{\Delta T}} - \frac{\Delta Q}{\sigma_{\Delta Q}}, \qquad (2)$$

where ΔT and ΔQ are temperature and precipitation anomalies, and $\sigma_{\Delta T}$ and $\sigma_{\Delta Q}$ are standard deviations corresponding to them.

It is shown that aridity growth is observed in the two territories which is particularly active in ATU. The value of the trend of a_1 of index S in the late 80es was 0.5 of σ_S in ATU and 0.25 of σ_S in ETU.

Checking of the trends being studied for stationarity by the method of series has shown however that the two series do not contradict the hypothesis on stationarity with 95% probability.

The extremely dry and extremely wet years in ETU and ATU have been examined. It has been found out, in particular, that 1991 is second in the 100-year observation series after 1955 in drought intensity in ATU.

A STUDY FOR SUMMER DRYNESS/WETNESS PROBABILITY IN NORTH OF CHINA

Huang Jiayou

Department of Geophysics, Peking University

Beijing,100871

(P.R.China)

Abstract

Using the data of summer (June - August) amount of precipitation during the period of 1951-1985 in eight stations in North of China, the wetness probability series based on Gamma distribution have been built. Main temporal and spatial chracteristics for the series of WP in the stations are extracted using empirical orthogonal functions (EOFs). The spatial patterns show that the dryness/wetness drought variations in the stations in the area are consist with each other and can be representative of main chracteristics by the WP series in Beijing.

In order to detect the response of dryness/wetnes in Beijing to the global warming, the correlation coefficient (r=-0.159) between the 120-yaer series of WP and the Northern hemispheric temperature shows that there is a closely relationship between the dry state and the global warming in the area. It surpass the statistical test at 5% level and is significant.. In fact, the dry tendency with the slope of -0.001 in the regression line of the WP series can be found.

Furthermore, the respose of the WP to the global warming is investigated in different climtic period. Usually, the climatic period cover about 30 years. Te comparisons in occurrence frequence of dryness/wetness grade derived from the WP series. It shows that the occurrence frequencies of drought was increasing with years. The seious drought also was represented in decreasing the occurrence frequencies of flood, especilly in very wet state.

The variation of the dryness/wetness phenomena may be relative to the the global temperature, especially the sea surface temperatures in Pacific ocean. The drought in the area associated with the extremely warming values of SST.

The persistence analysis of dryness/wetness in years has been completed using consecutive event probability.

The comparisions of the persistence of dryness/wetness (see

Table 5) in the consecutively two and three years in fifferent climatic periods show that the dry persistence has a increasing tendency..The tendency of the dry persistence was more stronger than the wet, and serious in recent 30 years.

In order to emphasize the natural pattern of periodic variation in data, the periodicity analysis has been completed using period pattern method, which is proposed in this paper, rather than the traditional methods(e.g. spectral analysis or variance analysis). The 120-year series of WP in Beijing is aranged in a two dimension matrix, in which the number of the column equal the year number of the test period. The test period ranges 2 to 60 years. The EOFs can be done for the matrix. The period patterns are etracted from the eigenvectors of the crossproduction matrix. The significance of the period pattern is depend on the significance of the correspoding component. According to North et al.(1982). the statistical tests on the components show that only the first components in the test periods are significant. The significant degree can be represented by the difference between the fisrt and second eigenlavue.

The significant period patterns can be selected from the period corresponding to the peak points in the differnce curve. They should have greater value (for example reach about 31.0) and be corresponding to higher explained variance. They can reflect the homogeneities of the test periods. Those period patterns can be representive of the main chracteristics of the variation of the corresponding period. They are corresponding to the periods of 2-,5-,7-,20-,24-,30-,40- and 60-year.

The rearranging series during 120 years are formed according to the period patterns are built. Under the supposition of the linear relationship between the series of WP and the new series, The WP can be estimated by a regression equation. The procedure are completed by stepwise regression with control value(F=3.0).

The wetness probability in the future ten yaers are estimated. If dyyness/wetness states forecast using the WP value greater/less 0.5, The accuracy reach 0.85 for correct forecasts. In the independent sample, in 1990, with the observed precipitation of 384mm and WP of 0.043, the estimation of dry year is correct. In future, the drought may be expected to relax in 1991-1994 and 1998-1999.

GLOBAL WARMING AND RAINFALL VARIABILITY - THE SRI LANKAN SITUATION

by

T. K. Fernando and L. Chandrapala

Department of Meteorology, Sri Lanka.

Introduction

Agro-ecologically, the Island of Sri Lanka which has an areal extent of about 65,610 square kilometres, is divided into three zones, namely, the wet zone, the intermediate zone and the dry zone.

The annual mean surface air temperature, the annual rainfall and the number of rainy days per year, recorded over a period of more than 100 years at 14 meteorological stations scattered throughout the Island have been analysed by filtering the time series using the Gaussian low-pass filter.

A linear regression analysis has been done in the case of mean surface air temperature and the annual rainfall, for the more recent period 1961 - 1990 in most cases.

The locations of the 14 stations are shown in fig. 1. The stations at Colombo, Ratnapura, Galle, Nuwara Eliya and Kandy are located in the wet zone; those at Jaffna, Trincomalee, Anuradhapura, Batticaloa, Puttalam, Mannar and Hambantota are in the dry zone while the remaining two stations at Kurunegala and Badulla are in the intermediate zone.

Data

The Department of Meteorology, Sri Lanka maintains 22 meteorological stations which are scattered throughout the island. The 14 stations mentioned earlier were selected on the basis of period of data, elevation of the stations and representativeness. The station at Puttalam is at the lowest elevation of 2.1 m, while that at Nuwara Eliya has the highest elevation of 1894.6 m above mean sea level.

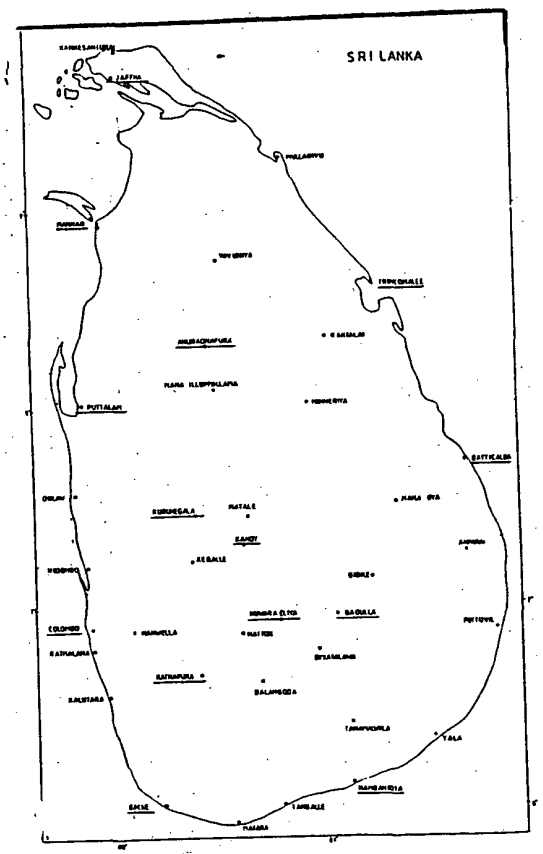


Fig.1 - Locations of stations (underlined)

Time Series Analysis

I. Surface Air Temperature

The time series of annual mean surface air temperature anomalies, relative to the average for 1951 - 1980, have been analysed by filtering the series using the 9 - point Gaussian low-pass filter. The resulting series indicate both increasing and decreasing trends. When the more recent period was considered, however, we found that the surface air temperature has been increasing at all the 14 stations considered. This increase has started as early as 1910 in the case of Ratnapura and as late as 1955 at Colombo, Galle, Kandy and Batticaloa. Increasing trends at the other 9 stations have commenced between 1910 and 1955.

II. Rainfall and Rainy days

The time series of annual rainfall when filtered using the Gaussian low-pass filter showed decreasing trends, in the recent past, at Colombo, Ratnapura, Galle, Kandy, Batticaloa, Hambantota, Trincomalee, Badulla and Nuwara Eliya.

Trends shown by the analysed time series of rainy days per year agree closely with those shown by the rainfall series.

Linear Regression Analysis

A linear regression analysis was done in the case of annual mean surface air temperature anomalies and annual rainfall for the more recent period, 1961 - 1990, in most cases. Table 1 shows the results obtained.

Selected graphs of filtered series of surface air temperature anomalies and annual rainfall are shown in fig. 2 and fig. 3 and maps of Sri Lanka showing isolines of increase in surface air temperature anomalies and variation in annual rainfall (with the R^2 - values given in parenthesis) are shown in fig. 4 and fig. 5 respectively.

TABLE 1

LINEAR REGRESSION ANALYSIS

STATION	TEMPERATURE Slope (°C/yr.)	R^2	RAINFALL Slope (mm/Yr.)	R^2
Ratnapura	+0.0175	0.88	-4.98	0.14
Badulla	+0.0217	0.85	-9.14	0.41
Batticaloa	+0.0114	0.83	-34.29	0.73
Trincomalee	+0.0189	0.83	-17.94	0.52
Hambantota	+0.0104	0.81	-12.77	0.91
Anuradhapura	+0.0364	0.79	+13.11	0.52
Kandy	+0.0185	0.72	-10.28	0.72
Colombo	+0.0164	0.67	-35.28	0.89
Galle	+0.0196	0.65	-20.52	0.69
Puttalam	+0.0106	0.64	+16.76	0.50
Jaffna	+0.018	0.61	+0.40	0.00
Nuwara Eliya	+0.0146	0.56	-4.91	0.20
Kurunegala	+0.0173	0.42	-7.30	0.20
Mannar	+0.0118	0.39	+9.18	0.45

Comments, Conclusions and Summary

Although factors such as relocation of stations, urbanization effects, exposure and height of instruments would have had some effect on the temperature and rainfall readings, it was assumed that such factors were unlkely to have had an effect on the trends.

This study revealed the following:

1. Surface air temperature has shown an increasing trend at all the 14 stations during the more recent period.
2. The annual rainfall has shown a

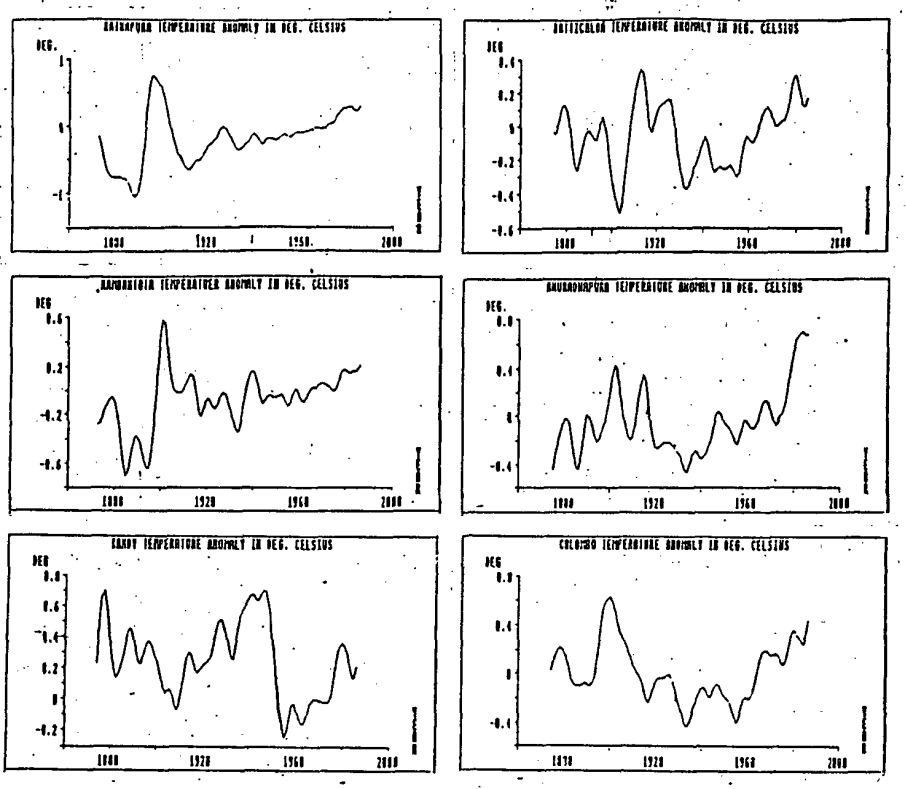


Fig.2 - Time series of filtered surface air temperature anomalies

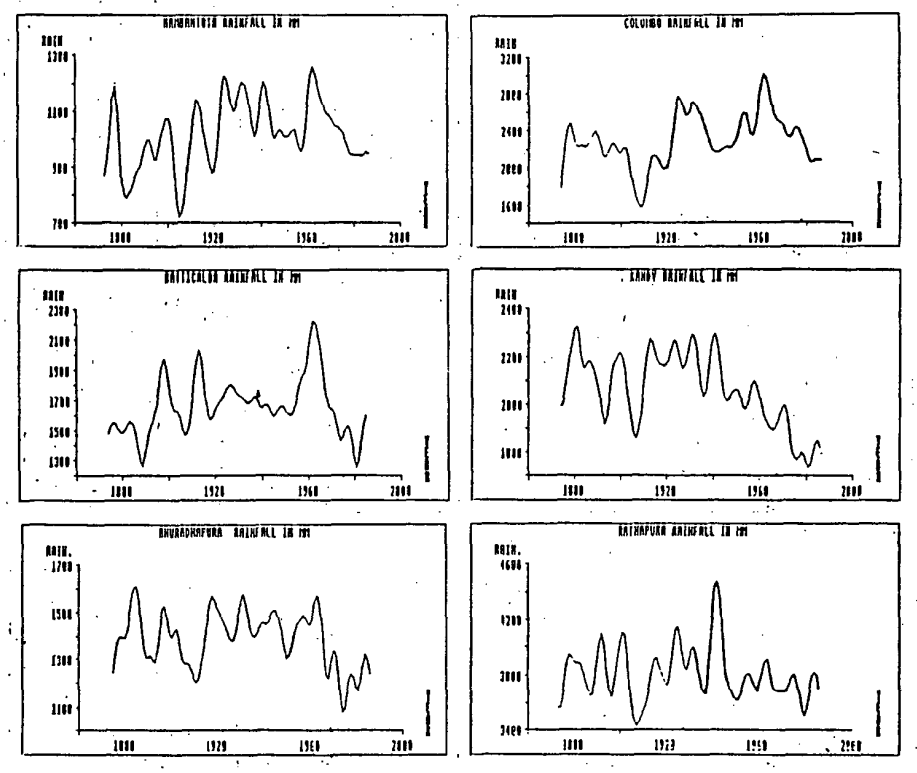


Fig.3 - Time series of filtered annual rainfall

decreasing trend, during the more recent period, at Colombo, Ratnapura, Galle, Kandy, Batticaloa, Hambantota, Trincomalee, Badulla, and Nuwara Eliya.

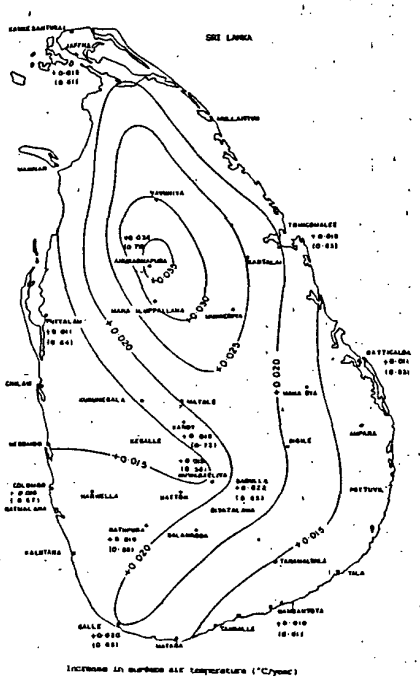


FIG. 4

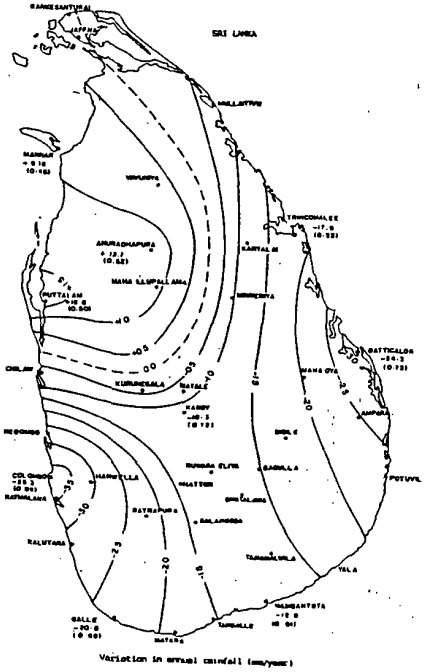


Fig.5

3. The surface air temperature has been rising at the rate of 0.036 deg C per year at Anuradhapura (the highest) and 0.010 deg C per year at Hambantota (the lowest), during the period 1961 - 1990.

4. The annual rainfall has been
   (a). decreasing at the rate of 35.3 mm per year at Colombo (the highest) and 10.3 mm per year at Kandy (the lowest), the period of anlysis being 1961 - 1990.
   (b). increasing at the rate of 16.8 mm per year at Puttalam (the highest) and 9.2 mm per year at Mannar (the lowest), the period of anlysis being 1975 - 1990.

Acknowledgements

The authors wish to thank most sincerely, Dr. A. W. Mohottala, Director of Meteorology, Sri Lanka for granting permission to use the Departmental computer facilities.

References

1. Agro-ecological regions of Sri Lanka, Land and water use division of the Department of Agriculture, Sri Lanka, 1979

2. WMO, 1966; Climatic Change, WMO Technical Note No. 79.

CHANGES IN ZIMBABWE'S RAINFALL REGIME AND THE EFFECT ON CLIMATOLOGICAL MEANS ('NORMALS')

LEONARD S. UNGANAI
Dept. of Met. Services, Box BE150, Belvedere, Harare, Zimbabwe.

ABSTRACT

This paper presents evidence of changes in Zimbabwe's rainfall pattern over the period 1896-1990. The effect of the changes on climatological rainfall means is addressed and possible shifts in the rainy season investigated.

Introduction

From the late 1940s, the frequency, persistency, areal extent and severity of drought episodes in Southern Africa has increased (Lewis and Berry, 1988). Still vivid in most people's minds are the meteorological droughts of 1946/47, 1967-73, 1982/83 and 1991/92 and their effects on agriculture, surface water supplies, the environment and the economy. Some parts of Zimbabwe never had more than three good seasons during the decade 1980-90. The question many would ask is whether these climatic extremes are a result of climate change, and if they are, what has been their effect on climatological averages ('normals')?. Rainfall averages are sensitive to the time period from which they come, should there be a sustained shift towards a drier or wetter regime as Todorov (1984) illustrated for the Sahelian region where some of the 'normals' became obsolete and uninformative as a result of persistent drought. In this paper precipitation records for the years 1896-1990 are analyzed in order to assess whether Zimbabwe's rainfall pattern has changed, and if it has what the effects have been on the climatic normals and the bounds of the rainy season. The methodology used here draws heavily on work by Todorov (1984) and Nicholson (1983).

DATA:

For the purposes of this study, monthly and annual rainfall totals for 13 meteorological stations with records stretching as far back as 1896 in some cases, were used. The stations were selected in such a way that as long a time frame as possible be covered, whilst at the same time representing different rainfall regimes in the country.

Table 1 is a list of the stations used and fig. 1 shows their geographical distribution. To avoid portraying an artificial climate change, only those stations whose sites have not moved to more than 5km away from the original location were used.

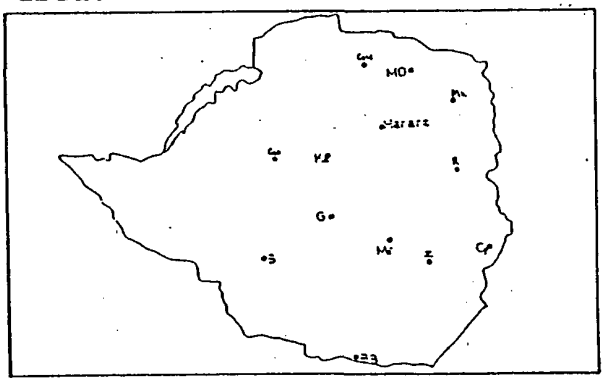


figure 1. Stations used

ANALYSIS:

Five year running means were plotted to find evidence of changes in the rainfall

pattern over the observation period. The same technique was used to investigate possible shifts in the start and/or end of the rains. Rainfall means for the time periods 1920-50, 1930-60, 1940-1970, 1950-1980, 1970-1990, 1980-1990 and 1920-90 (where the meteorological year runs from July to June of the following year) were computed so as to establish the effects of the changes in the rainfall regime. Zimbabwe's rainy season is bound by November(start) and March(end) (Climate handbook of Zimbabwe,1983), a decline in the total amount of rainfall in November would therefore imply a delay in the start of the rains , whereas an increase in rainfall for March would indicate a later than normal end of the season and the converse is also assumed true, a technique adopted from Todorov(1984).

The significance of the difference between any pair of means was tested by one non-parametric technique, the two-sample Wilcoxon test for independent data and the t-test using the Instat statistical package (at 95% significance level).

Results and Discussion

Figures 2(a-j) show that Zimbabwe's rainfall pattern has changed from about the 1950s. The observed changes have been different from one part of the country to another, with the northern areas of the country showing a decline in total annual rainfall as well as its inter-annual variability (fig 2b).

An antitheses of northern Zimbabwe is the south-eastern districts of the country (Chipinge, Masvingo and Beitbridge), middle and western Zimbabwe (Bulawayo, Gweru and Gokwe areas). In these areas,

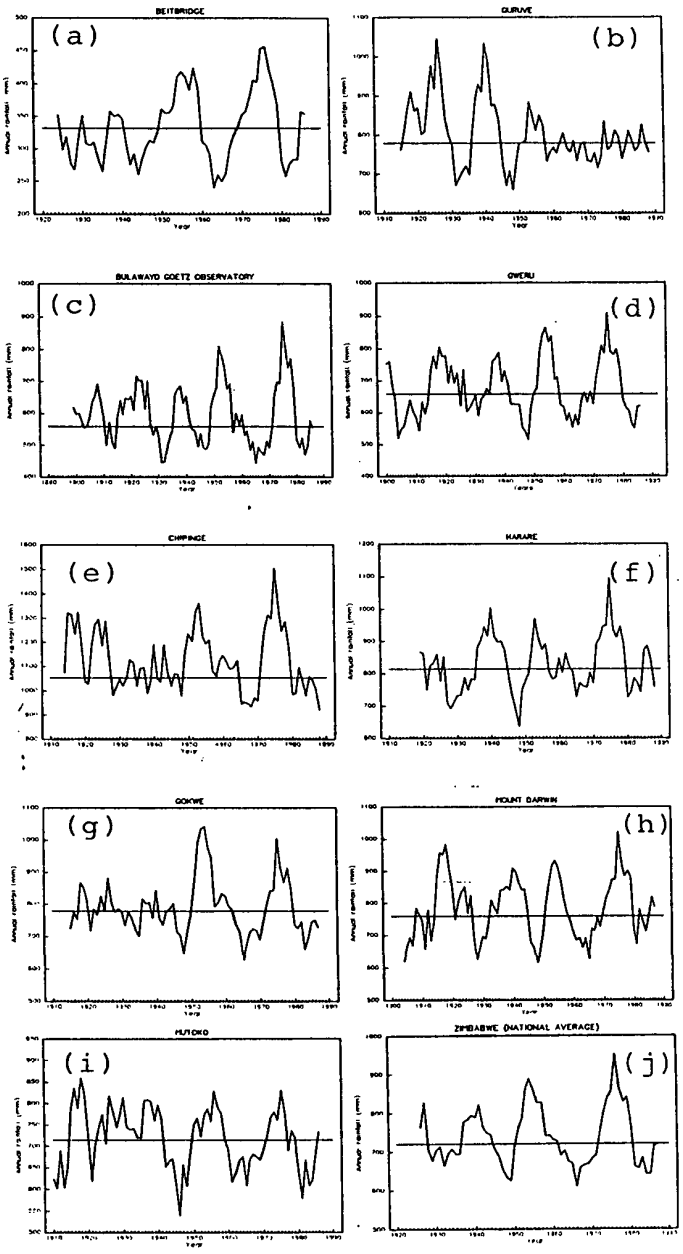


Figure 2. Five year running means on annual rainfall.

before the 1950s rainfall was generally less variable from season to season albeit low (fig. 2a,-c,-d,-e,-g). The results of this study also show that from the 1950s, the chances of extremes in weather such as very dry/very wet seasons has increased in the

South-eastern and midlands districts of Zimbabwe.

For the rest of the stations whose rainfall records were analyzed, no unusual shifts in the rainfall pattern were found. Fluctuations in the Harare/Mutoko/Mount Darwin areas(fig. 2f,-h,-i) are very much in line with the normal expected climatic oscillations, with some quasi-cyclicity in the pattern quite evident.

Aggregating all the rainfall data for the country results in the graph portrayed in fig.2j, which suggests that the whole country has experienced a shift in the rainfall pattern. It is worth noting that valuable information is lost by this type of averaging because the nature of the shifts have been different from one part of the country to the other, a situation that cannot be shown by a national average, unless there was a single rainfall regime in the country at any one point in time.

Shift of the rainy season

March which normally marks the end of Zimbabwe's rainy season has in the past been very wet for a series of years then turns dry for some time before wet conditions recur, as figure 3 clearly shows. The trend is however not well defined although there is slight indication that the 90s might experience wetter than normal conditions in March resulting in a late end of the season. Bulawayo and Chipinge provide an interesting picture of the trend in the total rainfall for March(fig. 3a,-d). In these areas, before the 1930s March was wetter than the period upto the 1990s.

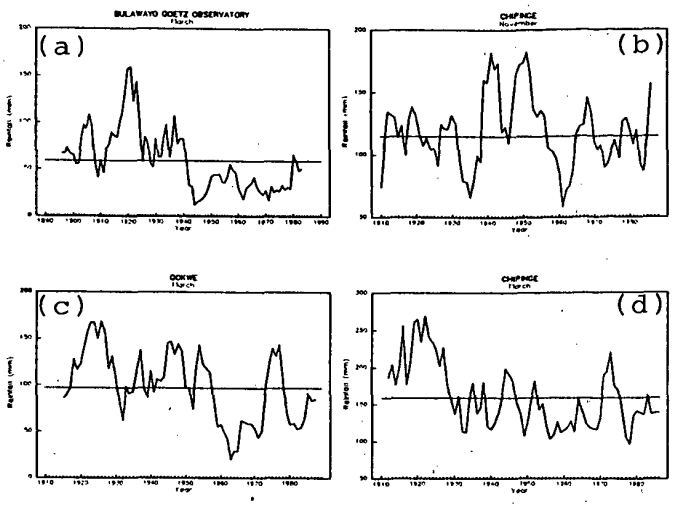


Figure 3. Five year running mean on monthly rainfall.

Variations in the November rainfall show no well defined pattern (fig 3b). The fluctuations over the observation period are typical of the expected climatic oscillations.

Rainfall Normals Used:

Tables 1 and 2 show annual and monthly rainfall means for different time periods from the 1920s. Although the means from different time periods appear different from each other, the t-test and the two-sample Wilcoxon independent data test showed that none of the differences in annual means are statistically significant. However, the March rainfall means in the Zaka, Bulawayo, Chipinge, and Gokwe areas were sensitive to the shift in the March rainfall pattern, with the 1920-50 mean being significantly different from the 1960-90 one. These results show that for the means to be affected significantly the change in the pattern has to be large and sustained for a long period of time as was the case in the Bulawayo March rainfall (fig. 3a).

Table 1: Stations used and the annual rainfall means (mm)

Station	Alt.(m)	Lat.	Long.	1920-90	1930-60	1940-70	1950-80	1960-90	1970-90	1980-90	1920-50
				70yrs	30yrs	30yrs	30yrs	30yrs	20yrs	10yrs	30yrs
Beitbridge	456	22°13′	30° 0′	331.6	344.6	336.0	338.1	331.8	355.9	305.4	331.6
Bulawayo	1343	20° 9′	28°37′	593.1	585.8	569.3	630.1	583.6	620.9	530.5	583.0
Chipinge	1131	20°12′	32°37′	1118.2	1110.0	1092.0	1157.0	1105.0	1148.8	1050.9	1102.7
Gokwe	1282	18°13′	28°56′	794.5	800.7	787.4	825.8	773.7	785.9	736.6	764.4
Guruve	1177	16°39′	30°42′	802.3	796.4	784.7	775.1	771.6	779.1	796.2	825.5
Gweru	1428	19°27′	29°51′	670.5	679.6	654.8	716.5	667.1	682.7	618.4	657.5
Harare	1506	17°48′	31° 3′	833.2	826.3	814.7	854.4	844.6	861.1	811.4	805.5
Kadoma	1149	18°19′	29°53′	741.7	766.8	732.7	769.9	732.2	767.9	729.4	722.5
Makoholi	1204	19°50′	30°47′	658.9	643.7	656.2	701.6	652.5	673.8	572.2	642.9
Mt. Darwin	965	16°47′	31°35′	790.8	797.9	768.2	801.5	783.0	828.4	787.8	770.5
Mutoko	1244	17°25′	32°13′	721.2	728.4	690.8	689.1	689.9	704.5	676.7	724.8
Rusape	1430	18°32′	32° 8′	825.5	837.9	808.2	791.4	791.4	804.5	750.4	831.5
Zaka	774	20°20′	31°28′	756.5	753.8	739.8	745.2	745.2	757.4	642.0	741.3

Table 2: March rainfall means (mm)

Station	1920-90	1920-50	1950-90	1960-90	1970-90	1980-90
Beitbridge	37	43.2	31.1	31.7	35.6	24.9
Bulawayo	51.7	80.0	35.5	36.2	34.4	44.0
Chipinge	150.1	185.1	135.9	139.6	145.2	140.0
Gokwe	95.7	125.5	78.1	72.7	83.0	69.0
Guruve	114.8	139.3	94.4	99.8	101.0	90.4
Gweru	73.3	97.8	65.3	55.9	57.5	43.4
Harare	103.1	116.8	97.2	103.3	121.3	111.7
Kadoma	89.3	114.5	77.9	74.8	83.8	77.3
Makoholi	79.1	112.6	116.0	62.2	70.3	61.5
Mt. Darwin	104.6	124.4	89.8	90.2	92.8	95.3
Mutoko	92.4	112.2	76.1	72.6	74.5	71.6
Rusape	94.9	124.0	76.5	80.1	81.9	65.3
Zaka	102.2	139.4	75.5	76.2	76.1	59.9

CONCLUSION

From this study, there is strong evidence that some parts of Zimbabwe have had shifts in their rainfall regimes. However, the nature of the shifts have been such that the long-term climatic averages (reference means) remained virtually unaffected. The shifts in the rainfall pattern maintained the sine wave form albeit with higher and deeper amplitudes than before the 1950s. A comparison of the means from the period 1920-50 (old regime) and 1960-90 (new regime) table 1, also shows that the quasi-cyclical nature of Zimbabwe's annual rainfall tends to mask the effect of the observed climate shift.

Very wet episodes have alternated with droughts more severe than before fairly regularly. The causes of such a trend, whose start coincides

with the 1946/47 drought can only be surmised. Deforestation, increase in atmospheric air pollution, sunspots, volcanic eruptions coupled with the inherent climatic oscillations are some but a few of the possible causes (Hounam, et. al. 1975). From this study, there is no basis for saying whether the observed shifts in the rainfall regimes are an indication of a gradual and sustained climate change in Zimbabwe.

Evidence of a shift of the rainy season is not strong. Only the Bulawayo and Chipinge areas might have been having an earlier than normal end of the season as a result of the drying out of March, the trend is however showing signs of reversing. More study is needed in this area because total rainfall per se is inadequate to define the start and/or end of the rains.

For Zimbabwe, the time period from which the mean comes is therefore not very relevant, as long as the period is long enough, the mean will be appropriate for use as reference.

Acknowledgements:
Thanks to Memory Dhorobheni for some of the typing, and to Dr M.C Zinyowera, Mr T. Ngara and Wish Marume for the critique.

References:

Climate handbook of Zimbabwe, (1981). Zim. Dept. of Met. Services, Harare

Ngara. T, McNaughton, D.L and S. Lineham (1983). Seasonal rainfall fluctuations in Zimbabwe. Agr. journ. **80** 4, Harare.

Nicholson S.E (1983). Sub-Saharan Rainfall in the years 1976-80: Evidence of continued drought. Monthly weather review. **111, 1646-1654.**

Lewis L.A and L.Berry (1988) African Environments and Resources. Unwin Hyman, London.

Panofsky, H.A. and G.W.Brier (1968). Some applications of Statistics to meteorology. Pennysylvania.

Todorov, A.V. (1984): Sahel: The changing rainfall regime and the 'Normals' used for its Assessment. J. of Clim. and Appl. Met. **24** 2.

Tyson, P.D (1986). Climatic change and Variability in Southern Africa. Oxford Press, Cape Town.

Tyson, P.D.and Dyer T.G.J 1980. The likelihood of droughts in the eighties in South Africa. South African Journ. of Sci, 76,340-341.

DESICCATION IN THE LOWVELD AREA OF ZIMBABWE ?

by

WISH MARUME

Department of Meteorological Services P.O. Box BE 150, Belvedere, HARARE, ZIMBABWE

INTRODUCTION

Zimbabwe experienced the worst dry period since the turn of the century in the period between 1981 and 1984. Since then, dry conditions have persisted in the southern districts of the country; Matabeleland South, Masvingo and southern parts of Midlands. This has led people to debate the possibility of climate change and desiccation in some parts of the country. This paper presents results of some investigations carries out on annual rainfall for some stations in this area.

RAINFALL PATTERN OF THE LOWVELD

Mean annual rainfall in the southern districts decrease progressively from north to south (Fig 1), except for the elevated area stretching from the Great Zimbabwe ruins to Bikita where rainfall is influenced by orography. Masvingo in the north has a mean annual rainfall of 628.2+-220.1mm and Beitbridge in the south averages 335.7+112.0mm. The rainy season decreases progressively southwards too.In this area, rainfall is very variable and unreliable. To the west, the coefficient of variability, (defined as standard deviation as a percentage of mean annual rainfall), is greatest (>40) and the annual rainfall range is greater than 900mm in some places.Filabusi has an annual range of 929.5mm ,Gwanda 989.8mm, Zvishavane 968.6mm, Masvingo 1212.6mm, and Beitbridge 556.5mm. Much of the rainfall occurs between October and April , the rainfall season in Zimbabwe(Climate Handbook of Zimbabwe).

MATERIALS AND METHODS

Long term total annual rainfall data for six stations in the southen region of Zimbabwe (Figure 2) (from when each station was opened based on rain-gauge data read daily) was obtained from the Department of Meteorological Services. This data were used to determine the annual rainfall patterns and hence the period considered in the analysis varied from one station to another (Table 1). Filabusi and Masvingo had the longest record dating from as far back as the 1903/04 season. A year was defined from July of one year to June of the succeeding year to ensure that annual rainfall totals were calculated from one continuous season since the rainy season in Zimbabwe straddle two successive years. The season labels on the graphs are the year in which the season ended.

A transformation used by Lamb (1983) and Nicholson (1983) was used to calculate the standardised annual rainfall departures at each station: ie

$$x_{ij} = (r_{ij} - r_i)/s_i$$

where r_{ij} is the annual rainfall for the station i and year j, r_i is the annual average rainfall at station averaged over the period of record and s_i is the standard deviation of annual rainfall at station i. The areal integrated rainfall departure for year j was calculated from

$$R_j = (x_{ij})/N_j$$

where N_j is the number of stations available for station j. Only those years, from 1921, when observation were done at all stations were used in the calculation.

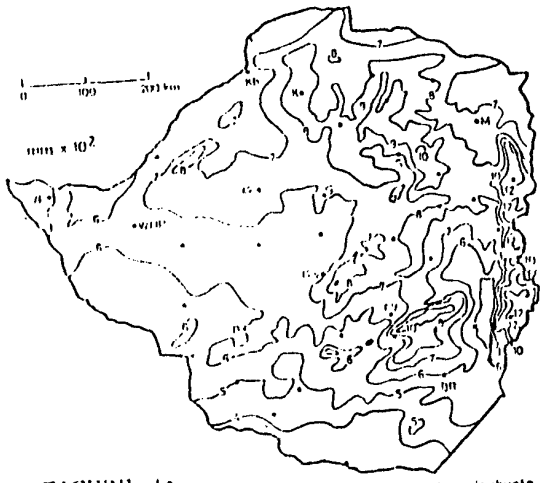


FIGURE 1: MEAN ANNUAL RAINFALL. Scale too small to show isohyets >1200 mm

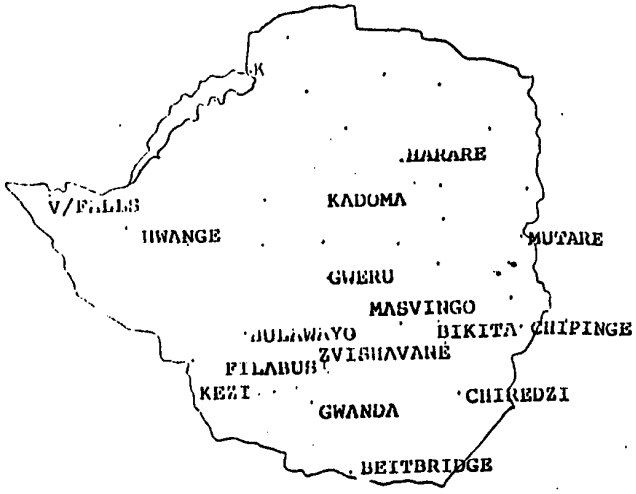


FIGURE 2: LOCATION OF STATIONS USED IN THIS ANALYSIS

Table 1: Station information.

Station	Latitude	Longitude	Altitude	Period
Filabusi	$20^0\ 33^1$S	$29^0\ 17^1$E	1070M	1903/04-90/91
Gwanda	$20^0\ 57^1$S	$29^0\ 01^1$E	990M	1910/11-90/91
Kezi	$20^0\ 55^1$S	$28^0\ 27^1$E	1015M	1920/21-90/91
Zvishavane	$20^0\ 19^1$S	$30^0\ 04^1$E	975M	1922/23-90/91
Masvingo	$20^0\ 04^1$S	$30^0\ 52^1$E	1095M	1914/15-90/91
Beitbridge	$22^0\ 13^1$S	$30^0\ 00^1$E	457M	1922/23-90/91

RESULTS

Station annual rainfall patterns are illustrated in figures 3(a)-(f). The patterns are more or less similar and do not show any trend. Instead the period of analysis can be divided into groups of years of relatively high rainfall, the period before 1930, 1951 to 1961 and 1971 to 1981, and low rainfall periods 1941-51, 1961-70 and 1982 to 1991. Some of the patterns are more organised than others, with the Masvingo and Gwanda rainfall deviating very little during the periods between wet years where as large annual variations occur with Filabusi and Kezi rainfalls. In the period before 1930, large annual fluctuations occurred at all stations.

The standardised areal rainfall departures (figure 4) have a pattern which is very similar to that of individual station rainfall patterns. Positive anomalies occurred in the period before 1930,1950 to 1960 and 1973-1982. Negative anomalies occurred in the periods 1941-51, 1961-70 and 1982-1991. The years 1926/27, 1946/47, 1968/69, 1972/73 and 1982/83 show large negative anomalies and are years remembered by many people for their large rainfall deficits.

DISCUSSION

The large annual rainfall fluctuations the beginning of the century can be attributed to a number of factors.The instruments used initially were not of the same standard. Most observers at the time were people who had developed an interest in keeping weather records but they were not trained in carrying out the accurate observations. Observers change during the course of observations bringing in a human factor in the error

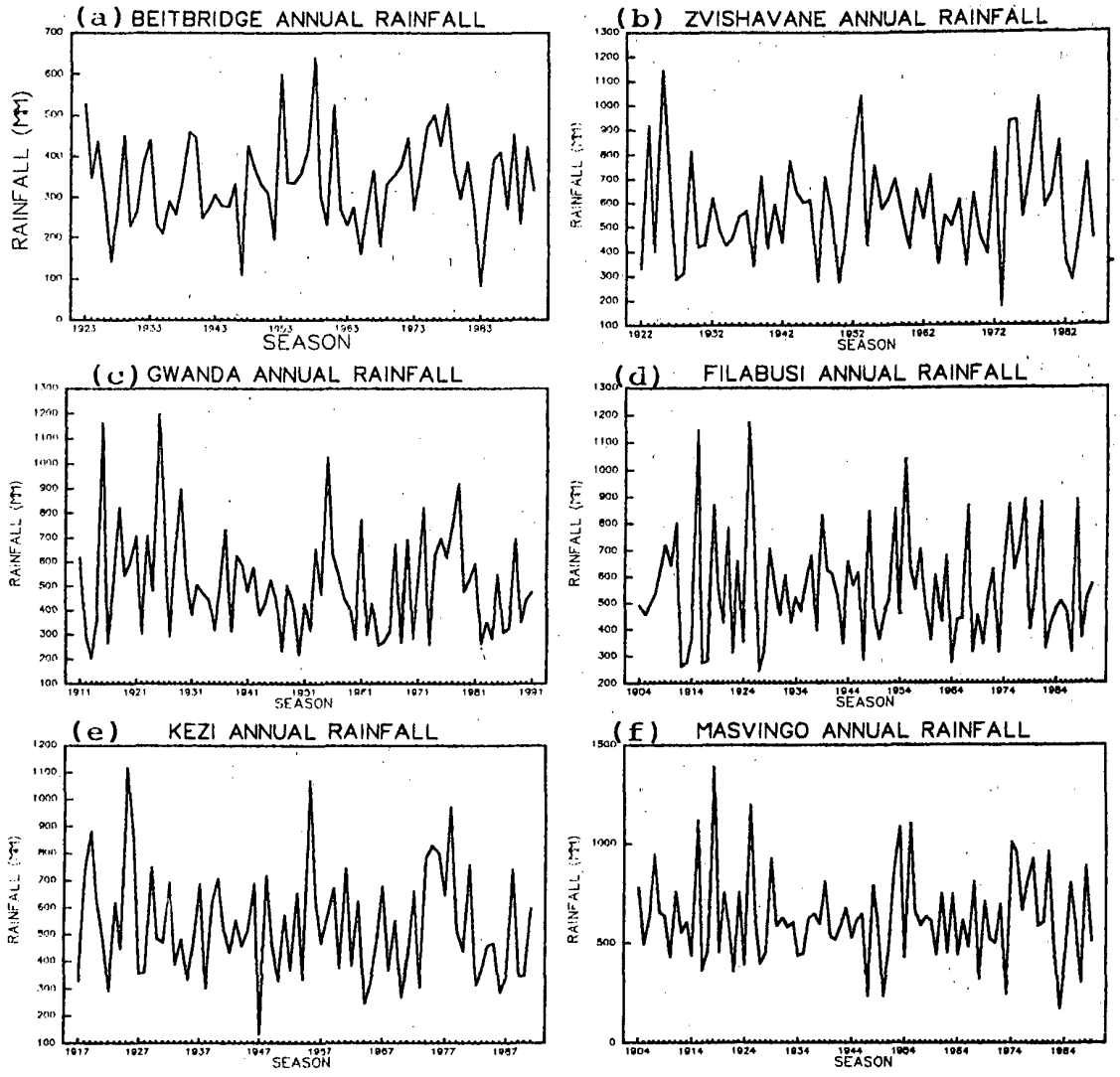


FIGURE 3: ANNUAL RAINFALL PATTERNS

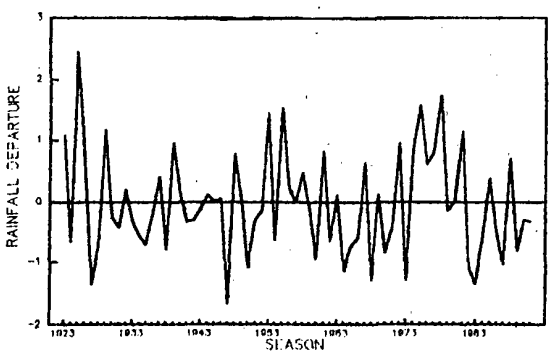


FIGURE 4: AREAL INTEGRATED STANDARDISED RAINFALL DEPARTURES

of observation. Some stations closed down because the persons who used to take man had moved to other locations and there was no-one at the old location to continue the work. In some cases, records have been replaced with those done a kilometre or more away in order to maintain continuity. Also some instrument locations were moved when expansion in buildings occurred. The convectional nature of the rainfall results in differences in rainfall amounts even over small distances depending on the location of development of the storm and its direction of motion (W.Marume 1988 unpublished MSc dissertation).

The results, for individual stations and when combined, do not indicate a downward trend as is expected in the event of desiccation. Instead an oscillatory pattern of about 10-12 years is strongly indicated a feature which conform with the analysis of Zimbabwe nation annual rainfall (Ngara et. al 1983). It is not quite clear what happened before the turn of the century since the data are not available. However, the cyclic nature of the rainfall also conforms with the findings of research work carried out in South Africa.(Reenen 1925, Tyson et.al. 1975, Dyer and Tyson 1977, Vines 1980). These results can, with some fine tuning, be extrapolated to cover this southern region of Zimbabwe.

However, when extrapolating these finding into the future one has to be cautious, because of the transient nature of the cycles. Climate has changed over the century and is still changing (Tyson 1991) coupled with the changes induced by the effects of human activities.

If the annual rainfall in the region has not changed, why then has the social-economic impact of rainfall deficits been so severe in the 1980s than was the case before, one may ask. Perhaps there are other factor, like reduced effective rainfall resulting from reduced number of rainy days (Booth 1989), increased runoff as a result of deforestation and poor land management, intensity and frequency of heavy storms increased demand on the water resources which has resulted in some boleholes drying up as pupolation increased and changes in human habits from growing more drought tolerant small grain crops to maize which requires more reliable rainfall and is lss tolerant to drought because it is more marketable, less labour intensive and the food is more palatable. Crop failure occurs more often with maize than small grains and people require food support more often now than during the time they grew small grains.

CONCLUSION

The is no evidence of desiccation in the Southern districts of Zimbabwe from annual rainfall. Instead cyclic oscillations of about 10-12 years are evident a result which conforms with studies carried out on annual rainfall in neighbouring South Africa. If desiccation has/is occurring perhaps it is cause by increased runoff due to deforestation, poor land management and tired soils after years of repeated cultivation and reduction in effective rainfall resulting from reduced number of rainydays.

ACKNOWLEDGEMENTS

I wish to thank the personnel in the Rainfall Section of the Zimbabwe Department of Meteorological Services for providing the rainfall data. Thanks too, to L.Unganai and A. MAkarau both of the Zimbabwe Department of Meteorological Services for their constructive comments on the manuscript.

REFERENCES

BOOTH,R.V.(1989)Indications of desiccation in the Matetsi/Deka catchment of north-western Zimbabwe. South Africa Journal of Science, Vol 85.

DYER,T.G.J. and TYSON,P.D. (1977) Estimating above and below normal rainfall periods over South Africa:1972-2000. Journal of Applied Meteorology, Vol 16.

LAMB,P (1983) Short Communications: Sub-Saharan rainfall update for 1982: Continued drought. Journal of Climatology, Vol 3, 419-422.

MARUME, W. (1988) Ths spatial ditribution of rainfall over Harare, Zimbabwe (unpublished Msc
dissertation)

NICHOLSON,S.E. (1983) Sub-Saharan rainfall in the years 1976-80: Evidence of continued drought. Monthly Weather Review, Vol 111.

TYSON,P.D. (1991) Climate Change in Southern Africa: Past Present conditions and Possible Future Scenarios. Climate Change, Vol 18; 241-258

TYSON,P.D and DYER T.G. (1978) The Predicted Above-Normal Rainfall of the Seventies and the Likelihood of droughts in the Eighties in South Africa. South African Journal of Science, Vol 74.

TYSON,P.D.,DYER,T.G. and MAMETSE,M.N.(1975) Secular changes in South African rainfall: 1880 to 1972. Quart.J.R.Met.Soc.101 ;817-833.

VINES,R.G. (1980) Analysis of South African Rainfall; South African Journal of Science, Vol 76.

ZIMBABWE DEPARTMENT OF METEOROLOGICAL SERVICES (1981) Climate Handbook of Zimbabwe.

NGARA,T.,McNAUGHTON,D.L. and LINEHAM,S (1983) Seasonal rainfall fluctuations in Zimbabwe. Zimbabwe Journal of Agriculture. Vol 30.

THE CLIMATE CHANGE FEATURE OF DYNESS/WETNESS FOR THE LAST 100-YEAR PERIOD IN CHINA

By You-Tang Liu, Area Observatory, Lin Fen 041000, Shanxi Province, China

ABSTRACT

Making up of the non-linear dynamics which will bring with new ideas on climate analysis.

It's presented that the dyness/wetness variation have had some quasi-periods in historical information. for example, they are quasi-two years(may be considerd as period-2),three,five,seven,eleven,twenty-two,thirty -six year's Periods(period-3,period-5 .....),etc.

In which,a few of them have been recognized by meteorologist,but,the periodic stage is instability , sometimes after which to persist for many years, it would be alternation to chaos,and vice versa,from the running spectrums of calculation of dyness/wetness grads year by year(1880-1990,including 50 citys).

The chaos to be resulted from periodic double-bifurcation,being fractal in spectrum dimension.

This paper has shown the period and chaos stages may alternate each other,so then, it's been making up that the chief feature of climate change.

In spit of these are disperse spectrums(a turbulence spectrum is continuous),but, which have been presented by obvious peak values,a wide band within spectrum is also chaos.

They are all provided that the periodic and chaotic stages with presistence,in general,to last 8-10 years, even more longest,which it's translation are catastrophe each other,may be considerated to be resulting from differental attractors.

The period or chaos had been having time-space variation,it'sdifferental that the field's distribution charts from period and chaos for per stages in national.

The dryness/wetness of periodic and chaotic stages are consistent with the general circulation of which of variations.

In experment,it's skill score is heigher that setting up statistical model during periodic stage(e.g. predictability) than that chaos stage.

RIVER REGIME IN RELATION TO CLIMATE CHANGE

A.A.L.N.Sarma & B.V.H.N.Sainath
Dept.of Meteorology and Oceanography, Andhra University
VISAKHAPATNAM-530 003, India

ABSTRACT

The regime of either a given landscape or a river basin is firstly dependent on its climate and secondly on the level of human activity over it. The level of human activity encompasses ~~m~~ such as injection of greenhouse gases through industrialisation, and deforestation, construction of different structures etc. Neither the natural nor the man made factors are constant. But of late the level of human activity might be responsible for the variability in the climate status of a given province about its mean. The world is by and large focussing its attention to view the recent changes from the man made angle. The authors have attempted to analyse the regime of the Krishna river basin, which is one of the largest rivers in south India in the context of climate change and for which the temperature, rainfall and river discharge date of it are made use of in drawing the conclusions.

The continentality index is used as the main parameter in finding whether there is any increase or decrease trend of the continental effect on the river regime with time at selected stations with special reference to rainfall. The characteristics of stream flow flood and drought events of the river basin is attempted through the theory of runs to get an insight into the variations in the intensity of extreme events with time. Attempts are also made to find the periodicity in the river flows. The harmonic approximation is giving a reasonable picture when the yearly fluctuations of the river flows are small about the normal. Further the available models are employed for the river flows to design them not only for the extremes but also for the different recurrence time intervals. The paper then proceeds in presenting the frequency analysis of stream flow volume for the determination of occurrence of low and high flow amounts using cube-root normal probability paper. Finally the paper deals with the mean decadal variations in the extreme flows to unfold the nature of aberrations in the respective flows in the light of hydrological regime change with reference to climate variability.

State Space Modeling of Paleoclimatic Time Series *

Robert H. Shumway
Kenneth L. Verosub
University of California, Davis

Abstract

We discuss preliminary results involving the use of time series methods as applied to the problem of reconstructing long term records of paleoclimate. We use, as an example, *varve chronologies* representing yearly deposition of clay and silt laid down in a lake by a receding glacier. The construction of accurate varve chronologies, which go back thousands of years, has the potential for creating accurate long-term high resolution proxies for paleoclimate.

We propose several *state space models* for handling problems associated with merging records from different sources that may also be susceptible to changes occurring at unknown time points. First, the underlying single varve thicknesses in New England, taken from the work of Ernst Antevs, are modeled using a simple first-order moving average *ARIMA* model on the log transformed differences. This suggests an equivalent simple first-order state space model for the vectors of partially observed varve chronologies which is then used to merge different series. A class of *non-Gaussian* state space models is proposed for handling regime changes occurring in paleoclimatic records.

1. Background

The first to recognize the potential of glacial varves for chronological studies was Gerard De Geer. Toward the end of the last century and into the beginning of this century, De Geer carried out a systematic study of natural and (hu)man-made exposures of varves in Sweden. These varves had formed in glacial lakes during the deglaciation of Scandinavia about 12,000 years ago, but the lakes are now drained. At each site De Geer measured the thickness of each varve in the sequence and compiled graphs of the thickness as a function of the position of a varve in the sequence. He then used the graphs from nearby sites to create a composite varve chronology that was longer than the record from any given site. By examining sites from one end of a glacial lake basin to the other, De Geer was able to compile composite varves chronologies incorporating several thousand varves (De Geer, 1921; 1940).

Shortly afterward, De Geer's student Ernst Antevs went to North America to study sequences of varves that had been deposited during the deglaciation of New England, New York and the St. Lawrence River valley (Antevs 1922, 1925, 1928). He too compiled varve chronologies for different basins, each containing several thousand varves. Because of the relationship between varve thickness and climate, the varve chronologies of De Geer and Antevs represent extraordinarily long records of paleoclimate at a resolution of one year. The situation is quite analogous to tree-ring series (see Fritts, 1991), and one would expect that with all of the interest in climate change, considerable attention would have been focussed on varve chronologies and on deciphering the record of paleoclimate that they contain.

Unfortunately, this has not been the case and, in fact, varve chronologies have been virtually neglected by paleoclimatologists. There are apparently two causes for this neglect. The first is that Antevs believed that the deglaciation of Sweden and New England occurred simultaneously with the same annual fluctuations in climate, and he tried to make a direct, trans-Atlantic correlation between the sequences from the two regions. His arguments, first put forth in the 1940's (see Antevs, 1962), were not very convincing, and they only served to raise doubts about varve studies in general. Later, when the first radiocarbon dates for the deglaciation of New England became available in the 1950's, there appeared to be too many varves for the time available, casting more doubt on Antevs methods and raising the question of whether the varves actually represented annual depositional cycles.

In the past ten years, the situation has changed significantly. Work by Verosub (1979 a, b) and by other workers in North America (Ridge and Larsen, 1990) as well as work by researchers in Sweden (Stromberg, 1985; Ringberg and Rudmark, 1985) has shown that the varve chronologies of Antevs and De Geer were quite good. In addition, new radiocarbon dates from sites in New England (Ridge and Larsen, 1990) have pushed back the date of the onset of deglaciation there, providing sufficient time for the accumulation of the observed number of varves. In Sweden, the varve chronology of De Geer has been extended up to the present with sites in the northern part of that country (Cato, 1985), and an absolute date has been calculated for the "zero" (first) varve of the Swedish chronology (Fromm, 1985). Finally, work on the varve chronologies in Sweden (Björk et al., 1988) and on sediments from existing proglacial lakes in Alaska (Glenn and Kelts, 1990) has confirmed that varves are indeed annual features.

In view of the developments above, it would seem that the time is right to "rehabilitate" varve chronologies as long, high-resolution proxy records of paleoclimate. The first step in this process would be to put the varve chronologies of DeGeer and Antevs on a firm statistical basis. In compiling their chronologies, Antevs and DeGeer simply "eyeballed" the records from different sites until they found what looked like a good match. They did not provide a measure of how good the chosen match was, compared to

* Presented at the 5th International Meeting on Statistical Climatology, 22-26 June, 1992, University of Toronto, Canada

other possible matches, or how the chosen match might be affected by the absence of one or two varves from the sequence. Therefore, as a first component of our collaboration we are addressing the question of the alignment of varve records from different sites.

Varve records from different sites can be regarded as yearly time series for which at least some sections contain common time intervals. For a common time interval, varves at different sites in a given proglacial lake will produce similar records that differ by additive noise that reflects local variations in the sedimentological regime. In addition, either because of geological processes or observational errors, some varves may be missing from any given record. Our goal in this project is to produce a single best record by combining records from different sites in their proper alignment. Because of the missing segments, the problems with which we are dealing are clearly in the realm of nonstandard time series analysis.

2. Modeling Varve Chronologies

By way of illustration, we show in Figure 1 below two of Antevs's varve chronologies, containing 1400 log transformed varve thicknesses, representing deposition for 1400 consecutive years at locations in Connecticut and Massachusetts respectively. Varves given are Antevs's numbers 3001 to 4400, with 3767-3871 (767-871 on the plot) defining the interval over which the thicknesses are observed at both sites. According to Ridge and Larson (1990), radiocarbon dating matches varve number 3001 with 15,600 years BP (before present).

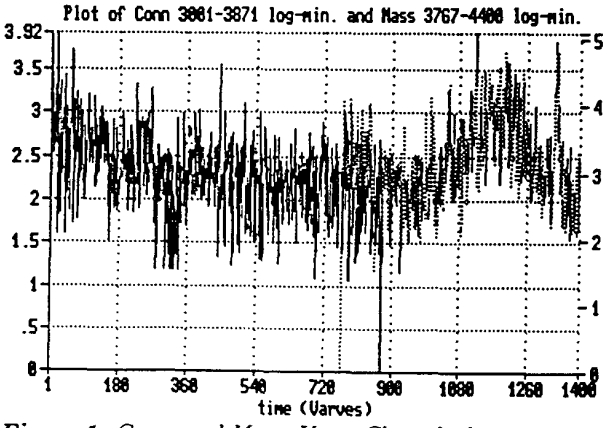


Figure 1: Conn. and Mass. Varve Chronologies.

The original untransformed records tend to have a number of years with very large thicknesses spaced through the series. In particular, the variation in the varve thickness tends to be proportional to the size of the varve, suggesting that logarithms might be applied to stabilize the variance. Histograms tend to confirm that the logarithms are more nearly normally distributed and we argue that for time series analysis purposes, the transformed series $y_t = \log V_t$, where V_t is the varve thickness in year t, tends to be more stationary and closer to normally distributed. Taking logarithms tends to discount the effect of the large varve thicknesses which often do not correlate well on common time intervals. The autocorrelation function of the transformed series in Figure 1 (not shown) decays very slowy, showing nonstationarity.

The nonstationarity in the Figure 1 varve series can be eliminated by differencing, yielding

$$y_t - y_{t-1} = \log \frac{V_t}{V_{t-1}} \tag{1}$$

as an underlying stationary series. The autocorrelation functions of the differenced series exhibit strong peaks at lag one for both the Conn. and Mass. series and both partial autocorrelation functions exhibit exponential decrease. These two observations imply that a first-order moving average, *ARIMA(0,1,1)* of the form

$$y_t - y_{t-1} = w_t - \theta w_{t-1} \tag{2}$$

can be expected to fit well where w_t is a zero-mean white noise process with variance σ_w^2 that drives the first-order linear system, characterized by the parameter θ. Estimating these parameters for Conn. and Mass. leads to the values $\theta = .94, .77$ and $\sigma_w^2 = .17, .23$ respectively. The residuals from these models were essentially white noise, showing that the first-order moving average provides an adequate description.

Two comments can be made at this point. The first is that the underlying model represents the percentage change (logarithms) as a first-order moving average, showing that the year to year changes are stationary and induced by the simple first-order driving system. The second comment is that the model given by Equation (2) is essentially equivalent to the simple state-space model proposed in the next section. The general form of the state space model, however, makes it more amenable to modeling the case where two separate varve chronologies need to be merged as will be considered in the next section.

3. Integrating Varve Records

We turn now to the problem of merging the two varve chronologies using the fact that they are observed at 105 points in common. Figure 2 below shows a snapshot of the common varve numbers for Conn. and Mass. and we see that they appear to be aligned properly. Antevs did this by eye; we checked by computing the cross correlation between the prewhitened series, i.e. using the estimated residuals $\hat{w}_t$ from the model defined in Equation (2). There was a strong peak at lag zero indicating that the series are correctly aligned.

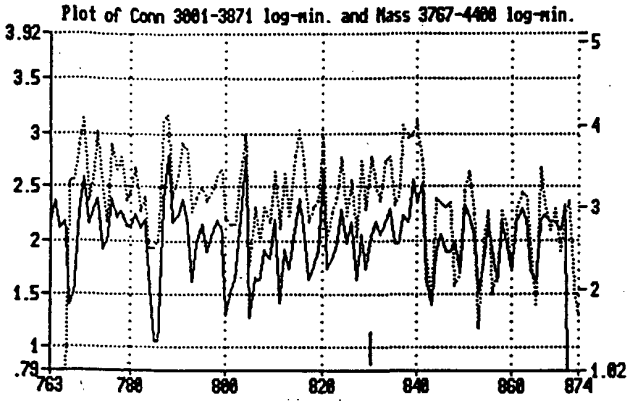


Figure 2: Snapshot of the Conn. and Mass. Varves

In the original data, shown in Figure 1, the scales on the right and left for the Conn. and Mass. varves respectively are different and we note from the common part shown in Figure 2 that there appears to be a constant shift between the thicknesses over much of the record. Regressing the Mass. varves on the Conn. varves over the common interval leads to an intercept of 1.2 and a slope of 1.0. This suggests a model of the form

$$y_{t1} = x_{t1} + e_{t1} \tag{3}$$

$$y_{t2} = x_{t1} + x_{t2} + e_{t2} \tag{4}$$

for the Conn. and Mass., varves y_{t1} and y_{t2}, respectively, where the common varve signal x_{t1} satisfies an equation of the form

$$x_{t1} = x_{t-1,1} + \epsilon_{t1}, \tag{5}$$

with ϵ_t having variance σ_ϵ^2. The additive process x_{t2} satisfies the equation

$$x_{t2} = x_{t-1,2} \tag{6}$$

so that the process is constant over time and measures the constant shift between the two separate varve records. Equations (3) and (4) are the so-called the *measurement equations* and (5) and (6) are the *state equations* in the general state-space model.

Each of the varve series separately, say y_{t1} and y_{t2}, are first order moving average processes of the form (2) as can be seen by differencing (3) and (4). The parameters $\hat{\sigma}_1^2 = .15$, $\hat{\sigma}_2^2 = .20$ associated with e_{t1}, e_{t2} as well as $\hat{\sigma}_\epsilon^2 = .005$ can be estimated by maximum liklihood. These small signal variance (.005) indicates that the underlying signal process is relatively smooth.

The general state space model is a matrix form (see Shumway, 1988) that includes the simple additive model above as a special case. The special features that simple additive models such as the above exhibit have been applied often in the literature to problems where series can be broken down into simple additive components (see, for example, Harrison and Stevens, 1976, Kitagawa and Gersch, 1984 or Shumway and Katzoff, 1991). Models of this kind are often confronted with data occurring in segments where one or the other pieces of some observed series are missing. Treatments of the estimation and smoothing problems for irregularly spaced data can be found in Parzen (1984) or in Shumway and Stoffer (1982) using the EM algorithm.

For the merged series, we may choose the *Kalman smoothed estimator* for x_{t1}, namely

$$x_{t1}^n = E\{x_{t1}|y_{t1}, y_{t2}, t = 1, \ldots, n\}. \tag{7}$$

Such a merged estimator is shown in Figure 3.

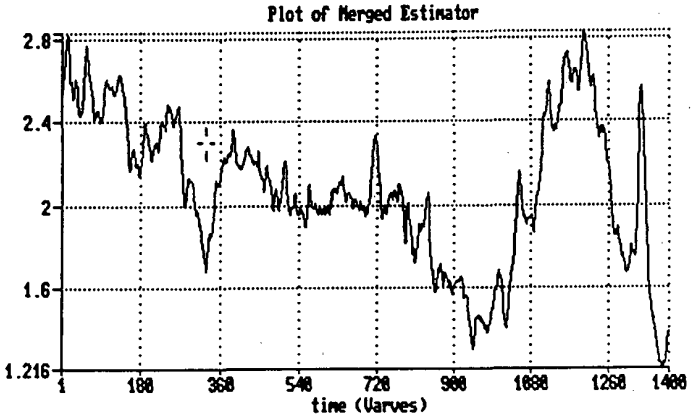


Figure 3: Kalman Smoothed Varve Signal

Note that the smoothed estimator may or may not be the best proxy for the merged varve series. Figure 3 shows that there are common higher oscillations present that may be worth adding into the model. For example, one might consider adding a common second component to each of the two series and then merging the result as the smoothed value of the sum of the components. The estimated shift process, say x_{t2}^n, turned out to be 1.20 with an estimated standard error of .05, confirming the value obtained by simple linear regression.

Modeling Regime Changes

The structural model given in the previous section is also convenient for describing possible changes in regime present in a time series. As a simple contrived example, consider the series in Figure 4 which contains a process with a simple added shift in the mean.

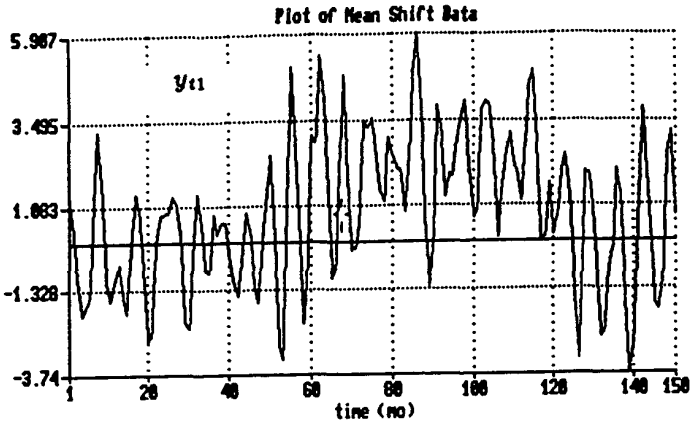


Figure 4: Process with *ARIMA(2,0,0)* and Mean Shift.

The data were generated using the model

$$y_t = x_{t1} + e_t \tag{8}$$

when the mean shift is not present, where the signal process is modeled as the second order autoregression, say of the form *ARIMA(2,0,0)*, written as

$$x_{t1} = \phi_1 x_{t-1,1} + \phi_2 x_{t-2,1} + \epsilon_{t1}. \tag{9}$$

Such a model can represent disturbed periodicities in paleoclimatic series as in Newton and North (1991) who express periodicities in the global ice volume series in terms of these models.

Suppose now that a change occurs in this process at some unknown time point that can be modeled as a simple mean shift added to the underlying model given above. That is, the observed series now becomes

$$y_t = x_{t1} + x_{t2} + e_t, \tag{10}$$

where

$$x_{t2} = x_{t-1,2} \tag{11}$$

The above process can be expressed as a special case of the overall mixture model

$$y_t = \sum_{j=1}^{m} a_{tj} x_{tj} + v_t \tag{12}$$

where the coefficients a_{tj} determine the presence of absence of the jth component x_{tj}. These coefficients are random and putting them into the matrix $A_t = (a_{t1}, \ldots, a_{tm})$, we may specify a probability structure for A_t, taken here to be a stationary Markov chain. For an approach where the A_t are independent processes, see Shumway and Stoffer (1991).

Letting the stationary and transition probabilities be

$$\pi_j = P\{A_t = M_j\}$$

and

$$\pi_{ij} = P\{A_{t+1} = M_j | A_t = M_i\},$$

we find that we are interested in the smoothed probability of being in the regime j at time t, defined as

$$\pi_j(t|n) = P\{A_t = M_j | y_1, \ldots, y_n\}. \tag{13}$$

Such smoothed probabilities can be approximated using various approaches (see Lindgren, 1978, Kitagawa, 1987, Shumway and Stoffer, 1991, Shumway, 1992). The equations used here are from the last reference. Of course, the problems associated with estimating the parameters $\phi_1, \phi_2, \sigma_\epsilon^2$ and σ_ϵ^2 along with the initial mean of x_{t2} must be solved as well. A maximum likelihood estimation procedure is given in Shumway (1992).

The smoothed probabilities for the data in Figure 4 are shown in Figure 5 below.

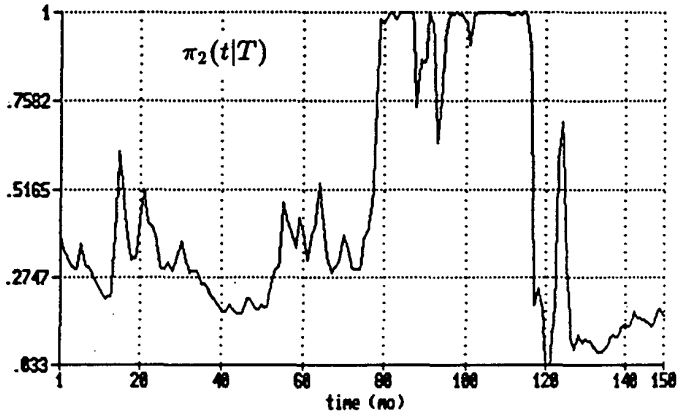


Figure 5: Smoothed Regime Probabilities

It is clear that the mean shift is detected with high probability between the time points $t = 80$ and 120. In this case, one would primarily be interested in the best estimator for the underlying signal, given by $x_{t1}^n = E\{x_{t1}|y_1, \ldots, y_n\}$ and shown in Figure 6.

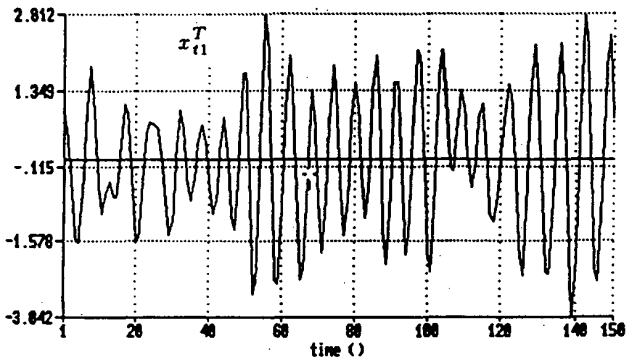


Figure 6: Smoothed Signal for Figure 4 Data.

It is clear again that the mean shift has disappeared from the estimated series, which remains an excellent replica of the original signal.

Discussion

In future research, we want to develop parameters that can be used to study periodicities and trends on annual, decadal and century time scales and to define other parameters that measure variations in the mean thickness of varves as well as the amplitude of the fluctuations about those means. We will investigate various models for stochastic trends and periodicities that can all be put into the state-space form given above. Although it is not clear at the present time how the varve chronologies will be modeled, we conjecture that one of the structural models involving some combination of smooth, seasonal and irregular components will be adequate. The various models and parameters which we develop will serve to demonstrate that more can be done with varve chronologies from different regions than simple varve-by-varve comparison of thicknesses.

We will also attempt to show that the varves may contain more information than has previously been assumed. Most workers have only measured the total thickness of each silt-clay couplet. We believe that the thicknesses of the silt and clay layers should be treated as separate parameters and that they may be influenced by different combinations of climatic variables. Therefore, we will focus on determining the extent to which the thicknesses of the silt and clay layers vary independently. Each pair of silt and clay thicknesses can be taken as an element in a bivariate time series that can be studied using classical time series methodology.

As an additional component of the research, we will attempt to determine which climate variables contribute to the thickness of a varve layer. Although it is generally assumed that spring temperatures and spring rainfall are the primary factors, this conclusion has never been confirmed. Because we now have thickness data from varves in Sweden (Fromm, 1985) and in Alaska (Perkins and Simms, 1983) that were deposited within the last 100 years, it is possible to compare these variations in varve thickness with recorded variations in climate to deduce the correct relation between them.

Our expectation is that the various analyses that we will undertake will demonstrate that varve chronologies have considerable potential as a paleoclimate database and that their acceptance and use will increase. One of the most important of these uses will be in the testing of global circulation models. Although such models are designed to predict the climate in the future, they can only be tested and validated by comparing their predictions for past climatic conditions to actual paleoclimate data. The varve chronologies provide an opportunity to test the global circulation models at several levels. For example, records from individual glacial lakes can be used to examine a model's ability to predict long-term and short-term local variability during the deglaciation while records from lakes in the same or adjacent region can be used to verify predictions of intermediate scale coherence and variability. Finally, the records from opposite sides of the Atlantic Ocean will provide a stringent test of whether a given model can correctly predict the climate at two places simultaneously.

A long-range statistical objective would be the development of statistical forecasting models for paleoclimate based purely on the varve chronologies.

Acknowledgements

This research was supported in part by Grant DMS-9115748 from the from the Division of Mathematical Sciences, National Science Foundation.

References

Antevs, E. (1922). The recession of the last ice sheet in New England. *Amer. Geog. Soc. Res. Series*, No. 11, 120p.

Antevs, E. (1925). Retreat of the last ice sheet in eastern Canada. *Canadian Geological Survey Memoir*, **146**, 1-142.

Antevs, E. (1928). The last glaciation, with special reference to the ice sheet in northeastern North America. *Amer. Geog. Soc. Res. Series*, No. 17, 292 p.

Björck, S., B.E. Berglund and G. Digerfeldt (1988). New aspects of the deglaciation chronology of south Sweden. *Geographia Polonica*, **55**, 37-49.

Cato, I. (1985). The definitive connection of the Swedish goechronological time scale with the present, and the new date of the zero year in Doviken, northern Sweden. *Boreas*, **15**, 123-125.

DeGeer, G. (1912). A geochronology of the last 12,000 years. *Report of the 11th Int. Geol. Congress*, **1**, 241-253.

DeGeer, G. (1940). Geochronologia Suecica Principles, *Svenska Betenskapsakad. Handl.*, **3:18**, 367 p.

Fritts, H.C. (1991). *Reconstructing Large-Scale Climatic Patterns From Tree-Ring Data: A Diagnostic Analysis.* University of Arizona Press.

Fromm, E. (1985). Chronological calculation of the varve zero in Sweden. *Boreas*, **14**, 123-125.

Glenn, C.R and K. Kelts (1989). Sedimentary rhythms in lake sediments. *Cycles and Events in Stratigraphy*, 1-42.

Harrison, P.J. and C.F. Stevens (1976). Bayesian forecasting (with discussion). *J. Roy. Statist. Soc.*, **B 38**, 205-247.

Kitagawa, G. (1987). Non-Gaussian state-space modeling of nonstationary time series. *J. Amer. Statist. Assoc.*, **82**, 1032-1063.

Kitagawa, G. and W. Gersch (1984). A smoothness priors state space modeling of time series with trend and seasonality. *J. Amer. Statist. Assoc*, 79, 378-389.

Lindgren, G. (1978). Markov regime models for mixed distributions and switching regressions. *Scand. J. Statist.*, 5, 81-91.

Newton, H.J. and G.R. North (1991). Forecasting global ice volume. *J. Time Series Analysis*, 12, 256-265.

Parzen, E. ed. (1984). *Time Series Analysis of Irregularly Observed Data*, Lecture Notes in Statistics 25. New York: Springer Verlag.

Perkins, J.A. and J.D. Sims (1983). Correlation of Alaskan varve thickness with climate parameters and use in paleoclimate reconstruction. *Quaternary Research*. **20**, 308-321.

Ridge, J.C. and F.D. Larson (1990). Re-evalution of Antev's new England varve chronology and new radiocarbon dates of sediments from glacial Lake Hitchcock. *Bull. Geolog. Soc. of Amer.*, **102**, 889-899.

Ringberg, B. and L. Rudmark (1985). Varve chronology based upon glacial sediments in the area between Karlskrona and Kalmar, southeastern Sweden. *Boreas*, **14**, 107-110.

Shumway, R.H. (1988). *Applied Statistical Time Series Analysis*, Englewood Cliffs: Prentice Hall.

Shumway, R.H. and D.S. Stoffer (1982). An approach to time series smoothing and forecasting using the EM algorithm. *J. Time Series Analysis*, **3**, 253-264.

Shumway, R.H. and D.S. Stoffer (1991). Dynamic linear models with switching. *J. Amer. Statist. Assoc.*, **86**, 763-769.

Shumway, R.H. and M.J. Katzoff (1991). Adjustment of provisional mortality series: The dynamic linear model with structured measurement error. *J. Amer. Statist. Assoc.*, **86**, 611-617.

Shumway, R.H. (1992). Nonlinear and non Gaussian modeling of mortality and morbidity series. Final Report. May 1, 1992, Office of Research and Methodology, National Center for Health Statistics, 6525 Belcrest Rd., Hyattsville, MD 20782.

Stromberg, B. (1985). revision of the late glacial Swedish varve chronology. *Boreas*, **14**, 101-105.

Verosub, K.L. (1979a). Paleomagnetism of varved sediments from western New England: Secular variation. *Geophys. Res. Letters*, **6**, 245-248.

Verosub, K.L. (1979b). Paleomagnetism of varved sediments from western New England: Variability of the paleomagnetic recorder. *Geophys. Res. Letters*, **6**, 241-244.

State Space Modeling of Time Series and Some Applications to Climatological Data

Genshiro Kitagawa
The Institute of Statistical Mathematics
4-6-7 Minami-Azabu, Minato-ku Tokyo, JAPAN 106

Abstract

A state space approach to the analysis of nonstationary time series is shown. Various types of nonstationary time series models can be expressed by using the state space model which facilitates the application of computationally efficient Kalman filter and smoothing algorithms. The method can be extend to a more general situation where the system is nonlinear or the noise distributions are non-Gaussian. The application of this general state space model is very broad. As examples of the use of the state space model, the analyses of the ground water level data, the maximum temperature data and the rainfall data are shown.

1. Introduction

A state space approach to the analysis of time series is shown. Two types of state space models and related filtering and smoothing algorithms are shown.

The first one is the ordinary state space model and is useful for the modeling of nonstationary time series with gradually changing parameters. The gradual change of the parameter θ can be expressed by $\Delta^k \theta_n = v_n$ where Δ is the difference operator defined by $\Delta\theta_n = \theta_n - \theta_{n-1}$ and v_n is a Gaussian white noise sequence with mean 0 and unknown variance τ^2. If $\tau^2 = 0$, the solution to the above model becomes a polynomial of order $k-1$. However, for nonzero τ^2, θ_n can express very flexible and smooth function. The model can be easily incorporated into the state space model, and the computationally efficient Kalman filter can be used for model fitting and state estimation.

Although the state space model is very powerful and has very wide range of applications, there are some important models that cannot be handled with this model. Non-Gaussian models and nonlinear models are such examples. The second model is developed for the analysis of such models.

In this paper, by using the data sets related to climatology, we shall exemplify the use of state space model. In the first example, we shall show the treatment of missing and outlying observations and dynamic regression. In the second example, the models for nonstationarity in the mean, varaince and covariance are shown. Discrete process and Quasi periodic process are also briefly discussed.

2. State Space Model and State Estimation

2.1 Linear Gaussian Case

Consider the state space model

$$\begin{aligned} x_n &= F_n x_{n-1} + G_n v_n \\ y_n &= H_n x_n + w_n, \end{aligned} \qquad (1)$$

where y_n is the observations, x_n is the unknown state vector, v_n and w_n are white noise sequences with

$$v_n \sim N(0, Q_n), \qquad w_n \sim N(0, R_n).$$

Many types of nonstationary time series models as well as the standard stationary times series models can be expressed in this form. The model is in particular useful for the system whose parameters are gradually changing with time.

Many important problems in time series analysis can be formulated as the estimation of the state vector of the state space model. And the problem of state estimation is reduced to the evaluation of $p(x_n|Y_m)$, the conditional density of x_n given observations $Y_m \equiv \{y_1, ..., y_m\}$. For $n > m, n = m$ and $n < m$, they are called the prediction, filtering and smoothing problems, respectively. For linear Gaussian state space model, $p(x_n|Y_m)$ is a Gaussian distribution $N(x_{n|m}, V_{n|m})$. Therefore, it is sufficient to specify the mean $x_{n|m}$ and the covariance $V_{n|m}$.

The well-known Kalman filter and the smoother are the recursive solution to these problems.

One step ahead prediction:

$$\begin{aligned} x_{n|n-1} &= F_n x_{n-1|n-1} \\ V_{n|n-1} &= F_n V_{n-1|n-1} F_n^t + G_n Q_n G_n^t \end{aligned} \quad (2)$$

Filtering:

$$\begin{aligned} K_n &= V_{n|n-1} H_n^t (H_n V_{n|n-1} H_n^t + R_n)^{-1} \\ x_{n|n} &= x_{n|n-1} + K_n (y_n - H_n x_{n|n-1}) \\ V_{n|n} &= (I - K_n H_n) V_{n|n-1} \end{aligned} \quad (3)$$

Smoothing:

$$\begin{aligned} A_n &= V_{n+1|n} F_{n+1}^t V_{n|n}^{-1} \\ x_{n|N} &= x_{n|n} + A_n (x_{n+1|N} - x_{n+1|n}) \\ V_{n|N} &= V_{n|n} + A_n (V_{n+1|N} - V_{n+1|n}) A_n^t. \end{aligned} \quad (4)$$

2.1 General State Space Model

Assume that the time series is expressed by a general state space model

$$\begin{aligned} x_n &\sim q(x_n|x_{n-1}) \\ y_n &\sim r(x_n|x_n), \end{aligned} \quad (5)$$

where q and r are the conditional distributions of x_n given x_{n-1} and of y_n given x_n, respectively. The initial state vector x_0 is distributed according to the density $p(x_0|Y_0)$. This general state space model contains various important time series models that cannot be expressed in ordinary state space model. Non-Gaussian noise models, nonlinear models and the discrete distribution models are typical examples.

It can been shown that for the general state space model , the recursive formulas for obtaining one step ahead prediction, filtering and smoothing distributions are given as follows (Kitagawa [5],[6]):

One step ahead prediction:

$$p(x_n|Y_{n-1}) = \int_{-\infty}^{\infty} q(x_n|x_{n-1}) p(x_{n-1}|Y_{n-1}) dx_{n-1} \quad (6)$$

Filtering:

$$p(x_n|Y_n) = \frac{r(y_n|x_n) p(x_n|Y_{n-1})}{p(y_n|Y_{n-1})} \quad (7)$$

where the denominator $p(y_n|Y_{n-1})$ is obtained by $\int p(y_n|x_n) p(x_n|Y_{n-1}) dx_n$.

Smoothing:

$$p(x_n|Y_N) = p(x_n|Y_n) \int_{-\infty}^{\infty} \frac{p(x_{n+1}|Y_N) q(x_{n+1}|x_n)}{p(x_{n+1}|Y_n)} dx_{n+1}. \quad (8)$$

These formulas (6), (7) and (8) show recursive relation between state distributions. Variety of non-Gaussian models, nonstationary models, nonlinear models and discrete distribution models can be handled within this frameworks. In the general case, however, the conditional distribution of the state $p(x_n|Y_m)$ becomes non-Gaussian and cannot be specified by using only the first two moments. It thus becomes necessary to use a numerical method for the realization of the formulas (Kitagawa [5], Sorenson and Alspach [9]).

2.3 Model Fitting

The state space model usually contains unknown parameters in Q_n, R_n, F_n, G_n or H_n. The best values of the parameters can be obtained by maximizing the log likelihood defined by

$$\begin{aligned} l(\theta) &= \log p(y_1, ..., y_N) \\ &= \sum_{n=1}^{N} \log p(y_n|Y_{n-1}). \end{aligned} \quad (9)$$

Here each $p(y_n|Y_{n-1})$ is the denominator in (7).

If we have several candidate models, the goodness of the model can be evaluated by the value of AIC defined by

$$\text{AIC} = -2\,\ell(\hat{\theta}) + 2(\text{number of parameters}). \quad (10)$$

Thus the best model can be found by looking for the one with the smallest value of AIC.

3. Ground Water Level Data

The data we analyze in this section is observed at Tokai area, Japan. The ground water level (y_n), barometric pressure (p_n), earth tide (et_n) and the precipitation (r_n) are observed at every 10 minutes since 1981 (Matsumoto et. al.[7]). The number of observations is about 500,000. Anomalous changes of the water level is an important earthquake precursors. However, actually it is difficult to detect the change of the water level related to the earthquakes, since it is seriously contaminated with various effects. Further, these data set contains huge amount of missing and outlying observations. In this section, by using this data set, we exemplify the use of the state space model for the treatment of missing and outlying observations and for the dynamic regression for eliminating the effects of other variables.

3.1 Missing Observations

In principle, the presence of missing observations does not cause any trouble in state space modeling.

If the observation y_n is missing, then $p(x_n|Y_n) = p(x_n|Y_{n-1})$ holds in equation (7). In the Kalman filtering this means that we just need to skip the filtering step and put

$$x_{n|n} = x_{n|n-1} \qquad V_{n|n} = V_{n|n-1}.$$

Therefore we can compute the log-likelihood even when there are many missing observations. Further, if necessary, we can interpolate the missed one by $y_{n|N} = H_n x_{n|N}$, where $x_{n|N}$ is obtained by the smoothing algorithm.

3.2 Outlying Observations

Some part of the observations of ground water contains many outlying observations due to a measurement problem. To treat these outliers, we considered the mixture model for measurement noise

$$w_n \sim \alpha N(0, \sigma^2) + (1-\alpha)N(\mu, \xi^2) \qquad (11)$$

Here σ^2 is the variance of the normal observation, μ and ξ^2 are the mean and the variance of the outlying observations, respectively, and α is the mixture weight. Since (11) is non-Gaussian, we apply the non-Gaussian filter and the smoother. Figure 3.2 shows a part of the ground water level after smoothing for many missing and outlying observations.

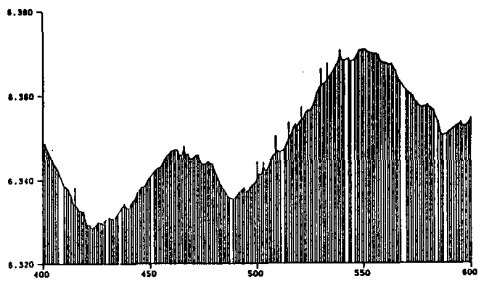


Fig. 3.1 Ground Water Level

3.3 Dynamic Regression

The data shown in Figure 3.2 has large deviation and it is difficult to find out the coseismic or preseismic effects. We then tried to detect the effect of earthquakes. To do that we consider the model

$$y_n = t_n + P_n + ET_n + R_n + w_n, \qquad (12)$$

where t_n, P_n, ET_n and R_n are the trend, air pressure effect, earth tide effect and precipitation effect, respectively. For the estimation of these effects, we consider the following models for these components

$$\begin{aligned} \Delta^k t_n &= v_n \\ P_n &= \sum_{i=0}^{m} a_i p_{n-i} \\ ET_n &= \sum_{i=0}^{\ell} b_i et_{n-i} \\ R_n &= \sum_{i=1}^{p} c_i R_{n-i} + \sum_{i=1}^{q} d_i r_{n-i} + v_n. \end{aligned} \qquad (13)$$

The basic model (12) and the component models (13) can be combined into a state space model. Therefore, the Kalman filter/smoother can be used for the estimation of these components. The unknown coefficients, a_i, b_i, c_i and d_i are estimated by the maximum likelihood method. The orders of the model, m, ℓ, p and q are determined by AIC criterion. Further, there are many possible extensions to the rain effect model. We can also find the best model by AIC.

Figure 3.3 shows the estimated trend and the effects of the air pressure, the earth tide and the precipitation. The trend component clearly shows upward tendency and occasional sudden decrease. These are considered as coseismic effects. By the analysis of the entire data set, it can be seen that the most of the nearby earthquakes have some coseismic effects.

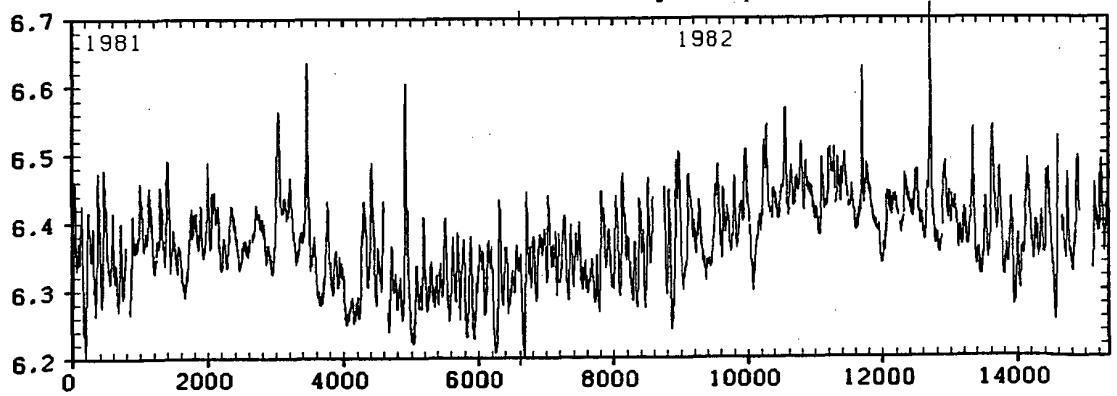


Fig. 3.2 Ground Water Level, 1981-1982

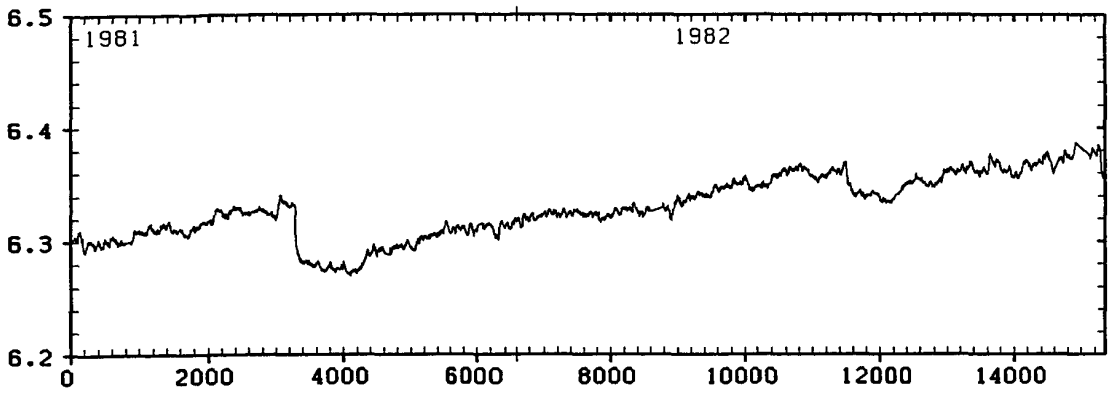


Fig. 3.3 Estimated Trend

4. Maximum Temperature Data

The second data set is the 21 years records of daily maximum temperature at Tokyo, 1971–1991. The state space methods for the estimation of the annual cycle, the time dependent variance and the time-varying AR model are shown with this data.

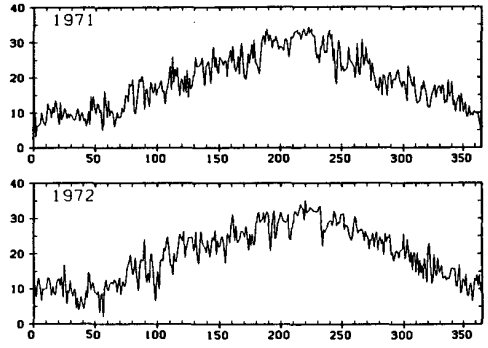


Fig. 4.1 Maximum Temperature, 1971–1972

4.1 Estimation of the Common Trend

Figure 4.1 shows the records of the first two years. Although the presence of the annual cycle is clear, it depends on the years. We consider the extraction of such an annual cycle (common trend). To estimate this annual cycle, we consider the model

$$\begin{aligned} \Delta^k t_n &= v_n \quad (k = 1 \text{ or } 2) \\ p_n &= \sum_{j=1}^{m} a_j p_{n-j} + u_n \\ y_n &= t_n + p_n + w_n. \end{aligned} \tag{14}$$

In this decomposition, t_n is the annual cycle and thus has the constraint that $t_n = t_{n-365}$. On the other hand, p_n is the stationary component expressed by an AR model and corresponds to the deviation from the trend. The model (14) with the constraint can be properly handled with the state space model. The estimated trend and the "rest" for the first two years are shown in Figure 4.2

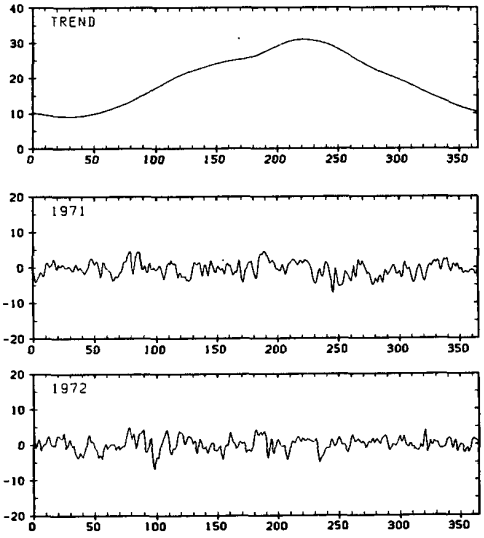


Fig 4.2 Estimated trend and the deviation from the trend, 1971-1972.

4.2 Changing Variance

The deviation from the trend shown in Figure 4.2 is stationary in the mean. However, its second moments are not homogeneous. The variance looks small in winter and in early August, since the weather in these season are stable in Tokyo. For the estimation of the time-varying variance, we consider the following model

$$\Delta^k(\log r_n) = v_n$$

$$y_n \sim N(0, r_n) \quad (15)$$

with the constraint $r_n = r_{n-365}$. This model is nonlinear and cannot be handled by the ordinary state space model. However, by using the general state space model given in (5), we can estimate the changing variance r_n.

Another way of estimating the changing variance is via the smoothing of transformed series. First, define u_n by

$$u_n = \log y_n^2, \quad (16)$$

then $u_n - \log r_n$ is an independent random variable distributed as the logarithm of a χ_1^2 distribution. Therefore, we can estimate the log-variance, $s_n = \log r_n$, by using the model

$$\begin{aligned} \Delta^k t_n &= v_n \\ u_n &= t_n + w_n, \end{aligned} \quad (17)$$

where the observational noise w_n is distributed as non-Gaussian distribution with the density $r(w) = 0.5exp\{(w - e^w)/2\}$. The non-Gaussian filter and the smoother is also applicable to this non-Gaussian model.

4.3 Time-varying Coefficients AR Model

Another feature of the deviation from the trend shown in Figure 4.2 is the time dependency of the wave form. This can be characterized by the time dependent covariance structure and we model such a series by using an AR model with time-varying coefficients

$$y_n = \sum_{j=1}^{m} a_{jn} y_{n-j} + w_n. \quad (18)$$

where w_n is a white noise with the variance σ^2.

For the estimation of the time-varying AR coefficients, we put the following constraint model

$$\Delta^k a_{jn} = v_{jn} \quad (19)$$

with $v_{jn} \sim N(0, \tau^2)$, and the constraint $a_{jn} = a_{j,n-365}$. The models (18) and (19) can be expressed in the state space model form. For example, for $k = 1$, it is given by

$$\begin{aligned} F_n = G_n = I, \quad H_n &= (y_{n-1}, ..., y_{n-m}) \\ x_n &= (a_{1n}, ..., a_{mn})^t. \end{aligned} \quad (20)$$

Once the time-varying AR coefficients are estimated by the smoothing algorithm, we can get the instantaneous spectrum of the nonstationary process by

$$p_n(f) = \frac{\sigma^2}{|1 + \sum_{j=1}^{m} a_{jn} \exp(2\pi i j f)|^2}. \quad (21)$$

Fig.4.3 shows the estimated changing spectrum of the series. The change of the spectrum is clearly seen in this figure. The estimated time-varying coefficient AR model is also useful for the prediction of the nonstationary time series.

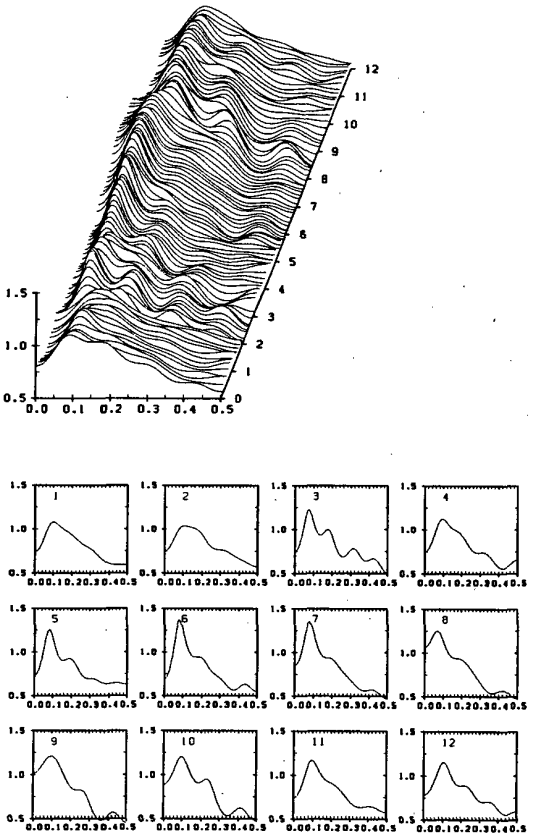


Fig 4.3 Estimated changing spectrum

5. Some Other Examples

5.1 Inhomogeneous Discrete Process

The general state space model can be applied to the estimation of time-varying mean of the discrete process. In Kitagawa [5] the estimation of the time-varying probability, p_n, of the occurrence of the rainfall in a specific calendar day of the binomial process was considered. The probability of rainfall was esti-

mated by the following model:

$$\begin{aligned} \Delta^k q_n &= v_n \\ m_n &\sim B(p_n, \ell_n). \end{aligned} \tag{22}$$

Here $q_n = \log\{p_n/(1-p_n)\}$, ℓ_n is the number of observations at the n-th time point, m_n the number of rainy days. The same method can be used for various discrete distributions such as the inhomogeneous Poisson distribution.

5.2 Quasi Periodic Process

Many of the climatological data exhibits the repetition of the same pattern but the period and the amplitude are not so definite. This type of quasi periodicity can be seen in many ecological data and air pollution data. For a time series with quasi periodic character, by using a model

$$\begin{aligned} \Delta^k t_n &= v_n \\ \Delta^k(\log c_n) &= u_n \\ y_n &= c_n h(\omega n + t_n) + w_n. \end{aligned} \tag{23}$$

we can estimate phase and amplitude of the model. $h(t)$ is a cyclic function and is expressed, for example, by a Fourier series.

REFERENCES

1. Akaike, H. , "Likelihood and Bayes Procedure," in *Bayesian Statistics,* J.M.Bernardo, M.H. De Groot, D.V. Lindley and A.F.M. Smith, eds., University Press, Valencia, Spain, 1980, 143-166.

2. Anderson, B.D.O. and Moore, J.B. , "Optimal Filtering", New Jersey, Prentice-Hall, 1979.

3. Kitagawa, G. , "Changing Spectrum Estimation", *Journal of Sound and Vibration*, 89, No.4, 1983, 433-445.

4. Kitagawa, G. and Gersch, W., "A Smoothness Priors-State Space Approach to the Modeling of Time Series with Trend and Seasonality", *Journal of the American Statistical Association*, 79, No.386, 1984, 378-389

5. Kitagawa, G. , "Non-Gaussian State Space Modeling of Nonstationary Time Series", *Journal of American Statistical Association,* Vol.76, No.400, 1987, 1032-1063.

6. Kitagawa, G. , "Numerical Approach to Non-Gaussian Smoothing and its Applications," *Computing Science and Statistics, Proceedings of the 20th Symposium on the Interface,* eds., E.J. Wegman, D.T. Gantz, J.J. Miller, Fairfax, Virginia, 1988, 379-388.

7. Matsumoto, N., Takahasi, M. and Kitagawa, G. , "Development of Linear multivariate regression model to detect coseismic changes of ground-water level, Bull. Geol. Surv. Japan, Vol.40, 1989, 613–623, (in Japanese with English abstract).

8. Sage, A. P. and Mersa, J. L. , "Estimation Theory with Applications to Communications and Control", McGraw-Hill Series in System Science, McGraw-Hill, New York, 1971.

9. Sorenson, H.W. and Alspach, D.L. , "Recursive Bayesian Estimation Using Gaussian Sums", *Automatica*, 7, 1971 465-479.

A Hydrological Rainfall Model Using Atmospheric Data With Application to Climate Model Impact Evaluation

Peter Guttorp, James P. Hughes, and Paul D. Sampson
Department of Statistics, University of Washington

ABSTRACT

In order to evaluate local effects on climate from various scenarios, such as CO_2-doubling, it is necessary to link the global-scale circulation models to regional effects. Since the circulation models are not developed with regional accuracy as a goal, regional models that use the global model outputs need to be developed. In the context of precipitation effects we develop a model that uses smooth atmospheric fields, such as temperature profile and surface pressure, to determine relatively homogeneous weather classes, within which simple spatial stochastic precipitation models are appropriate.

1. Background

In order to assess the impact of various scenarios for changes in anthropogenic emissions on climatic factors, such as temperature or precipitation, we commonly use the output from general circulation models, and attempt to interpret such outputs for local regions. However, these models have not been developed with regional accuracy as a goal, and precipitation, in particular, is usually poorly captured. This may not be important from the point of view of assessing global effects, but is clearly a problem when regional assessment of climate effects is needed. Generally it appears that spatially smooth fields, such as temperature and pressure, are better captured by the circulation models than are non-smooth fields such as precipitation.

Over the last decade there has been substantial work in the hydrological community on stochastic models for rainfall. Such models are important in order to be able to simulate realistic precipitation scenarios as input to runoff models, dam capacity calculations etc. In this paper we review some of the work done in this area, indicate some drawbacks with the current approaches, and outline a fairly general stochastic model that incorporates meteorological information. One possible use for this type of model is to use output from a circulation model, computed under various scenarios, to assess the regional impact on precipitation. The advantage with this approach would be that one uses only the "better" parts of the GCM output (i.e., the pressure and temperature fields), and translates these into precipitation using a regionally determined model.

In this paper we first describe the oldest class of stochastic precipitation models, namely Markov chain models. In Section 3 we present an alternative approach using point process theory. Section 4 summarizes some recent work on incorporating atmospheric data, directly or indirectly, while section 5 outlines a new structural model of precipitation. Finally, in section 6 we discuss some details of this approach.

2. Markov Chain Models

The simplest model for daily precipitation at an observed site is the Markov chain model, which assumes that in order to predict the occurrence of precipitation tomorrow from historical data we only need to know whether or not it rained today. Such a model was first suggested by Quetelet (1852) and in modern literature by Gabriel and Neuman (1962). A relatively recent review is the paper by Stern and Coe (1984).

In the simplest Markov chain model, let X_t be 1 or 0 according to whether or not precipitation occurred on day t (we use "day" in a generic sense meaning a fixed time period over which we have observations). The model is determined by the numbers p_d and p_w, where $p_d = P(X_t = 1 | X_{t-1} = 0)$, and $p_w = P(X_t = 1 | X_{t-1} = 1)$. Note that this probability is assumed not to depend on t. This is the assumption of stationary transition probabilities, which is commonly met by looking only at meteorologically homogeneous time periods such as a month or a season. The paper by Stern and Coe discusses ways of eliminating this assumption.

Example: The US Weather Service maintains a large number of precipitation monitors throughout the United States. One station is located at the Snoqualmie Falls in the foothills of the Cascade Mountains in western Washington. A day is defined as wet if at least 0.01 inches of precipitation falls during a precipitation day: 8 am through 8 am the following day. Using data from 1948 through 1983, and looking at January rainfall only, there were 325 dry and 791 wet days. Taking into account successive days we get the following observed counts.

Yesterday	Today Dry	Today Wet	Total
Dry	186	123	309
Wet	128	643	771
Total	314	766	1080

Natural estimates of the parameters are

$$\hat{p}_w = \frac{643}{771} = 0.834 \quad \hat{p}_d = \frac{123}{309} = 0.398$$

If we try to model several stations, we simply need to replace the simple indicator variable X_t by a binary vector, indicating the presence or absence of rain at each of the stations. If we have k stations, this model has $2^k(2^k-1)$ parameters, corresponding to the probability of each pattern of rainfall given a certain pattern the previous day. The rapid increase in parameters with network size make this model less feasible for regional modeling. In addition, the Markov chain model typically does not fit very well to network data.

A further drawback with the Markov chain model is that it does not take into account any of our knowledge of meteorological processes. In the next section we discuss a different approach, which draws upon such knowledge.

3. Using A Conceptual Rainfall Structure

Hobbs and Locatelli (1978) describe mesoscale rainfall activity in cyclonic storms roughly as follows. Synoptic scale weather fronts contain large mesoscale regions, rainfall bands, where precipitation activity is possible. In turn, these bands contain moving rain cells, which are the points of higher rainfall rates. Observing this from a fixed point in space (e.g., a rain gauge), we see varying amounts of rainfall over time, with precipitation tending to come in clusters. Mathematically, this may be described by a cluster point process. A simple such process was introduced by Neyman (1939), and is called a Neyman-Scott cluster process. It has two parts: a primary process (corresponding to weather fronts in this application) of cluster starts, and a secondary process of points, laid out in an iid fashion around the cluster starts. LeCam (1961) was first to suggest modeling rainfall at a location by a cluster point process, while Kavvas and Delleur (1981) suggested a Neyman-Scott Poisson cluster process, in which the primary process is a nonhomogeneous Poisson process, and were the first to fit it to observed data. Rodriguez-Iturbe and co-workers (Rodriguez-Iturbe et al., 1984; Valdés et al., 1985; Rodriguez-Iturbe et al., 1987) have studied different versions of this model, usually made stationary by considering only a short time period each year, such as a month. As it happens, it is not possible to use the simple approach of assuming precipitation amounts independent of the occurrence. Rather, the successful models of this type associate with each secondary event both a duration and an average rain rate. The resulting model can have precipitation from different rain cells overlapping (and accumulating) at the same point in time. Thus, the conceptual model no longer corresponds precisely to the physical

knowledge.

In most versions of cluster point process analysis of precipitation, the primary process is assumed unobserved. This may be reasonable if only rain gauge data is used. However, one would often be able to assess the arrival of weather fronts using different types of data. Kavvas and Herd (1985) used radar data to assess the process of front arrivals, and used rain gauge data to fit the secondary process. Guttorp (1986) used so-called event-based data from an acid rain monitoring network to assess features of the secondary process.

Although several workers have obtained a reasonable fit using modified versions of the Neyman-Scott Poisson cluster process, the actual fitting procedure uses method of moments, which is an inefficient estimation method at best. Foufoula-Georgiou and Guttorp (1987) discuss some of the statistical problems related to such fits. Also, the generalization to more than one station is not trivial. Rodriguez-Iturbe et al. (1986) discuss a spatial model, while Waymire et al. (1984) extend this methodology to a space-time process, but there does not appear to be much successful work on fitting their model to data.

4. Incorporating Atmospheric Processes

A different type of conceptualization was given by Zucchini and Guttorp (1991), who thought that a very simple model should be adequate to describe precipitation in meteorologically homogeneous time-space regions. They assumed that given some unobserved weather states, assumed to follow a Markov chain model, the precipitation occurrence at different stations would be independent random events, with probabilities depending on the underlying weather state. Analysis using their model shows a much improved fit over the Markov chain model, both in terms of dry spells and in terms of spatial structure.

While the conceptual weather states in the Zucchini-Guttorp model do not explicitly include atmospheric variables, but are estimated from the precipitation data themselves, it is interesting that the average meteorological structure within each weather state appears reasonable. For Snoqualmie Falls January rainfall, there are three weather states, one corresponding to high pressure with low probability of rain, one to low pressure with high probability of rain, and one to a changing pattern with intermediate precipitation probabilities.

Coming from a different direction, Hay et al. (1991) defined six different weather types (high pressure, coastal return, maritime tropical return, frontal maritime tropical return, cold frontal overrun, and warm frontal overrun), and assumed that the weather types followed a Markov chain. Conditional upon weather state, the precipitation at a station was estimated from historical data. They do not discuss extension to more than one station. A slightly different approach was taken by Wilson et al. (1991), who used a clustering algorithm to determine homogeneous weather states, and assumed that rainfall followed a Markov chain within each such weather state. The resulting model did not reproduce the temporal structure as well as did the Zucchini-Guttorp model. However, the type of weather states obtained from the clustering were fairly similar to those obtained from the Zucchini-Guttorp model.

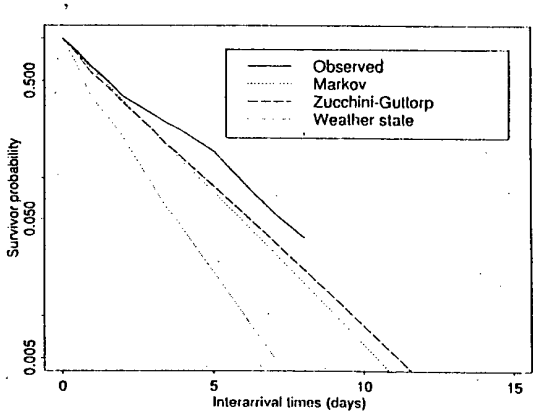


FIGURE 1: The observed and estimated probability of storm interarrivals exceeding a given number of days plotted on a log scale for three models of January precipitation at Snoqualmie Falls, WA.

Figure 1 shows the estimated survival probability (essentially the dry spell distribution) for Snoqualmie Falls January precipitation using the Markov chain model, a weather state model similar to that used by Wilson et al., and the Zucchini-Guttorp model.

As with the point process approach, it does not seem possible to adhere faithfully to a physical description while maintaining a simple probabilistic structure of space-time precipitation fields. Ideally one would like to combine the relatively faithful description of precipitation sequences, both temporally and spatially, of the Zucchini-Guttorp model with the meteorologically based approach of Hay et al. and Wilson et al. This is the topic of the next section.

5. A Generalization

We now present a more general model which contains both of the approaches presented above. The model utilizes atmospheric data (in a very general way) while maintaining the concept of a hidden or unobserved weather state which is temporally persistent. Let C_t denote the unobserved weather state at time t, while A_t stands for atmospheric data at the same time. As before, X_t is the precipitation state, which is multivariate for spatial data. We assume that X_t depends on the history of precipitation, weather states, and atmospheric variables (up through the present day) only through the current weather state:

$$\mathbf{P}(X_t|X_1^{t-1},C_1^t,A_1^t) = \mathbf{P}(X_t|C_t).$$

where, e.g., C_s^t is the history of weather states $(C_s,\ldots,C_t)$. Furthermore, C_t depends on past weather states and past or future atmospheric data only through the current atmospheric data and the previous weather state:

$$\mathbf{P}(C_t|C_1^{t-1},A_1^T) = \mathbf{P}(C_t|C_{t-1},A_t) \propto \mathbf{P}(A_t|C_{t-1}^t)\mathbf{P}(C_t|C_{t-1}).$$

If we eliminate the direct dependence on atmospheric variables the Zucchini-Guttorp model obtains. By identifying A_t with the six weather states used by Hay et al., and setting $\mathbf{P}(A_t|C_t)=1$ whenever $A_t=C_t$ the Hay et al. model is obtained. It is straightforward to allow C_t to have higher order dependence on past weather states, e.g., using the parsimonious parametrization suggested by Raftery (1985).

Using recursive procedures similar to those suggested for hidden Markov chains by Baum (1972) it is possible to compute the likelihood function under some additional assumptions. Given the atmospheric history up through the present, the history of rain and the current weather state are assumed not to depend on future atmospheric variables. In addition, we need to assume that given today's weather state and all atmospheric data (past, present and future), future precipitation does not depend on yesterday's weather state, nor on atmospheric variables up to and including today, i.e.,

$$\mathbf{P}(X_{t+1}^T|C_{t-1}^t,A_1^T) = \mathbf{P}(X_{t+1}^T|C_t,A_{t+1}^T).$$

6. Discussion

The model in the previous section does not explicitly deal with the joint spatial distribution of precipitation over the stations. The Guttorp-Zucchini model assumed independence between the stations, conditional upon weather state, as did the model by Hay et al. A more realistic assessment of dependence between stations, given the climate state, can be obtained from the literature on spatial point processes (Diggle, 1983; Ripley, 1988). A relatively simple model, borrowed from the Ising model of ferromagnets in statistical physics (*e. g.*, Preston, 1976), assumes for a vector-valued $X_t=(X_{1,t},\ldots,X_{K,t})$ that

$$\mathbf{P}(X_t|C_t=c) \propto \exp(\sum_i \alpha_{c,i}X_{i,t} + \beta_c\sum_{i,j} f_{ij}X_{i,t}X_{j,t})$$

where the f_{ij} are weights which would typically depend on the distance d_{ij} between sites i and j. The simplest weights are obtained by introducing a dependence distance d_0, and setting $f_{ij}=1(d_{ij}\le d_0)$. By letting f_{ij} be a function of the locations of sites i

and j, rather than just of their distance, it is possible to incorporate spatial inhomogeneity and anisotropy. Qualitatively, if the $\alpha_{c,i}$ are positive then rain is a likely event given climate state c, whereas if they are negative rain is unlikely (anywhere in the network). The parameter β_c introduces spatial dependence: if β_c is positive neighboring pairs are likely to have rain concurrently.

Another important aspect of the model in the previous section is the determination of appropriate atmospheric variables. From a practical point of view it is not possible to include all measured variables over all grid squares that may be thought to influence precipitation in a given region. The discussion in section 1 suggests using only spatially smooth variables such as temperature profile and surface pressure. These are also variables that directly relate to the occurrence of precipitation. Because of the smoothness of these spatial fields, it may be worthwhile to summarize them using some multivariate technique such as principal components.

In order to fit models of this type to historical data, one must use standard meteorological data. It is important to remember that such data typically are not measurements, but spatially interpolated values. Therefore analysis of such models should take into account the uncertainty in these atmospheric measures. In addition, daily precipitation is integrated over a precipitation day, while the meteorological data are snapshots once or twice a day. Therefore they may not match completely with the precipitation process. In spite of these problems, a good stochastic precipitation model developed along these lines may be applicable to probability forecasts of precipitation, using forecasted atmospheric variables from a standard numerical weather prediction model.

References

Baum, L. E., An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes, *Proc. 3rd Symposium on Inequalities*, 1-8, Academic Press, New York, 1972.

Diggle, P. J., *Statistical Analysis of Spatial Point Patterns*, Academic Press, London, 1983.

Foufoula-Georgiou, E. and P. Guttorp, Assessment of a class of Neyman-Scott models for temporal rainfall, *J. Geoph. Res. D, 92*, 9679-9682, 1987.

Gabriel, K.R. and J. Neumann, A Markov chain model for daily rainfall occurrences at Tel-Aviv, *Q.J.R. Meteorol. Soc., 88*, 90-85, 1962.

Guttorp, P., Analysis of event based precipitation data with a view towards modeling. *Water Resources Research, 24*, 35-44, 1986.

Hay, L. E., G. J. McCabe, D. M. Wolock, and M. A. Ayers, Simulation of precipitation by weather type analysis, *Water Resour. Res., 27*, 493-501, 1991.

Hobbs, P. V., and J. D. Locatelli, Rainbands, precipitation cores and generating cells in a cyclonic storm, *J. Atmos. Sci., 35*, 230-241, 1978.

Kavvas, M. L., and J. W. Delleur, A stochastic cluster model for daily rainfall sequences, *Water Resour. Res., 17*, 1151-1160, 1981.

Kavvas, M. L., and K. R. Herd, A radar-based stochastic model for short-time-increment rainfall, *Water Resour. Res., 21*, 1437-1455, 1985.

Le Cam, L. M., A stochastic description of precipitation, in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 3, edited by J. Neyman, pp. 165-186, California, Berkeley, Calif., 1961.

Neyman, J., On a new class of 'contagious' distributions, applicable in entomology and bacteriology. *Ann. Math. Statist., 10*, 35-57, 1939.

Preston, C., *Random Fields*, Lec. Notes. Math. vol. 534, Springer, Berlin, 1976.

Quetelet, A., Sur quelques propriétés curieuses que présentent les résultats d'une série d'observations, faites dans

la vue de déterminer une constant, lorsque les chances de rencontrer des écarts en plus et en moins sont égales et independantes les unes des autres, *Bull. Acad. Royale Belg., 19, Parte 2*, 303-317, 1852.

Raftery, A., A model for high-order Markov chains, *J. R. Statist. Soc. B, 47*, 528-539, 1985.

Ripley, B. D., *Statistical Inference for Spatial Processes*, Cambridge University Press., Cambridge, 1988.

Rodriguez-Iturbe, I., V. K. Gupta, and E. Waymire, Scale considerations in the modeling of temporal rainfall, *Water Resour. Res., 20*, 1611-1619, 1984.

Rodriguez-Iturbe, I., D. R. Cox, and P. S. Eagleson, Spatial modelling of total storm rainfall, *Proc. R. Soc. London A, 403*, 27-50, 1986.

Rodriguez-Iturbe, I., B. Febres de Power, and J. B. Valdés, Rectangular pulses point process models for rainfall: analysis of empirical data, *J. Geophys. Res. D, 92*, 9645-9656, 1987.

Stern, R.D. and R. Coe, A model fitting analysis of rainfall data (with discussion), *J. Roy. Statist. Soc. A., 147*, 1-34, 1984.

Valdés, J. B., I. Rodriguez-Iturbe, and V. K. Gupta, Approximations of temporal rainfall from a multidimensional model, *Water Resour. Res., 21*, 1259-1270, 1985.

Waymire, E., V. K. Gupta, and I. Rodriguez-Iturbe, A spectral theory of rainfall intensity at the meso-β scale, *Water Resour. Res., 20*, 1453-1465, 1984.

Wilson, L. L., D. P. Lettenmaier, and E. F. Wood, Simulation of daily precipitation in the Pacific Northwest using a weather classification scheme, in Wood (ed.), *Land Surface — Atmospheric Interactions for Climate Modeling: Observations, Models, and Analysis*, Surv. Geophys., *12*, 127-142, Kluwer, Dordrecht, 1991.

Zucchini, W. and P. Guttorp, A hidden Markov model for space-time precipitation. *Water Resources Research, 27*, 1917-1923, 1991.

On relationship between the variations of the precipitation in France and Tunisia and the 700 hPa height field

S. CLINET
Environnement Climatique (CNRS), Grenoble, France

A. DOUGUEDROIT
Institut de Géographie, Université d'Aix-Marseille II, and Environnement Climatique, Grenoble, France

ABSTRACT

Monthly indices of 700 hPa height field patterns are related to monthly precipitation indices of areas in France and Tunisia from 1964 to 1979 for the month of January. The results demonstrate that some pattern indices of the three previous months are significantly correlated with the regional precipitation in January. Multiple regressions linking two or three 700 hPa pattern indices with the January precipitation of each area were calculated for 1964-1977 and precipitation were estimated for 1978 and 1979.

1. Introduction

Teleconnections between precipitation and the 700 hPa height field are investigated here. Time indices representing both fields are first determined by a method close to the Rotated Principal Component Analysis (RPCA) with a Varimax rotation. Such a method has been used to investigate the variability of the climate since Steiner (1965) and of the atmosphere since Lorenz (1956) ; rotation has been widespread more recently, mainly since 1981 (Wallace and Gutzler, etc.). The main interest of the RPCA consists in condensing the information with a minimum loss into a small number of space patterns (Richman 1986, O'Lenic and Livezey 1988, etc) with an associated time index. It is supposed to represent an underlying simple structure.
Most of the teleconnection studies focused exclusively on the atmosphere circulation at different levels, especially at the 700 hPa height (Kimoto and Ghil, 1986, Barnston and Livezey 1987, etc). Investigations on teleconnections between climate and atmosphere fields are more recent (Douguédroit, 1989, Leathers and al, 1991, etc).

2. Data and methodology

Monthly precipitation data of France and Tunisia were collected from the national meteorological services of both countries. The 700 hPa field data are from the HEMIS data-set of the French meteorological office ; it is limited to the period 1963-1979.
A model of two way analysis of variance (Martin and al, 1986) which gives results close to those obtained by a Varimax RPCA was applied to both data-sets. Based on a covariance matrix, it gives for each space pattern a time index which is composed of anomalies estimated in millimeters for precipitation and in meters for the geopotential height. The retained time-scale is the month and the series limited to 17 years due to HEMIS series length. The examples of space precipitation patterns have been limited here to four areas in France and three in Tunisia (Fig. 1). The 700 hPa height field was investigated by Clinet (1990) who determined the low-frequency patterns and the associated indices. Most of them can be compared with those described by Barnston and Livezey (1987).
Each monthly time index of each precipitation pattern was regressed against the indices of patterns of the 700 hPa height field, limited to the three months preceding the month of the precipitation. Here, we only display results concerning January precipitation which were correlated with the October, November and December values of the atmospheric patterns. Two stages were carried out successively. First, simple linear correlations between all the anomalies of the precipitation of a chosen

months and of the atmospheric patterns of the three previous month were calculated two by two. Secondly, each monthly precipitation index (January for instance) was regressed against the 700 hPa monthly pattern indices of the three previous months (December, November and October for January precipitation).

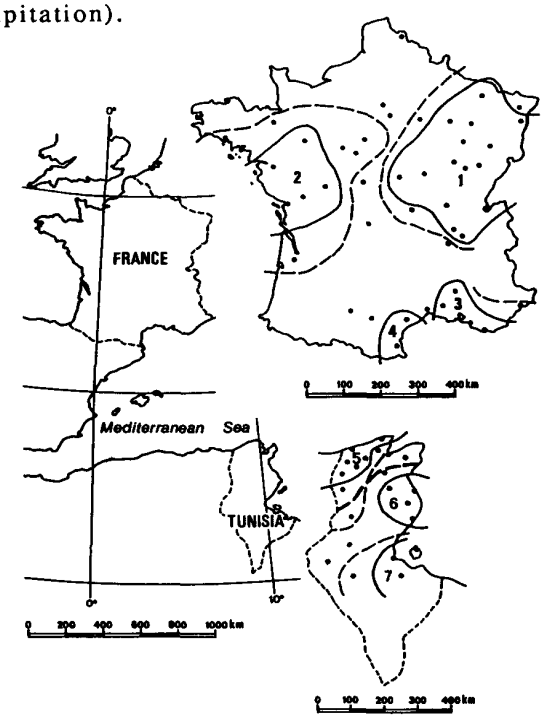


Fig. 1
Location of the precipitation areas
1 to 7 = number of the area. Explained variance by the model = > 50% (solid line), > 40% (dotted line)

3.Teleconnections between January precipitation and the 700 hPa height field

The January precipitation anomalies (1964-1979) of the seven areas previously determined have been correlated with the height anomalies of the low-frequency 700 hPa patterns of December, November and October of the previous year (Table 1). The patterns with a significant relationship at the 95% and the 90% levels have been determined. They are between two and seven and belong to the three chosen months according to the precipitation areas. The n° 1 area (Eastern France) is only correlated with patterns located in the east or the north-east (NL, Eur, NAs = see list in Annex 1). On the contrary, the n° 2 area (Western France) is mainly correlated with western patterns (NEP/EC, NCan, NAO). The two French mediterranean areas (n° 3 and 4), the January precipitation of which being significantly correlated, are both linked with NAs.

Such teleconnections can be interpreted in terms of large-scale dynamical structure of the atmosphere taking the form of propagating Rossby wave trains. The significant correlations between precipitation and the pattern anomalies of the previous months suppose an investigation of the pattern persistency and of the correlations between the patterns. Autocorrelation studies have displayed that patterns could persist with one to three month lag, with highest values corresponding with large patterns spreading over the oceans as NCP, NEP/EC, NEP, CAtl, NL. And the variations of the pattern anomalies can be correlated with other patterns. For instance, NAs values of December are significantly correlated with NEP/EC (-.49) and with Eur (-.70) of the same month and with EAtl (-.51) of the previous month, November.

4.Attempt to predict monthly precipitation

Individual monthly precipitation indices have been regressed against the 700 hPa pattern indices of the three previous months for the seven areas during three periods of records (1964-77, 1964-78, 1964-79 for precipitation and 1963-76, 1963-77 and 1964-78 for the 700 hPa field for January). In all cases, we have kept the significant higher correlated equations, the concerned patterns of which being common with the three attempts, and kept the coefficients of the 1964-77 regression (Table 2, January).

Regressions depend mainly on the significant correlated patterns previously determined (see Table 1). Regressions concerning areas 7 and 8 in Tunisia have been eliminated because they were based on the NAf which includes a bias in the HEMIS data-set (Clinet, 1990).

Areas	Patterns											
	a	b	c	a	b	c	a	b	c	a	b	c
1	NL	11	58	Eur	12	57	Eur	11	53	NAs	12	-51
2	NEP/EC	10	62	NCan	10	53	NAO	10	-52	Eur	10	-52
3	NAs	12	51	NCP	10	45						
4	NCan	11	-58	NAs	12	49						
5	Scan	12	64	NL	12	-62	NL	10	50	NEP/S	11	50
	NL	10	49	NPO	10	48	EAtl	12	-47			
6	NEP/EC	12	55	NEP	12	-55	NCan	10	50	EAtl	12	47
7	CAtl	12	61	NAf	11	-58	NAs	12	52	NEP/EC	12	50
	NAf	12	-50	NAf	10	-50						

Table 1 - Significant correlation (at 95% and 90% level = .50 and .43) between the variations of January precipitation (1964-1979) and the variations of the previous month 700 hPa patterns (10 = October, 11 = November, 12 = December).
1 to 7 = Area number (Fig. 1) ; a = name of the pattern (see annex 1) ; b = month ; c = correlation coefficient (x 100) ; $\bar{x}$ = Mean January precipitation (1964-1979).

1964-77 regressions have been used to calculate predicted monthly precipitation for 1977, 1978 and 1979 precipitation. Observed and predicted results have been compared (Table 3). As in an area, anomalies do not present equal values in millimeters while having the same trend, existing departures between observed and predicted values can be compared to the averages and appreciated through the repartition into the five quints of each series. All the obtained results are very fragile because of the limited length of the series.

Area	%	Regression
1	75	-1.08 - .53 (NWP/S 12) + .54 (NEP/EC 11) + .41 (NL 11)
2	69	7.74 +.42 (NEP/Ec 10) +.72 (NCan 10)
3	72	12.18 + .49 (NAs 12) + .71 (NWP/S 12) + .72 (NcP 10)
4	75	7.32 +.36 (NAs 12) + 2.0 (NCan 11) + 6.7 (NWP/S 11)
5	67	-6.7 + 1.32 (Scan 12) + 1.77 (NWP/S 11)

Table 2 - Regressions linking January precipitation to the 700 hPa patterns of the previous months.
1 to 5 = number of the area (Fig. 1)
Eur12 = name of the pattern (see Annex 1) and number of the month (10 to 12 = October to December).%=Percentage of explained variance.

Areas		a	b	c	d	e
1	$\bar{x}$		31	28	31	40
	1977	24	53	42	51	34
	1978	3	24	9	16	35
	1979	46	17	-7	17	-22
3		a	f	g	h	
	$\bar{x}$		46	55	69	
	1977	36	37	19	53	
	1978	19	140	19	104	
	1979	24	36	58	75	

Table 3 - Estimated and observed precipitation anomalies for 1977, 1978 and 1979 January. Estimated anomaly for the whole area (a). Observed anomalies in Besançon (b), Langres (c), Luxeuil (d), Neufchateau (e), Marseille (f), Orange (g) and Toulon (h) - $\bar{x}$ = Mean January precipitation (1964-1979).

5. Conclusion

Variations of January precipitation index in several areas of France and Tunisia have been compared with 700 hPa field pattern indices of the three previous months. All the indices have been obtained by the model of two way variance close to the Varimax RPCA. Significant correlations exist between the precipitation and the atmospheric patterns of the previous months during the 1964-1979 period. Multiple regressions calculated for the 1964-1977 period have been used to estimate the 1977,

1978 and 1979 January precipitation. Predicted and observed values have been compared. But the small length of the series strongly limits the interest of the obtained results.

Annex 1.

List of the 700 hPa height field patterns
NAf = Northern Africa
NCP = North Central Pacific
EAtl = Eastern Atlantic
NAO = North Atlantic Oscillation
NAs = Northern Asia
NPO = North Pacific Oscillation
NEP/EC = North East Pacific/East Canada
Scan = Scandinavian
WEu = Western Europe
Eur = Eurasian
NEP = North East Pacific
CAtl = Central Atlantic
NL = Northern Latitudes
NCan = Northern Canada
NEAm = North East America
NWP/S = North West Pacific Siberia

References

Barnston,A.G.,Livezey,R.E.,1987. Classification; Seasonality and Persistence of Low-frequency atmospheric circulation pattern. *Mon. Wea. Rev.* 115, 1083-1126.

Clinet, S., 1990. *Variabilité du champ de pression sur l'hémisphère nord extratropical. Analyse des géopotentiels décadaires et mensuels à 700 et 500 hPa.* Thèse. Grenoble. 381 p. Annexes.

Clinet, S., Martin, S., 1992. 700 hPa Geopotential Height Anomalies from a statistical analysis of the French HEMIS data-set. *Intern. J. Climatol.* (in press).

Douguédroit, A., 1989. Relations between monthly precipitations in north-western Africa and the 700 hPa height field. *Annales Geophysicae*, Special issue, XIII General Assembly, 89.

Kimoto, M., Ghil, M., 1986. *A statistical study of 700 mb Height Anomalies using Rotated Principal Components and cluster Analysis Rechearch.* Report n 1 of the joint NMC/UCLA Project on Climate Dynamics,Statistics and Forecasting, 48 p.

Leathers DI, Yarnal B., Palecki M.A., 1991. The Pacific/North American Teleconnection Pattern and United States climate. Part 1 : Regional Temperature and Precipiation Associations. *J. Climate*, 4, 5, 517-523.

Lorenz E.H., 1956. *Empirical orthogonal Functions and statistical weather predictions.* Sci. Report 1. Statistical Forecasting Project. Dpt of Meteorology, M.I.T., Cambridge.

Martin, S., Antoniadis, A., Degerine, S., Grégoire, G, Lebreton, A., Charre, J., 1986. Models of two ways analysis of variance for the study of precipitation data. Extended abstracts. *25th International Geographical Congress.* Paris, IGU, 310.

O'Lenic, E.A., Livezey, R.E., 1988. Practical considerations in the Use of Rotated Principal Components Analyses (RPCA) in Diagnostic Studies of Upper-Air Height Fields. *Mon. Wea. Rev.* 116, 1682-1689.

Richman, M.F., 1986. Rotation of principal components. *J. Climatol.* 6, 7, 511-520.

Steiner, D., 1965. A multivariate statistical approach to climatic regionalization and classification. *Tijdischrifte van het koninkilk nederlansch aardrijskundig genootschop*, 82, 4, 329-347.

Wallace, J.M., Gutzler, D.S., 1981. Teleconnections in the Geopotential Height Field during the Northern Hemisphere Winter. *Mon. Wea. Rev.*, 109, 784-812.

On the relationship between the atmospheric circulation in the Atlantic-European area and some seasonal meteorological parameters in Central Europe

Peter C. Werner
Potsdam-Institut für Klimafolgenforschung
Telegrafenberg, O-1561 Potsdam, Germany

1. Introduction

In the present study we consider the space-time variability of Central European temperature in winter in the interval 1901-80. The area "Central Europe" is represented by 11 stations: Potsdam, Hohenpeißenberg, Frankfort, Jena and Hamburg (Germany), Prague (Czechoslovakia), Uccle (Belgium), Vienna (Austria), Geneva, Zurich (Switzerland), Fanoe (Denmark). We define "winter" as January/February. In a further step we derive Empirical Orthogonal Functions (EOFs) from the JF mean temperatures at the 11 stations and the winter mean of the circulation. As parameter that represents the large-scale circulation we chose the sea-level pressure (SLP) field on a 5^{o} x 5^{o} grid from 35^{o} N to 75^{o} N and from 50^{o} W to 40^{o} E.
In section 3, we analyse the relationship of the Central European temperature field and the large-scale circulation by means of a Canonical Correlation Analysis (CCA).
In section 4, we examine the consistency of trends in the large-scale circulation and in the temperature.

2. Canonical correlation analysis of the regional temperature and the sea-level pressure field

The CCA (Anderson, 1984; Mardia et al., 1979) is done in subspaces spanned by the first few EOFs. This procedure has the double advantage that the input-data sets are independent (orthogonal functions) and that much of the noise in the data-field is eliminated. The estimated correlation coefficients obtained in a CCA are positively biased: The more EOFs are used the larger the correlation coefficient become. We try to avoid the bias by applying the following concept: We expect the largest correlation coefficient to grow rapidly with increasing number of EOFs as long as the added EOFs represent additional information. If the extra EOFs represent mostly noise, however, the correlation coefficient is expected to grow slowly. This procedure lead us to the choise of 5 EOFs (or less) of temperature and sea-level pressure.
The CCA are carried out with the sea-level pressure and temperature data from the "estimation" interval 1901-40; the explained variances and the correlation coefficients are calculated from the complete interval 1901-80 including the "test" interval 1941-80.
The SLP pattern of the first CCA pair describes an anomalous southwesterly flow into Central Europe (Fig. 1a). This southwesterly flow advects anomalously warm maritime air so that the temperature pattern is positive everywhere with maximum values of almost 2K along the Alps and minimum values of 0.7K at Fanoe (Fig. 1b). In Fig. 1b also the rate of explained local variance, as derived from the test sample 1941-80, is given. This rate is larger than 40% everywhere except for Zurich (31%), Hamburg (37%) and Fanoe (22%). Apparently, the 1st CCA pair controls mostly the southern part of the analysis area.
The second CCA pair (Fig. 2) is in control of the northern part of the analysis area. The sea-level pressure field represents a strong northwesterly flow that affects mostly the northern part of Central Europe. There the typical temperature anomalies are almost 2K (at Potsdam, Hamburg and Fanoe), and the rates of explained local variance is more than 40%. Along the Alps, at Geneva and Zurich, the typical anomalies are 0.6K with explained variances of less than 10%.

3. Consistency of trends in the circulation and in the temperature

We may use the results of the CCA of the previous section to indirectly derive temperature anomalies exclusively from the anomalous circulation. This "CCA-model" is given by Storch et al. (1992). The power of this model is assessed by calculating the explained variance and the correlation in the 1941-80 test subset. All variances and correlations are calculated with anomalies relative to the 1901-40 means. Not surprisingly, the first CCA pair is sufficient for the southern stations (e.g., Geneva), whereas the second CCA pair is need for the northern stations (e.g., Fanoe). The addition of the third CCA pair does not improve the skill of the CCA model. Therefore, we have to chose the first two modes (m).
With the m = 2 model more than 50% of the variance are described in the north. The maximum rate is 64% at Hamburg, the minimum is 18% at Zurich. The correlation is everywhere high, with average values of 80%. This high level indicates that the year-to-year fluctuations are captured quite well by CCA model, whereas the variability of explained variance indicates that the CCA model has, at least at some locations, difficulties to reproduce the actual size of the anomalies. That this is indeed so is shown by the time series of observed and reconstructed temperature at the 11 Central European stations (Fig. 3). The similarity of the high frequency variations is good not only for the estimation interval 1901-40 but also for the test interval 1941-80. That the CCA model does not correctly describe the trends is obvious:
The differences between observations and reconstructions are large at all stations.
According to the in-situ observations did the temperatures decrease in the north (-0.010 K/year at Hamburg) and increase in the south (0.022 K/year at Zurich and 0.011 K/year at Geneva). One might speculate that these trends might be contaminated by the spurious warming by urbanization. Two of the stations are not from urban areas: Fanoe

and Hohenpeißenberg. The trends of urban stations are consistent with the trend at these two non-urban stations so that urbanization likely is not controlling our data.

The reconstructed trends by SLP field (Fig. 4a) deviate marked from the trends in Fig. 4b. According to the changes of the circulation, the temperature should have decreased everywhere in Central Europe by -0.01 to -0.03 K/year in the last 80 winters. The largest difference is at Zurich, namely -0.038 K/year, and we propose that the Zurich time series is contaminated by an urbanization effect.

4. Conclusions

The main conclusions to be drawn are:

a) The variations of the winter mean large-scale North Atlantic/European circulation control large parts of the variations of Central European temperature. This is in particulare so for the year-to-year variations.

b) The 80-year trends of the circulation and of the regional temperature field is of unimportance in case of the used statistical methods.

c) The regional temperature on time scales of several decades is controlled not only by the circulation.

References

Anderson, C.W. (1984): An introduction to multivariate statistical analysis. 2nd Ed. Wiley & Sons

Mardia, K.V., Kent, J.T., Bibby, J.M. (1979): Multivariate analysis. Academic Press

Storch, H.v., Zorita, E., Cubasch, U. (1992): Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in Wintertime. J. Climate (in press)

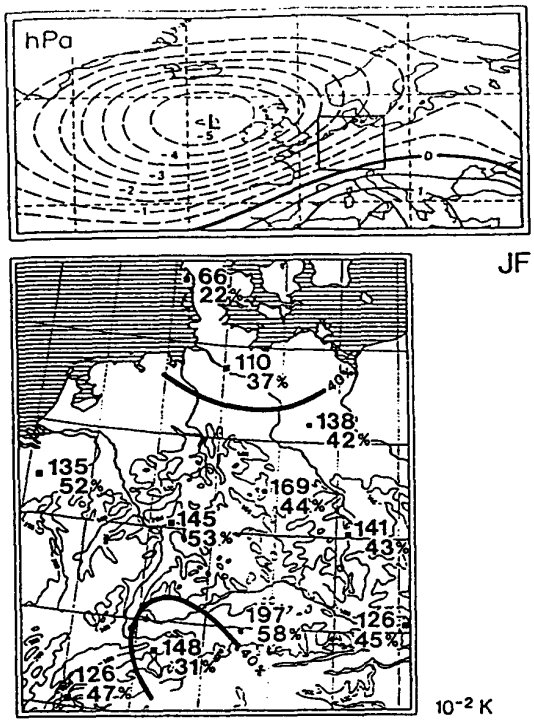


Fig. 1 First CCA pair of SLP (a) and seasonal mean temperature (b) for JF derived from the 1901-40 subset

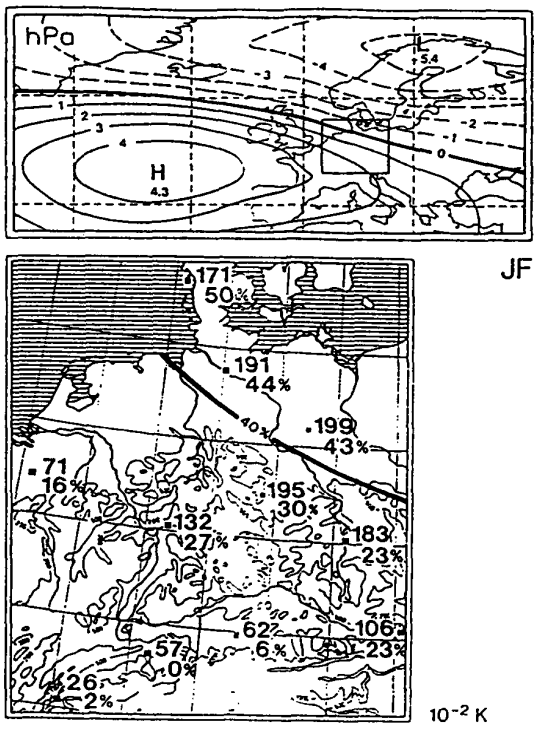


Fig. 2 Second CCA pair of SLP (a) and seasonal mean temperature (b) for JF derived from the 1901-40 subset

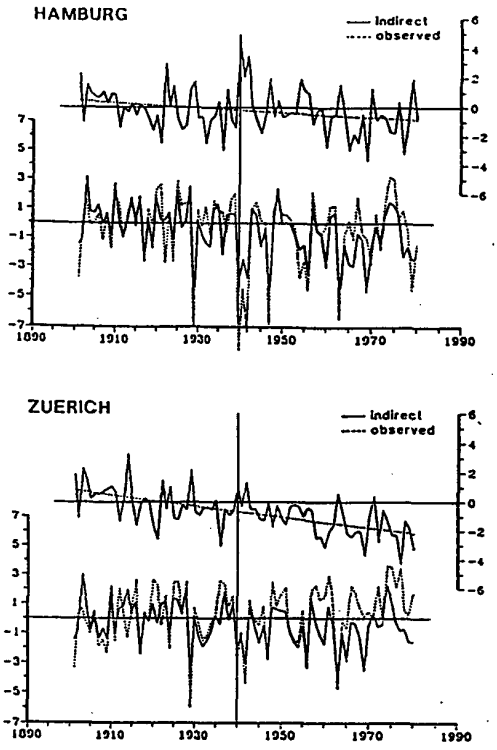
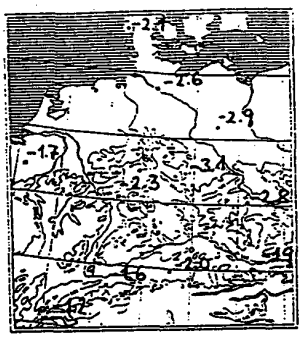


Fig. 3 In-situ observations and reconstructions, by means of concurrent sea-level pressure anomalies and model with n = 2, of local JF mean temperature at 2 Central European stations. The time interval 1901-40 was used to fit the model, and the time interval 1941-80 represents independent data. A linear trend derived from all data is added.
Top: Difference between reconstructions and in-situ observations.
Bottom: Anomalies derived from in-situ observations (dashed) and reconstructed anomalies (solid).

a) From in-situ observations



b) From reconstructed temperatures

Fig. 4 1901-1980 trends of temperature at 11 Central European stations (K/100years).

A Method for Mapping Mesoscale Climatological Fields in Data-Sparse Arctic Regions

John D. Jacobs

Department of Geography, Memorial University of Newfoundland
St. John's, Newfoundland, Canada A1B 3X9

ABSTRACT

The method of kriging is applied to the mapping of mean climatological fields in an arctic area with few permanent observing stations. Records from short-term stations are incorporated to improve the resolution of the interpolation model.

1. Introduction

In high latitudes, the density of meteorological stations is generally much less than that considered adequate for regional representativeness, and the length of record from many of these stations is too short to provide the baseline against which regional trends might be measured. Extensive areas of the Canadian Arctic, for example, are unrepresented by the networks, even in terms of the basic climatic elements: temperature and precipitation. These mesoscale areas (10^3 to 10^5 km^2 in this context) encompass significant topographic features, such as wetlands, lakes, and ice caps, that are of interest intrinsically and in relation to regional climatic change. As topoclimatic systems, they respond to any significant change in regional climate and likely provide some feedback to that process.

This paper describes an approach to improving the mesoscale climatology of arctic regions, with application to Baffin Island. This area of nearly 500,000 km^2 currently has permanent climatological stations at 16 sites, all but one a coastal location. In addition, records of varying length exist for 8 stations that are no longer in operation. This archive is supplemented by short-term records from expeditions and by automated climate stations installed for purposes of this study at four inland locations.

Seasonal statistics and a measure of trends have been obtained for a nominal 30-year period using records from the longest-standing stations. These provide a framework for establishing the regional temperature and precipitation fields, using objective methods. The temporary and short-term station records are then accommodated, using the previously established parameters as a first-guess structure, and a revised structure function is produced by trial and error. The resulting function is used to produce a regional mapping of the climatological fields, including standard errors.

2. Analytical Approach

The general problem of determining the spatial structure of a continually varying random variable from observations at irregularly spaced, discrete points using an interpolation function derived from the spatial covariance of the variable has been treated extensively in the geophysical literature (Journel, 1989). This technique, known as kriging, is computationally intensive but generally accessible in geostatistical computer software (Englund and Sparks, 1988). The method is increasingly being used in climatology and hydrology (e.g. Finkelstein,

1984; Bigg, 1991).

Kriging involves examination of the covariance C(h) of a variable Z(x) at different lag or separation distances h, to produce a semivariogram of the form:

$$2g(h) = [1/N(h)]\Sigma[Z(x)-Z(x+h)]^2$$

where N(h) is the number of pairs with lag separation h. A regression model is found that provides the best fit to the semivariogram. This may be one of several possible forms; however, in the present case it is a *linear-sill* model:

$$g(h) = C_0 + C \qquad \text{for } h > A$$

$$g(h) = C_0 + (C/A)h \qquad \text{for } h \leq A$$

where C_0 is the "nugget", C is a measure of the structural variance called the "sill", and A is the range, i.e. the lag distance at which g(h) approaches a constant value. Since values of the parameters are obtained directly from the observational data, they provide both an objective means of interpolating values of the variable at intermediate points within the region and a measure of the error in the interpolation.

3. Application

In this example, the 1951-80 July mean daily temperatures were analysed for a 28,000 km^2 area of Baffin Island, centred approximately at 70°N. An orthogonal coordinate system with origin at 60°N, 90°W was used. The published normals for the five permanent stations in the area (see Table 1) were first mapped using a simple inverse-square routine (Fig. 1a). Because elevations range from sea level to over 1 km, potential temperatures were similarly mapped for comparison (Fig. 1b).

The records were searched to find those years in which the greatest number of stations existed in the area. During 1962-64, observing stations operated at two short-term radar sites, at a geological camp, and at a glaciological research site. Kriging was employed on the mean daily temperatures from all the stations for each of those three years. Figs. 1c and 1d, respectively, show the resulting temperature plots and standard deviations for 1963. Finally, by taking mean values of the model parameters, a generalised structure function was obtained which was applied to the normals to produce a revised temperature map (Fig. 1e) and a map of the standard deviation of those estimates (Fig. 1f).

TABLE 1. Listing of stations used in the example. C refers to first-order and synoptic stations, the same code followed by * means a discontinued station, while E refers to expedition camps.

C6 Cape Hooper	C20* Rowley Island
C7 Clyde River	C21* West Baffin
C8 Dewar Lakes	E1 Barnes Ice Cap
C11 Hall Beach	E5 Mary River
C13 Longstaff Bluff	

4. Discussion and Conclusion

The example shown demonstrates the usefulness of interpolation by kriging, combined with short-term station data, to produce an improved mapping of the climatological field. Each of the maps derived from kriging involves gridpoint interpolations at 20 km intervals. This should not be taken too seriously, in view of the amount of dispersion in the error maps. However, it points to the power of the technique as more information is added.

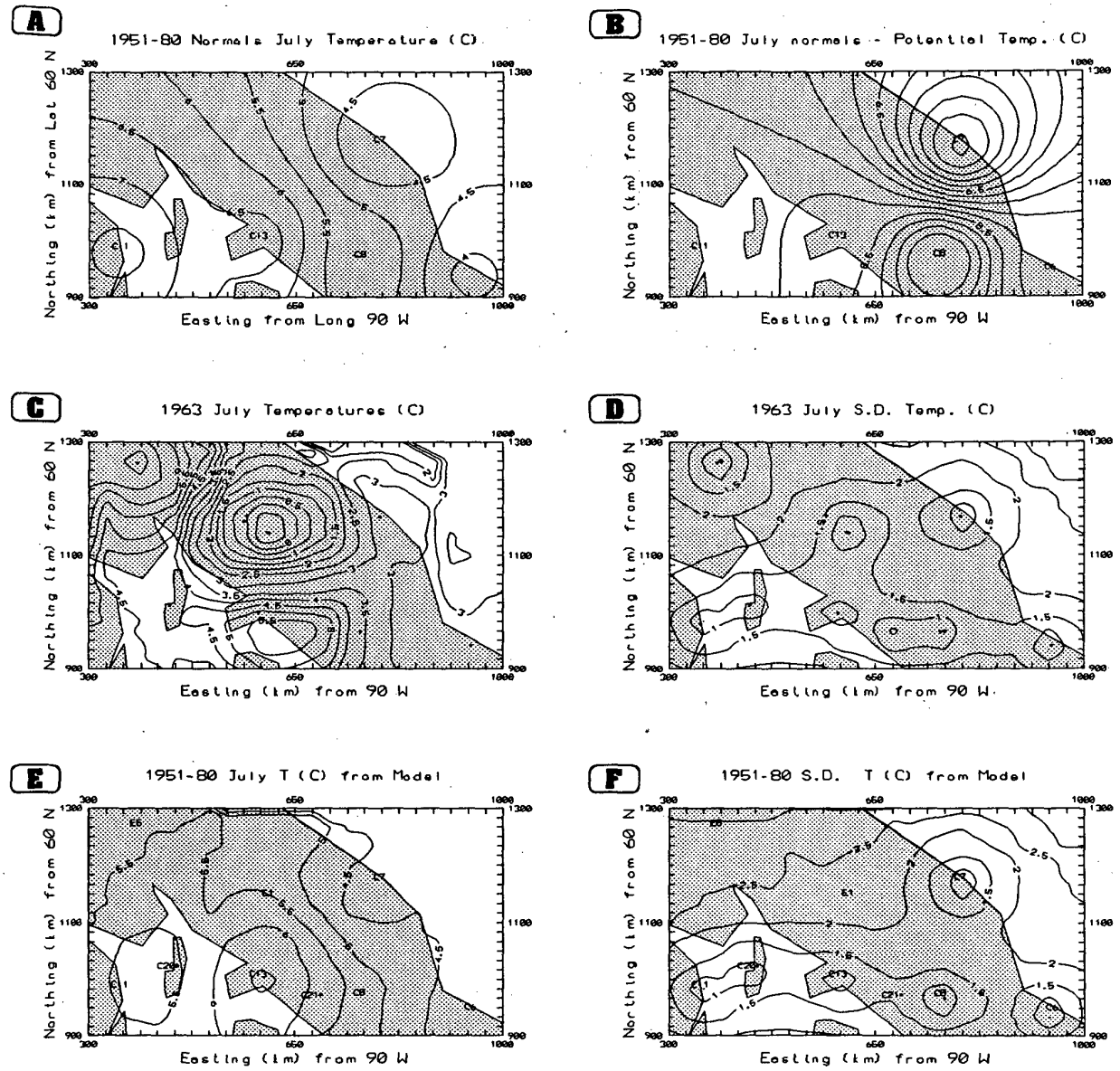


FIG. 1. Mapping of July temperature fields for the central Baffin Island region: (a) mean temperatures and (b) potential temperatures from 1951-80 normals based on permanent stations only; (c) 1963 July temperatures and (d) standard deviations including short-term and expedition stations; (e) adjusted 1951-80 means and (f) standard deviations based on kriging model.

While this approach is useful in defining mesoscale temperature patterns and, equally importantly, providing a measure of the uncertainty of the interpolated values, it is clear that a strictly objective mapping will always be limited by the degree to which the observing sites actually represent the diversity of topoclimates in the region. Surface isotherms which cross coastlines without inflection are clearly unrealistic, for example.

The short-term use of automated stations at selected sites was instituted as part of this study in 1987 (Jacobs, 1990). They will continue to be essential, particularly in the underrepresented inland areas, but there are obvious practical limits to their deployment.

An alternative method for estimating relative departures from regional patterns involves use of simple physical models of the surface climate, supplemented by satellite-derived surface temperatures and albedos. All such data, whether real or synthetic, can be accommodated in the model presented here. The result is a synthetic mesoscale climatology which may be used in the absence of subsequent observations from within those areas.

Acknowledgments. This research was funded with a grant from the Natural Sciences and Engineering Research Council of Canada. The Baffin Island automatic climate station project is supported by the Canadian Climate Centre, with field support from the Polar Continental Shelf Project.

REFERENCES

Bigg, G. R., 1991: Kriging and intra-regional rainfall variability in England. *Int. J. Clim.*, **11**, 663-675.

Englund, E., and A. Sparks, 1988: *Geo-Eas User's Guide*. U. S. Environmental Protection Agency, Las Vegas.

Finkelstein, P. L., 1984: The spatial analysis of acid precipitation data. *J. Clim. Appl. Meteor.*, **23**, 52-62.

Jacobs, J. D., 1990: Integration of automated station data into objective mapping of temperatures for an arctic region. *Clim. Bull.*, **24**(2), 84-96.

Journel, A. G., 1989: *Fundamentals of Geostatistics in Five Lessons*. American Geophysical Union, Washington.

REGRESSION APPROACHES TO RETAINING SPARSE DATA IN SPATIAL SERIES

Paul A. Kay
Geography, University of Nebraska, Lincoln, NE 68588-0135
Istvan Bogardi
Civil Engineering, University of Nebraska, Lincoln, NE 68588-0531
Istvan Matyasovszky
Civil Engineering, University of Nebraska, Lincoln, NE 68588-0531
(on leave from Meteorology, Eotvos University, Budapest, Hungary)

Hydrological modelling of impacts of climatic change and assessment of precipitation enhancement programs are two applications that may require time series of spatial data, such as areal means, totals, or gradients. A seemingly dense station network may yield sparse data if many of the stations operated intermittently or only for a few years. Thus it would be advantageous to retain as much information as possible in the development of such spatial time series, even to the point of including stations that operated for only one year. Most existing methods of spatial interpolation discard stations with sparse or short records, thus reducing the density of sites considered in the modelling effort. This paper considers two approaches to the inclusion of short records to maintain high spatial density.

The first approach considers the mean and variance of precipitation, assuming they are functions of elevation only. Strong correlation of the first order trend surface of elevation with latitude and longitude (that is, the elevation gradient) indicates such a situation, but is not generally necessary for the technique. The mean precipitation m_i at elevation z_i is estimated by kernel regression from the precipitation data c_{kj} for k stations in j years, as

$$\hat{m}_i = \frac{\sum_k \sum_j c_{kj} K\frac{z_i - z_k}{b}}{\sum_k \sum_j K\frac{z_i - z_k}{b}}$$

where K is an appropriate kernel function. The bandwidth, b, is estimated by crossvalidation by minimizing

$$\sum_i \sum_j (c_{ij} - \hat{m}_i(i))^2$$

where

$$m_i^{(i)} = \frac{\sum_{k\neq i} \sum_j c_{kj} K\frac{z_i - z_k}{b}}{\sum_{k\neq i} \sum_j K\frac{z_i - z_k}{b}}$$

The estimation of variance proceeds in a similar fashion, using $(c_{ij}-m_i(i))^2$ in place of c_{ij}.

The second approach considers the correlation between station records, assuming it is a function of the distance between stations only. The assumption of a homogeneous and isotropic precipitation field is rough. Thus, the (x,y) coordinates are first transformed to

$$x' = (x \cos \phi + y \sin \phi)$$
$$y' = -(x \sin \phi + y \cos \phi) \sqrt{\lambda}$$

where ϕ is the angle of rotation of the original grid and $\lambda>0$ is the elongation in the y' direction. The correlation matrix is then approximated by

$$R_{ij} = \exp(-\alpha d), \text{ where } d = \sqrt{x'^2 + y'^2}$$

The parameters α, ϕ, and λ are determined by minimizing

$$\frac{1}{n} \sum_{ij} [r_{ij}-R_{ij}]^2$$

where r_{ij} are correlations estimated directly from the sample, n_{ij} are the numbers of such pairs (i,j) when precipitation occurs, and n is the sum of n_{ij}.

The two approaches are illustrated by considering the Golan region of Israel. In the period 1968-69 to 1983-84, 49 stations recorded data in the area of approximately 1100 km^2, but only seven of them have records of at least ten years, and only four have data for 14 of the 16 years. The Golan plateau slopes upwards towards the north-northeast with a linear gradient of about 25 m/km. The coefficient of determination for the first-order trend surface is 0.80, so the assumption that elevation is a function of latitude and longitude is essentially valid. The results of a study of annual total precipitation are reported. The first approach yields a mean squared error of estimate that is about 50% of the total variance estimated directly from the station data. The second approach gives a weighted squared error that is about equal to the average variance of the directly estimated correlations. These results suggest that good approximations may be obtained from optimal interpolation techniques when means, variances, and correlations are known.

INVESTIGATION OF THE RELATIONSHIP BETWEEN THE MOISTURE BUDGET COMPONENTS OVER CENTRAL NORTH AMERICA IN SUMMER 1979, BY THE USE OF SPECTRUM ANALYSIS

Abraham Zangvil[1], Diane H. Portis, and Peter J. Lamb

Cooperative Institute for Mesoscale Meteorological Studies (CIMMS)
University of Oklahoma
Norman, OK 73019

The water vapor budget for summer 1979 is studied for a rectangular area (1300 x 750 km) centered over east central Illinois and extending across most of the upper Midwest. The computations include: (1) area-averaged rainfall (obtained from approximately 590 recording raingauges); (2) area-averaged time changes of total precipitable water (obtained from the NWS rawinsonde network interpolated into a 190 km grid); (3) area-averaged vertically integrated water vapor flux divergence. The budget is calculated for consecutive 12 hour periods between 00 and 12 UT, 12 and 00 UT etc. Area averaged evapotranspiration is obtained as the residual of the water vapor budget equation. Quite realistic values are obtained for the time mean daily evapotranspiration and also for night versus day values.

A pronounced diurnal variation is found in the water vapor flux divergence, evapotranspiration and the total precipitable water, and to a lesser extent in precipitation. Specturm and cross-spectrum analysis is employed to identify the dominant time scales of the water vapor budget components and their coherence and phase relationships. It is found that precipitation, water vapor flux divergence and total precipitable water vary on pronounced time scales of 1, 2, 4-5, and 11 days, while evapotranspiration showed time variations mainly in the 1-3 day range.

High coherence is found between precipitation and the velocity divergence part of the water vapor flux divergence near 11 and 3 days between 700 mb. Above 700 mb, a high coherence is found between precipitation and the water vapor advection part of the flux divergence. Precipitation is also found to have high coherence with the time change of precipitable water near 1, 2, 4-5, and 11 days. For the 1 day cycle a 180° phase difference is found, suggesting that the weak nocturnal maximum in rainfall is associated with a depletion of precipitable water. Precipitable water is highly coherent with precipitation at 11 and 5 days with almost an in-phase relationship. High coherence (near 0.95) is found between the flux divergence and precipitable water below 700 mb at 11 and 5 days and also between evapotranspiration and the time change of precipitable water (0° phase) at the diurnal period (evaporation is increasing precipitable water). Also, evapotranspiration is coherent with precipitation at the 11 day period, with the maximum evapotranspiration occurring 2 days after the precipitation maximum. The above results establish the usefulness of water vapor budget calculations on short time scales of days.

---

1 On leave from the Jacob Blaustein Institute for Desert Research, Ben Gurion University of the Negev, Israel.

GCM Response to an Arctic and an Antarctic Polynya

Rita Glowienka-Hense, Andreas Hense, Meteorologisches Institut
Universität Bonn, Auf dem Hügel 20, W-5300 Bonn 1

April 29, 1992

1 *Introduction*

Sea ice is an important aspect in the coupled atmosphere-ocean-sea ice system. Its anomalies exist over several months and are therefore much more persistent than atmospheric anomalies. If these anomalies have a significant influence on the atmospheric circulation they might be a helpful tool for longer range predictions. To find out, how the atmosphere is affected by a polynya, we have performed two different GCM experiments. The first experiment has been one with a polynya in the Arctic ocean (N-polynya) and the second one with a polynya in the Weddell sea (S-polynya) which is part of the Antarctic ocean. A polynya is an open water region within the sea ice bounds which has an extend of several thousand square kilometres. Both experiments have been performed in permanent single month mode, which was January for the Arctic and July for the Antarctic polynya run. This provides a relatively large number of realizations of single months, which is needed for the test statistic that helps to single out the atmospheric reaction to the polynya from noise. Both experiments are compared to a corresponding control run with standard ice distribution. The GCM used is a descendent of the ECMWF weather forecast model with a T21 resolution. The N-polynya run has been performed with the socalled cycle 17 version of the model and the S-polynya run was done with the later ECHAM1 version both derived at the MPI for Meteorology and the Meteorological Institute in Hamburg. The Arctic polynya consists of 13 model grid points and has an area of about 850000 km^2. It is situated in the Kara sea east of Novaya Zemlya. The Antarctic polynya consists of 15 model grid points and 2135000 km^2 and was placed in the eastern Weddell sea. At this location a polynya was actually observed during the southern winters 1974-76.

2 *Methods of Analysis*

To compare objectively the mean fields of control and polynya runs we made use of a hierarchical Hotelling-T^2-statistic as described in Hense et al (1990). The first step of this method is to reduce the spatial degrees of freedom by projecting the grid fields onto an a priori chosen set of basis functions. In most cases we used spherical harmonics, which are the eigenmodes of a diffusion equation. Analysing the 300 hPa streamfunctions we also used the eigenmodes of the linearized barotropic vorticity equation. The modes were sorted according to increasing eigenvalue because those modes which have the smallest eigenvalues are most easily incited by a forcing. In this vector space spanned by the new basis functions with reduced spatial degrees of freedom we then tested the Nullhyothesis of equal means between control (μ_c) and polynya (μ_p) runs against the Alternativhypothesis of different means.

$$H_0 : \mu_p = \mu_c$$
$$H_1 : \mu_p \neq \mu_c$$

In case the Nullhypothesis was rejected at the 5% error level, we accepted the Alternativhypothesis and performed an additional test of

global and local recurrence of the significant signal. We did this because only those signals are of practical importance which allow with a high confidence to distinguish a single new control from a polynya realization. For the northern and southern polynya experiments we analysed the mean circulation fields and the eddy statistics as characteristics of the circulation.

3 *Response of the Mean Flow*

A direct impact of the polynyas on the atmosphere is an increase of the local fluxes of sensible and latent heat. In case of the Arctic polynya a local increase of the heat fluxes by about 80 W/m^2 exist above the polynya which is similar in the Antarctic polynya run, however for the latter differences of similar amplitude are present on the whole southern hemisphere. A comparable picture is seen in the differences of the relative topography 850/1000 hPa control minus polynya for both experiments. Again the response is only localized above the N-polynya whereas the S-polynya excites a hemispheric response. Also the temperatures fields are only locally changed in the N-polynya experiment. In the S-polynya experiment the lower atmosphere south of 60°S is warmed by the additional heat from the polynya whereby at the same time the extrapolar continents are cooled by up to 2°C. In the upper troposphere at 300 hPa cooling extends from the region equatorward of about 60°S.

Using spherical harmonics as test functions we find no significant response for the geopotential and SLP fields between control and N-polynya run. However we find highly significant and recurrent differences between the S-polynya and the associated control run. Consistent with the cooling of the area above the extrapolar continents in case of the S-polynya the extra polar continents is are characterized by anomalous high pressure in the lower troposphere and low pressure in the upper troposphere, where this zone extends south of 60°S.

The eigenmodes of the linearized barotropic vorticity equation are then used as guess pattern for testing the 300 hPa streamfunction differences of the N-polynya experiment. A response in this direction of phase space indicates that the model response to the N-polynya are compatible with barotropic vorticity physics. Indeed we find a large scale, significant and highly recurrent change of the streamfunction fields.

4 *Synoptic Eddy Activity*

The synoptic eddy activity plays an important role for the meridional heat transport which is necessary to balance the solar differential. We next looked whether the synoptic eddy heat transport is affecteed by the polynyas. The test procedure showed a significant decrease of the meridional heat transport due to the N-polynya above the western Atlantic at 700 hPa with a maximum of 1 Km/s (Fig. 1a). The transports at the other levels up to 300 hPa seemed unaffected. This means that the vertical gradient of the meridional eddy heat transport is changed which is the baroclinic forcing term in the quasigeostrophic vorticity equation. Also the barotropic eddy forcing term $\nabla \overline{v'\zeta'}$ at 300 hPa has significantly changed between the N-polynya run and the control. The response is different for the S-polynya experiment. Here the meridional eddy heat transport even strongly increased at 850 hPa in the anomaly run compared to control (Fig. 1b). The increase occurs in a circumpolar zone at midlatitudes with a maximum downstream of the polynya of 1.5 Km/s whereby the maximum transport is near 4 Km/s in the control run. On the other hand no significant change of the barotropic eddy forcing term at 300 hPa is seen for the S-polynya. Also the vertical heat transports are influenced by the heat sources. For the Northern and the Southern polynya experiments we find a significant change of the synoptic transports at 500 hPa. The N-polynya produces an anomalous circulation with upward heat transport above the western Pacific and downward heat transport above the eastern Pacific upstream of the region of the anomalous southward transport in the western Atlantic. The S-polynya run is characterized by increased vertical eddy heat transport above the Antarctic continent and a region of reduced up-

ward heat transport above New Zealand. For the S-polynya run we also tested the stationary eddy response and here we found also a significant increase in vertical heat transport with a maximum downstream of the S-polynya 40-60°S and regions of reduced vertical heat transport above the tropical Atlantic and the Indian Ocean Western Pacific region.

5 *Conclusions*

The comparison of the two model experiments shows that the model atmosphere reacted rather differently to the Arctic and the Antarctic polynya. The Arctic polynya run differed only marginally from its associated control run, whereby the additional heat flux from the polynya seems to be balanced by reduced meridional transports in the Atlantic strom track area. A change of the barotropic part of the flow is indicated by the fact that the projection of the 300 hPa streamfunction onto the eigenmodes of the barotropic vorticicy equation changed significantly between control and N-polynya run. The Antarctic polynya in the Weddell sea results in large scale large amplitude changes in many variables in the lower and in the upper troposphere. The heat fluxes from the polynya result in a cooling of the subtropical continents and anomalous low level highs and reduced pressure equatorward of 60°S in the upper troposphere. The poleward eddy heat flux in the lower troposphere even increased as a response to the Weddell polynya. The large response to the S-polynya indicates that sea ice variations might affect predictability. At this stage we however do not know why the responses to the Arctic and to the Antarctic polynyas differ. The reasons might be the size of the polynyas or their different position relativ to the mean flow. The Arctic polynya was situated in a very special region at the exit of the Atlantic storm track.

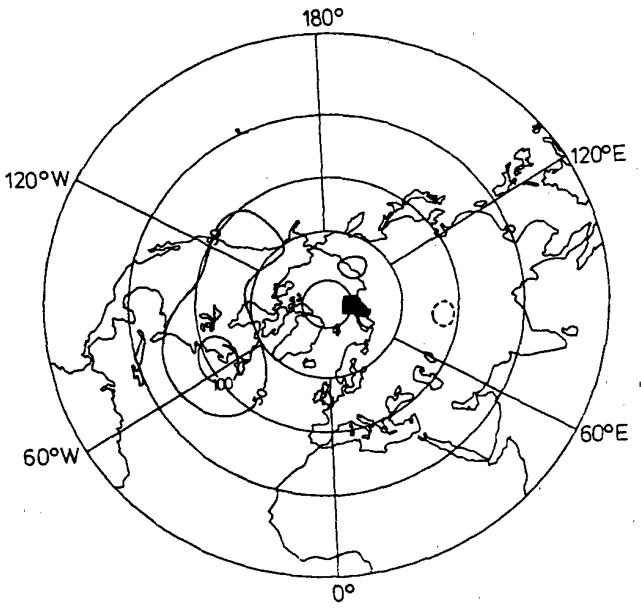


Fig. 1a: Significant difference of meridional heat fluxes N-polynya minus control at 700 hPa. The units are 10^{-2} Km/s.

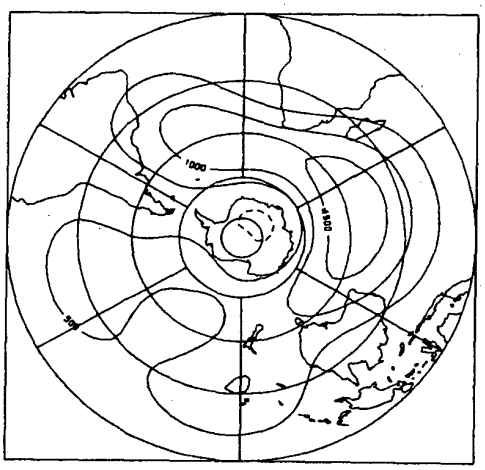


Fig. 1b: Significant difference of meridional heat fluxes control minus S-polynya at 850 hPa. The units are 10^{-2} Km/s.

A Grey Box Model for the Variations of Air Temperature

H.T. Søgaard and H. Madsen

The Institute of Mathematical Statistics and Operations Research
The Technical University of Denmark, building 321
DK-2800 Lyngby, Denmark
Tel: +45 42 881433; Fax: +45 42 881397
email: hm@imsor.dth.dk

1 Introduction

In this paper the formulation and estimation of a simple continuous time model for the variations of the surface layer air temperature is discussed. In the formulation of the model physical knowledge of the most important relations is used, and the model is formulated as a stochastic differential equation. The continuous time formulation permits a direct physical interpretation of the parameters. Models of this type, where prior knowledge is used in the model formulation, are called *grey box* models ([Graebe, 1990], [Bohlin, 1984], [Bohlin, 1991]). The grey box approach possesses attractive qualities compared to the following two traditional approaches:

1. *Purely deterministic and physical based models.* An example is found in [Zangvil, 1982]. The model structure is obtained by coupling well established models of involved subprocesses, and the associated parameters are determined from physical characteristics. The dynamics is typically specified as differential equations. An advantage of a physical based description is that a wide range of conditions can be studied through variation of physical interpretable parameters. However, since the models require a great number of parameters, boundary conditions and exogenous variables it may not be well suited for simulation and forecasting. Furthermore, insufficiencies of the descriptions of the subprocesses may add up, resulting in an overall model being too far from the true mechanisms.

2. *Purely statistical or black box models.* Black box models are built without incorporating any physical knowledge. The model formulation relies solely on statistical methods. Both general linear models and Box-Jenkins transfer function models belong to this class of models. As an example considering the variations of air temperature, Fourier coefficients are estimated in order to describe the annual and diurnal variations of the air temperature by means of harmonics (see, e.g., [Hansen and Driscoll, 1977]). The deviation from the harmonic functions is considered as noise. The correlation of the noise is typically described through ARIMA type models. The problem about over-parameterisation is not present, since it is possible, by statistical methods, to deduce whether a given parameterisation is reasonable, and if not, how it can be improved. However, it is frequently very difficult to give a physical interpretation of the parameters.

The grey box approach may be seen as a method in between the above traditional approaches. The structure of the grey box model proposed below is based on physical considerations involving the energy balance at the surface of the earth, cf. Section 2, and most of the parameters may be given a physical interpretation. The actual values of the parameters, however, are estimated using meteorological observations.

2 Energy Balance at the Surface of the Earth

The source of energy for the climate system is the incident solar radiation. At the surface of the earth, the sum of the incoming short-wave radiation (S_{in}) from the sun and the long-wave radiation (L_{in}) from the clouds and the atmosphere makes up the positive contribution to the net radiation, R_n. The reflected part (S_{out}) of the incoming short-wave radiation together with emitted long-wave radiation (L_{out}) from the ground make up the negative contribution to the net radiation. Consequently the net radiation obeys the relation

$$R_n = S_{in} + L_{in} - S_{out} - L_{out} .$$

Since no accumulation of energy takes place at the surface of the earth, the net radiation energy is balanced by a net heat flux away form the surface. The net heat flux is a sum of three different components: The latent heat flux (water vapour), $\mathcal{L}E$, the sensible heat flux, H, and the soil heat flux, G. The following equation states the energy balance:

$$R_n = \mathcal{L}E + H + G . \tag{1}$$

The sensible and the latent heat fluxes are called turbulent fluxes.

The sensible heat flux is strongly related to the vertical gradient of the air temperature, T_a. Correspondingly the latent heat flux is related to the vertical gradient of the humidity, q. The humidity is here expressed by the water vapour pressure. [Berkowicz and Prahm, 1982] introduce so-called resistance equations as approximations to these relationships. Combining the resistance equations with (1) they obtain the Penman-Monteith formula for the latent heat flux:

$$\mathcal{L}E = \frac{(R_n - G)\Delta r_a \gamma^{-1} + [q_s(T_a(z_t)) - q(z_t)]\rho c_p \gamma^{-1}}{r_s + (1 + \Delta\gamma^{-1})r_a} , \tag{2}$$

where ρ, c_p and γ are the air density, the specific heat of the air at constant pressure and the psychrometric constant, respectively. Furthermore r_a is an aerodynamic resistance for fluxes of heat and water vapour and r_s is a surface resistance for the flux of water vapour. $q_s(T_a(z_t))$ is the saturated vapour pressure at the temperature $T_a(z_t)$, while $q(z_t)$ is the actual vapour pressure at height z_t and $\Delta = \partial q_s(T_a)/\partial T_a$.

The expression for $\mathcal{L}E$ in (2) is non-linear in the parameters r_a, γ, ρ, c_p and r_s. However, the estimation method used later on requires an expression which separates the parameters and the input-output variables. Therefore a Taylor expansion of $f(\Delta) = 1/(r_s + (1 + \Delta\gamma^{-1})r_a)$ is introduced. If only the first two terms of

this series expansion are used ($f(\Delta) \approx f_0 + f_1\Delta$) the expression for $\mathcal{L}E$ can be rewritten as

$$\mathcal{L}E = \alpha_0 I_0 + \alpha_1 I_1 + \beta_0 J_0 + \beta_1 J_1 , \qquad (3)$$

where α_0, α_1, β_0 and β_1 are new parameters[1] entering linearly in the expression and

$$\begin{aligned} I_0 &= (R_n - G)\Delta \\ I_1 &= (R_n - G)\Delta^2 \\ J_0 &= q_s(T_a(z_t)) - q(z_t) \\ J_1 &= (q_s(T_a(z_t)) - q(z_t))\Delta . \end{aligned}$$

As the data set used for the parameter estimation later on contains hourly measurements of R_n, G, Δ, $q_s(T_a(z_t))$ and $q(z_t)$, hourly values of the variables I_0, I_1, J_0 and J_1 can be computed once and for all. Thus the expression in (3) is directly applicable as a part of the air temperature model described below.

3 Dynamic Relation between Air Temperature and Net Radiation

[Madsen *et al.*, 1987] argues that a model for the variations of air temperature should be dynamic and have the net radiation as an input variable. The arguments include a comparison of the mean diurnal and annual variations of the air temperature with the corresponding mean variations of net radiation (based on 69384 hourly observations of air temperature and net radiation). Both in case of the diurnal and the annual variations, the changes of the air temperature follow the changes of the net radiation with some time-delay. This indicates that the air temperature is driven by the net radiation through a dynamic system. [Madsen *et al.*, 1987] directs attention to the fact that a second order linear dynamic (possibly time-varying) system with net radiation as an input variable would explain a major part of the variation of the air temperature. The diurnal and annual phase shifts between the net radiation and the air temperature indicate that, as a first approximation, the system should contain two time constants of about 2 hours and 50 days for the diurnal and annual variation, respectively. Furthermore, [Madsen *et al.*, 1987] suggests that the smallest time constant is determined by the capacity of air and the resistances through which heat is exchanged with the surroundings. The largest time constant is due to the heat capacity of the sea. Their conclusions, however, indicate that besides the air and the sea, the heat dynamics of the soil should also be included in a model of air temperature. That is why the model presented in the next section is of order three.

4 A Linear Stochastic Model of the Air Temperature

Fig. 1 shows the model which is investigated in this paper. As an approximation it has been assumed that the local variations of air temperature is governed by a linear dynamic system with three heat capacities: The air over land, the surrounding sea water and the local land mass. Thus a system with three time constants is obtained.

Along with the model a number of quantities have been introduced. The variables are:

[1] Relations to the original parameters exist.

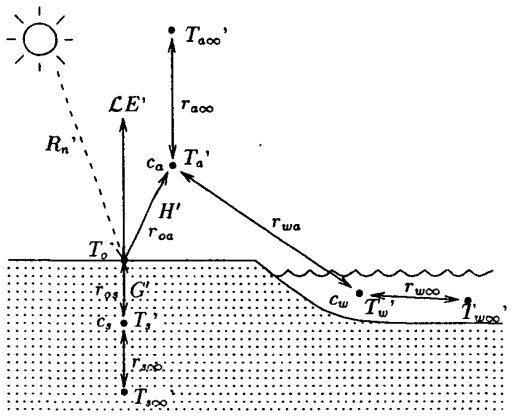


Figure 1: *A linear dynamic model of the air temperature based on a partitioning of the net radiation in the latent, the sensible and the soil heat fluxes.*

T_a' Air temperature 2 m above the surface of the earth [°C].
T_s' Temperature of the upper layer of the soil [°C].
T_w' Temperature of the sea [°C].
$T_{s\infty}'$ Constant temperature of some deeper layers of the soil [°C].
T_o' Temperature at the surface of the earth [°C].
$T_{a\infty}'$ Constant temperature, possibly in the upper part of the atmosphere [°C].
$T_{w\infty}'$ Constant temperature of some deeper ocean [°C].
R_n' Net radiation [W/m^2].
H' Sensible heat flux [W/m^2].
$\mathcal{L}E'$ Latent heat flux [W/m^2].
G' Soil heat flux [W/m^2].

The unknown model parameters to be estimated are:

c_a Heat capacity of the air [Wh/ °Cm^2].
c_s Heat capacity of the upper layer of the soil [Wh/ °Cm^2].
c_w Heat capacity of the sea [Wh/ °Cm^2].
r_{os} Resistance against the exchange of heat between the surface of the ground and the soil below [°Cm^2/W].
$r_{s\infty}$ Resistance against the exchange of heat between the upper layer and the deeper layers of the soil [°Cm^2/W].
r_{oa} Resistance against the exchange of heat between the surface of the earth and the air [°Cm^2/W].
$r_{a\infty}$ Resistance against the exchange of heat between the air over land and the constant temperature $T_{a\infty}'$ [°Cm^2/W].
r_{wa} Resistance against the exchange of heat between the sea and the air over land [°Cm^2/W].
$r_{w\infty}$ Resistance against the exchange of heat between the sea and some deeper ocean [°Cm^2/W].

As indicated in Fig. 1, the net radiation is partitioned into the latent, the sensible and the soil heat fluxes, agreeing with (1). The soil heat flux enters a first order dynamic system representing the soil. The temperature and the capacity of the upper layer of the soil are T_s' and c_s, respectively. An exchange of heat with the deeper layers of the soil occurs through the resistance $r_{s\infty}$. The amount (and direction) of heat transferred between the surface and the upper layer of the soil is determined by the temperature difference $T_o' - T_s'$. The latent heat flux is assumed to be governed by (3). Exchange of heat between the surface of the earth and the air is determined by the resistance r_{oa}. The rate of this exchange, i.e. the sensible heat flux, is determined by

the temperature difference $T_o' - T_a'$. The sensible heat flux enters a dynamic system consisting of two first order systems in parallel representing the air and the sea, respectively. Notice that a possible loss (gain) of energy to (from) both the atmosphere and the ocean takes place through the resistances $r_{a\infty}$ and $r_{w\infty}$.

Since the model is equivalent to an electrical RC-network (heat flux $\sim$ current; temperature $\sim$ potential) the formal model equations are easily obtained by using Kirchhoff's law for the nodes represented by T_o', T_s', T_a' and T_w':

$$R_n = \frac{T_o - T_s}{r_{os}} + \mathcal{L}E + \frac{T_o - T_a}{r_{oa}} \quad (4)$$

$$c_s \frac{dT_s}{dt} = \frac{T_o - T_s}{r_{os}} - \frac{T_s}{r_{s\infty}} \quad (5)$$

$$c_a \frac{dT_a}{dt} = \frac{T_o - T_a}{r_{oa}} + \frac{T_w - T_a}{r_{wa}} - \frac{T_a}{r_{a\infty}} \quad (6)$$

$$c_w \frac{dT_w}{dt} = \frac{T_a - T_w}{r_{wa}} - \frac{T_w}{r_{w\infty}} \quad (7)$$

All heat fluxes and temperatures are corrected by their long term averages (e.g., $T_a = T_a' - \bar{T}_a$).

In (4), $\mathcal{L}E$ should be substituted by the expression in (3). T_s, T_a and T_w are the states of the model described by (4)-(7). As measurements of the air temperature and the soil heat flux are present in the data set, both of these quantities are considered as output of the model. The air temperature is one of the states while the soil heat flux is expressed indirectly by the states as

$$G = \frac{T_o - T_s}{r_{os}} . \quad (8)$$

By solving (4) with respect to T_o and using the result to eliminate T_o in (5), (6) and (8), the model in its final state space form is readily obtained:

$$d\boldsymbol{T} = \boldsymbol{A}\boldsymbol{T}dt + \boldsymbol{B}\boldsymbol{U}dt + d\boldsymbol{w}(t) \quad (9)$$

$$\boldsymbol{Y}(t) = \boldsymbol{C}\boldsymbol{T}(t) + \boldsymbol{D}\boldsymbol{U}(t) + \boldsymbol{e}(t) , \quad (10)$$

where

$$\boldsymbol{T} = (T_s\ T_a\ T_w)^T \quad (11)$$

$$\boldsymbol{U} = (R_n\ I_0\ I_1\ J_0\ J_1)^T \qquad (\textit{Input vector}) \quad (12)$$

$$\boldsymbol{Y}(t) = (T_r(t)\ G_r(t))^T \qquad (\textit{Output vector}) \quad (13)$$

$$\boldsymbol{w}(t) = (w_1(t)\ w_2(t)\ w_3(t))^T \quad (14)$$

$$\boldsymbol{e}(t) = (e_1(t)\ e_2(t))^T \quad (15)$$

$$\boldsymbol{A} = \begin{pmatrix} -\frac{1}{c_s}\left(x_1 + \frac{1}{r_{s\infty}}\right) & \frac{x_1}{c_s} & 0 \\ \frac{x_1}{c_a} & -\frac{1}{c_a}\left(x_1 + \frac{1}{r_{wa}} + \frac{1}{r_{a\infty}}\right) & \frac{1}{c_a r_{wa}} \\ 0 & \frac{1}{c_w r_{wa}} & -\frac{1}{c_w}\left(\frac{1}{r_{wa}} + \frac{1}{r_{w\infty}}\right) \end{pmatrix} \quad (16)$$

$$\boldsymbol{B} = \begin{pmatrix} x_4 & -\alpha_0 x_4 & -\alpha_1 x_4 & -\beta_0 x_4 & -\beta_1 x_4 \\ x_5 & -\alpha_0 x_5 & -\alpha_1 x_5 & -\beta_0 x_5 & -\beta_1 x_5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (17)$$

$$\boldsymbol{C} = \begin{pmatrix} 0 & 1 & 0 \\ -x_1 & x_1 & 0 \end{pmatrix} \quad (18)$$

$$\boldsymbol{D} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ x_2 & -\alpha_0 x_2 & -\alpha_1 x_2 & -\beta_0 x_2 & -\beta_1 x_2 \end{pmatrix} . \quad (19)$$

with

$$x_1 = \frac{1}{r_{oa} + r_{os}} , \quad x_2 = r_{oa} x_1 , \quad x_3 = r_{os} x_1 ,$$

$$x_4 = \frac{x_2}{c_s} , \quad x_5 = \frac{x_3}{c_a} .$$

$T_r(t)$ and $G_r(t)$ are the hourly recorded values of air temperatures and soil heat fluxes, respectively (corrected by their long term average). In (9) $\boldsymbol{w}(t)$ is introduced in order to describe the deviation between the model and the true system. Likewise, $\boldsymbol{e}(t)$ is introduced in (10) to describe the measurement error. $\{\boldsymbol{e}(t)\}$ is assumed to be a sequence of independent $N(\boldsymbol{0}, \boldsymbol{R}_2)$ stochastic variables while $\{\boldsymbol{w}(t)\}$ is assumed to be a Wiener-process with incremental covariance $\boldsymbol{R}_1' dt$.

Table 1: *Estimation results.*

Parameter	Estimate		Standard deviation
c_s	104.6	Wh/ °Cm^2	1.7
c_a	26.08	Wh/ °Cm^2	0.39
c_w	2304	Wh/ °Cm^2	214
r_{os}	0.2093	°Cm^2/W	0.0010
$r_{s\infty}$	1.122	°Cm^2/W	0.050
r_{oa}	0.008171	°Cm^2/W	0.000237
$r_{a\infty}$	$+\infty$ (*)	°Cm^2/W	
r_{wa}	0.03234	°Cm^2/W	0.00043
$r_{w\infty}$	0.05067	°Cm^2/W	0.00309
α_0	0.01424	°C/Pa	0.00008
α_1	$-5.474 \cdot 10^{-5}$	°C^2/Pa^2	$5.3 \cdot 10^{-7}$
β_0	-0.3185	W/(m^{2}Pa)	0.0050
β_1	$7.546 \cdot 10^{-4}$	(W °C)/(m^{2}Pa^2)	$2.52 \cdot 10^{-5}$
τ_1	0.725	h	
τ_2	20.6	h	
τ_3	118	h	
Estimate of the one-step-ahead prediction error variance for air temperature: 0.3903^2 °C^2			

(*): The value of $r_{a\infty}$ has been fixed to ∞ due to statistical insignificance of $1/r_{a\infty}$.

5 Estimation of the Parameters

The parameters, c_s, c_a, c_w, r_{os}, $r_{s\infty}$, r_{oa}, $r_{a\infty}$, r_{wa}, $r_{w\infty}$, α_0, α_1, β_0, β_1, $\boldsymbol{R}_1'$ and $\boldsymbol{R}_2$, of the model are estimated by using a maximum likelihood (ML) method. The estimation is based on 69384 hourly observations[2] of the input and output variables. A detailed description of the ML estimation of the parameters in continuous time state space models can be found in, e.g., [Melgaard and Madsen, 1991].

Since it has not been possible to determine an explicit expression for the ML estimation, numerical methods have to be considered. In the present case a computer software package, CTLSM (Continuous Time Linear Stochastic Modelling) has been used ([Madsen and Melgaard, 1991]).

The obtained ML estimates are shown in Table 1.

After computing the eigenvalues, λ_i $(i = 1, 2, 3)$, of the estimated $\boldsymbol{A}$-matrix in the continuous time model (9), the time constants are obtained as

$$\tau_i = -\frac{1}{\lambda_i} , \quad i = 1, 2, 3 .$$

The time constants are shown in Table 1.

[2]Climate observations measured at a climate and water balance station, Højbakkegård, near Copenhagen, Denmark. The Observations cover the period from February 1st, 1966, to December 31st, 1973. See [Madsen, 1985] and [Hansen *et al.*, 1981].

6 Discussion and Conclusion

For the above model the variance of the prediction error for the air temperature is 0.3903^2 °C^2. Using the same climate data [Madsen *et al.*, 1987] estimated a more simple second order model with net radiation as input and air temperature as output. The prediction error variance for this model is 0.4686^2 °C^2. Hence, the model proposed in this paper improves the prediction performance considerably.

The time constant of 2 hours, as mentioned in Section 3, is presumably a mix of τ_1 and τ_2 since it lies in between them. This indicates that a second order model is more appropriate than a first order model for an approximation of the short term dynamics of the underlying distributed system governing the diurnal variations of the air temperature. The large time constant of about 50 days, however, does not have a counterpart among those in Table 1. One reason for that may be that estimation of a large time constant in a system with both large and small time constants is difficult. Therefore τ_3 should rather be attributed to some deeper soil layers than to the heat capacity of the sea.

The magnitudes of the estimated soil parameters seem to be reasonable. A typical value of the heat capacity of soil is $\rho c = 700$ Wh/m^3 °C. Therefore the estimate of c_s corresponds to an upper soil layer of thickness $\hat{c}_s/\rho c = 0.15$ m. Since the thermal conductivity of soil is typically about $\lambda_h = 1.3$ W/m °C, the estimate of r_{os} corresponds to a depth of $\hat{r}_{os}\lambda_h = 0.27$ m. The estimates of c_s and r_{os} are in reasonable accordance with figures in [Hansen *et al.*, 1981] which indicate that the diurnal variation in the soil is measurable down to a depth of 0.25 - 0.30 m. For the observed air temperatures the average heat capacity of atmospheric air is about 0.35 Wh/m^3 °C. This means that the estimated value of c_a on average represents the heat capacity of air up to a height of about 26/0.35 = 74 m. However, it is more likely that the estimated capacity also includes the upper part of the soil.

References

[Berkowicz and Prahm, 1982] Berkowicz, R. and L. P. Prahm, 1982: Sensible Heat Flux Estimated from Routine Meteorological Data by the Resistance Method. *J. Appl. Meteor.*, **21**, 1845-1864.

[Bohlin, 1984] Bohlin, T., 1984: Computer-Aided Grey-Box Validation. Tech.rep. TRITA-REG-8403, Dept. of Automatic Control, Royal Institute of Technology, Stockholm, Sweden.

[Bohlin, 1991] Bohlin, T., 1991: Grey-Box Identification: A Case Study. Tech.rep. TRITA/REG-91/00001, Dept. of Automatic Control, Royal Institute of Technology, Stockholm, Sweden.

[Graebe, 1990] Graebe, S., 1990: *Theory and Implementation of Gray Box Identification.* Ph.D. thesis. Dept. of Automatic Control, Royal Institute of Technology, Stockholm, Sweden.

[Hansen and Driscoll, 1977] Hansen, J. E. and D. M. Driscoll, 1977: A Mathematical Model for the Generation of Hourly Temperatures. J. Appl. Meteor., **16**, 935-948.

[Hansen *et al.*, 1981] Hansen, S., S. E. Jensen and H. C. Aslyng, 1981: *Jordbrugsmeteorologiske Observationer, Statistik Analyse og Vurdering, 1955 - 1979.* Hydrological Laboratorium, Royal Danish Veterinary and Agricultural University, Denmark.

[Madsen, 1985] Madsen, H., 1985: *Statistically Determined Dynamical Models for Climate Processes.* Ph.D. thesis. The Institute of Mathematical Statistics and Operations Research, The Technical University of Denmark, Lyngby, Denmark.

[Madsen and Melgaard, 1991] Madsen, H. and H. Melgaard, 1991: CTLSM - A Program for Parameter Estimation in Linear Stochastic State-Space Models in Continuous Time. Res.rep. 8/91, The Institute of Mathematical Statistics and Operations Research, The Technical University of Denmark, Lyngby, Denmark, 28 pp.

[Madsen *et al.*, 1987] Madsen, H. P. Thyregod and J. Holst, 1987: A Continuous Time Model for the Variations of Air Temperature. *Proc. 10th Conference on Probability and Statistics in Atmospheric Sciences*, Edmonton, Canada, 52-58.

[Melgaard and Madsen, 1991] Melgaard, H. and H. Madsen, 1991: The Mathematical and Numerical Methods used in CTLSM. Res.rep. 7/91, The Institute of Mathematical Statistics and Operations Research, The Technical University of Denmark, Lyngby, Denmark, 23 pp.

[Zangvil, 1982] Zangvil, A., 1982: A Study of the Effect of Variable Surface Conditions on the Outdoor Heat Stress Using a Mathematical Model of the Atmospheric Boundary Layer. *Energy and Building*, **4**, 285-288.

The Empirical Orthogonal Functions of 3-, 5- and 7-day Monsoon Rainfall over India

S.V. Singh and K.K. Singh
Indian Institute of Tropical Meteorology, Pune-411008, India

1. Introduction

Precipitation is the most important weather element from the point of view of economy particularly for the tropical countries. But it is highly unreliable and unpredictable as compared to other meteorological variables. This problem is more serious in monsoon regimes where rainfall is seasonal. It is known that large part of rainfall in the tropics occurs during large bursts of small spatial scales (i.e. meso scales) embedded and larger dynamical scale features making the rainfall field rather spotty or smaller time scales and therefore difficult to predict. Apart from the dynamical circulation features the distribution of rainfall is influenced to a large extent by the geographical features particularly the orography which interacts with the circulation regimes in rather intricate ways. Yet the rainfall exhibits some order in its spatial and temporal occurrence. It is important to recognise the orderly patterns.

The empirical orthogonal function analysis is the most commonly practical procedure for identification of large scale spatio-temporal patterns in rainfall and other climatic variables. The rotation of the optimal functions has often found to provide further insight in the structure of variability. These functions have been determined for the seasonal monsoon rainfall by several authors. The patterns for shorter periods, 3-, 5- and 7- day scale have been studied only to limited extent.

Kripalani, Singh and Arkin (1992) determined the EOF of 5-day rainfall by using 2.5°Lat./Long. grid data. However for studying the finer structure of rainfall variability, we have used 1° Lat./Long. grid data prepared by Prof. Hartmann of the University of Washington. This data utilises more than 3000 rainfall stations. There are about 10 stations in each 1°x1° bin and because of spatial

averaging of station data they are suitable for study of large scale structures on shorter time scales. We have determined EOF's of 3-, 5-, and 7-day rainfall for 4 seasons over India. The EOF's are also determined after applying several type of power transformations to data. They are rotated by varimax method and Harris-Kaiser method and a regionalisation of Indian rainfall on shorter time scales arrived.

2. Brief Description of unrotated EOF's

The eigenvalues and the variances explained by the first 4 EOF's obtained through the eigenanalysis of the correlation matrix are shown in the Table. Though the higher order EOF's (upto 10) show variance explained more than 1%, only four have been described here. It can be seen from the table that the variance explained by the first EOF is 3 times of the variance explained by the second EOF. After cube root or log transformation the variance explained by the first EOF doubles. The first EOF explaining more than half of the variance (after transformation) portrays well the large scale active/break phases of monsoon. In this EOF most of the country has same sign of loadings. Highest loadings are found along the east-west direction over central India along 22°N latitude with maximum values near 80° Longitude. High loadings are also located over upper west coast along 18°N latitude. This kind of pattern is expected as the large scale variation of rainfall is controlled by the east-west oriented monsoon trough at lower tropospheric levels. The second EOF shows one sign (negative) of loadings in northern India and opposite sign (positive) of loadings over most of central and southern India. Though the variance explained is one third of the first, the spatial structure shows large scale feature. Both the negative loadings over north India and positive loadings over south India continue to show east-west elongation. The zero line dividing the positive and negative loadings runs through the region where the first EOF shows maximum loadings This feature is obvious from the principles of orthogonality and maximisation of the variance

Table : Variance explained in percentage by first 4 EOFs.

Time period	EOF order	Untrans-formed	Cuberoot trans-formed	Log transformed
3-day	1.	27	51	49
	2.	9	8	8
	3.	8	6	6
	4.	5	4	4
5-day	1.	26	53	53
	2.	9	8	8
	3.	8	6	6
	4.	4	3	4
7-day	1.	28	54	55
	2.	9	9	9
	3.	9	6	6
	4.	4	3	3

followed. Similar features are shown by the third EOF. The first 3 EOF's have in general east-west orientation of loadings and spatial disposition of the loadings suggests a large scale wave in the north-south direction. In fact presence of such a wave related with the MJO, has already been pointed out by us through extended EOF and other analysis.

3. Rotation and Regionalisation

For obtaining simple structures of loadings we have rotated the first four interpretable EOF's by the varimax and Harris-Kaiser rotation. Both procedures give nearly identical patterns. The rotation creates more even distribution of variance explained by the functions and the loadings structure tends to be simpler. There is no major change in the

structure of the first EOF which is the most important function. The second and third EOF have now only one prominent centre of loading as against 2 or 3 centres in corresponding unrotated patterns. The second, third and fourth EOF in the reduced order of variance explained, show concentrated loadings in the central sub-Himalayan region, the dry northwest and the mid-Central peninsula respectively. These areas represent the homogeneous regions of variability.

4. Discussion

The EOF analysis is able to identify the most important modes of neutral variability of EOF's. The most important evolutionary features have also been identified in this and previous studies through extended EOF analysis of rainfall data filtered in various frequency bands. We have also determined the EOF's of 5-day rainfall over Thailand and SriLanka and OLR fields over Indian region. The most important coherent regions of convection and precipitation are identified. The rotation does not change the most important structures, but helps us to label the loadings according to their importance. The varimax and Harris Kaiser rotation have nearly same effect.

APPLICATION OF THE LANDSAT IMAGERY IN ATMOSPHERIC DIFFUSION

Y. Viswanadham, Lúcia H. R. Machado and J. A. Torsani

Instituto Nacional de Pesquisas Espaciais - INPE
Secretaria da Ciência e Tecnologia da Presidência da República
12201 - São José dos Campos, São Paulo, Brazil

1. Introduction

The meteorology and topography of urban coastal regions generally differ considerably from the features of inland urban areas. The transport and dispersion of pollutants over coastal terrain is usually more difficult to characterize than that over inland regions. In such areas, remotely sensed meteorological measurements are particularly useful where the application requires a spatial and temporal data density that cannot be obtained economically with a multitude of conventional point sensors, or where the region of the atmosphere to be explored is not accessible to normal instrument platforms. In situ and remote-sensing instruments are complementary rather than competitive in most of these cases. For example, Landsat detected numerous smoke plumes from industrial at the southern end of Lake Michigan drifting more than 150 km downwind (Lyons and Pease, 1973). Lyons (1975) summarized some satellite observations of both manmade and natural aerosols. While the satellite technique shows great promise, the Landsat images are made over any given region on the earth only 9 or 18 days in a narrow strip 180 km wide, and thus do not make it a likely candidate for an operational monitoring system. However, the authors used Landsat imagery to study in fact some meaningful interrelationships (e. g., Ernst, 1975; Lyons and Husar, 1976; Mekler et al., 1977; Viswanadham and Torsani, 1982 and others). The purpose of this work was to demonstrate the usefulness of Landsat imagery to present the applicability of empirical methods for determining the eddy diffusivity coefficients and the rate of kinetic energy dissipation from the diffusion parameters under complex meteorological conditions over long oceanic distances.

2. Theoretical Considerations

As an introduction to the working formulae, it is useful to consider certain convenient procedures. Consider idealized cases of instantaneous and continuous sources in a homogeneous air stream, with the x axis in the direction of the mean wind, supposed horizontal, the y axis across wind and the z axis vertical. The mean wind assumed to be constant in space, while the mean components of velocity in the crosswind and vertical directions are zero. Ideally it is assumed that the instantaneously generated cloud moves in a straight line parallel to the x axis, expanding in all directions, while the fixed continuous source generates a symmetric plume around the fixed x axis, with an expanding cross section in the y, z plane. The practical success of explicit diffusion formulae thus depends on the extend to which correct choices of the distribution shape, the lateral standard derivation S_y, and the vertical standard derivation S_z are contained therein. Comparison with the observations then enable the correctness of the original diffusion treatment to be assessed, and parameters such as diffusion coefficients and intensities of turbulence to be evaluated.

The major contribution to the theory of relative diffusion is that of Batchelor (1952). Viswanadham and Torsani (1982) have presented the theoretical procedure of Batchelor (1952) to calculate the apparent lateral eddy diffusivity coefficient (K_y) and the dissipation rate of turbulent energy ε from the Gaussian diffusion smoke plume model. Also, they illustrated to estimate the standard deviation of the crosswind distribution of material in the plume (S_y) from the Landsat imagery.

3. Sites And Data Sets

The Landsat imageries around 12 GMT (i. e., 0830 LST-Local Standard Time) were obtained from three areas:(i) Fábrica Nacional de Alcalis (National Chemical Industry), Cabo Frio (22° 59'S; 42° 02'W), Rio de Janeiro (RJ); (ii) Aracruz Celulose, São Mateus (18° 44'S; 39° 51'W), Espírito Santo (ES) and (iii) Riocell, Guaíba (30° 06'S; 51° 00'W), Rio grande do Sul (RS), Brazil. The necessary 0830 LST radiosonde data sets were collected from the meteorological stations at Galeão (22° 00'S; 42° 30'W), RJ; Ilha da Trindade (20° 31'S; 29° 19'W), ES, and Porto Alegre (30° 06'S; 51° 11'W), RS of the National Meteorological Department, Brazil. Determination of necessary auxiliary parameters from meteorological data and the examples of processed Landsat imagery are presented by Viswanadham and Torsani (1982) and Machado (1991).

4. Results

The foregoing classification of smoke plumes from Landsat imagery provides a useful

approach to the broad specification of diffusion in terms of general weather conditions. They may be used in assessing the quality of diffusion effects on particular occasions. However, apart from the qualitative nature of the results, there are numbers of limitations in the present approach due to meteorological data sets, topography and climatology of the sites. For the achievement of quantitative and reliable estimates of diffusion it is necessary to have accurate simultaneous turbulence measurements in space and time. To be exact, the characteristic variation of wind structure with height above ground should be taken into account. But the over-all effect of mesoscale motions would most likely be an increased diffusion over water. In Table 1, reference number of Landsat image cases, the gradient Richardson number (Ri) and the Brunt-Vaisala frequency (N) are presented for a qualitative description of the smoke plumes. A deep surface super-adiabatic layer failed to form, however, due to the relatively weak sunshine. The Richardson gradient numbers that were computed by taking temperature differences below the 600 m level of the radiosonde show that quasi-neutral and stable conditions predominated. The radiosonde date detected surface inversion below 600 m and also elevated ones beginning about 1200 m, extending to about 2000 m, in some cases (Table 1).

TABLE 1. Reference number of Landsat cases (n) and stability parameters.

n	Date	Local	Ri	$N^2(s^{-2})$ $(X10^{-4})$
1	23-06-76	RJ	0.084	2.36
2	11-03-77	RJ	0.011	0.94
3	22-05-77	RJ	0.035	1.84
4	31-10-77	RJ	0.007	0.39
5	01-07-78	RJ	0.064	1.43
6	15-07-77	RJ	0.035	1.84
7	08-04-84	RS	0.097	2.20
8	17-06-84	ES	0.001	0.02
9	12-07-87	ES	0.005	0.01
10	13-08-87	ES	0.002	0.04

The dispersion parameter S^2_y obtained from Landsat imagery is plotted in Fig. 1 as a function of downwind distance (x) from the source for the case 8. In Fig. 1, the straight line estimated by regression analysis and the statistical parameters are mentioned. There is a very little scatter of points about the line. A part of scatter may be attributed to the technique used in estimating S_y. The value of correlation coefficient r (= 0.9) indicates that the basic relationship between S^2_y and x is valid in each case. The goodness of fit of the regression line can be tested by means of the analysis of variance. As a test, the F value and the standard error of estimate $S_{y.x}$ (y = S^2_y and x = x) are calculated. The F value obtained is highly significant about a 99% significance level. It is interesting to note that the S^2_y values can indeed be used to calculate average lateral diffusion coefficient in quasi-neutral and stable conditions from Eq. (2) which is given in Viswanadham and Torsani (1982). The coefficient is then estimated to be 10 $m^2\ s^{-1}$ for quasi-neutral case with the average wind speed being taken as 2.1 m s^{-1} (Fig. 1).

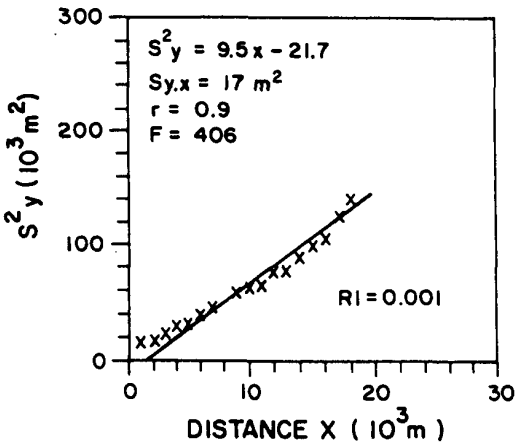


Fig. 1. Variance of the crosswind distribution of material in the plume (S^2_y) from the Landsat Imagery as a function of downwind distance (x) in quasi-neutral condiction at about 1130 GMT near Aracruz Celulose, São Mateus, Espírito Santo, Brazil.

In the Table 2, values of K_y estimated from Eq. (2) (see Viswanadham and Torsani, 1982) are more consistent and they do not show any abnormal and brusque changes with x. The calculations indicate that the values of K_y vary from 2 to 388 $m^2\ s^{-1}$ over the Atlantic Ocean (Table 2). These results compare well with the previous published values obtained over water surfaces (Pond and Bryan, 1976; Viswanadham and Torsani, 1982).

The rate of kinetic energy dissipation by viscous processes can be regarded as a good indicator of the strength of turbulence. With the aid of certain approximations the Landsat data sets can be used to make estimates of ε in these cases. Thus, we assumed an isotropic turbulence and inertial range spectrum to estimate ε from Eq. (3) (in Viswanadham and Torsani, 1982). The values we have obtained for ε are consistent with the values of Bunker (1977). The results are shown in Fig. 2 together with N. The best fit of the points to the form $\varepsilon = aN^y$ yields $a = 7.1 \times 10^{-2}$, y = 1 and $a = 2.1 \times 10^{-2}$, y = 0.5. In either case, the slope is distinctly positive and ε has a

TABLE 2 Estimated lateral standard derivation S_y (m) and lateral eddy diffusivity coefficient $K_y(m^2\ s^{-1})$ from Landsat images.

n	Ri	$\bar{u}$, m s^{-1}		Distance from source in km 1	4	8	12	16	20	Mean
1	0.084	4.4	S_y	72	109	200	245	327	454	234.50
			K_y	4.91	15.47	11.01	25.80	54.55		22.35
2	0.011	5.8	S_y	116	494	727	887	1032	1162	736.33
			K_y	222.89	206.26	187.22	201.73	206.78		204.98
3	0.035	4.1	S_y	80	261	742	1143	1203	1283	785.33
			K_y	42.18	247.25	387.39	72.14	101.93		170.18
4	0.007	4.1	S_y	79	173	388	551	632	734	426.17
			K_y	16.19	61.82	78.44	49.11	71.41		55.39
5	0.064	2.7	S_y	84	148	243	285	338	412	251.67
			K_y	6.68	12.54	7.48	11.14	18.73		11.32
6	0.035	4.1	S_y	90	129	168	199	258	-	168.80
			K_y	5.84	5.94	5.83	13.81	-		7.85
7	0.097	2.7	S_y	63	190	63	-	-	-	105.33
			K_y	14.49	-10.83	-	-	-		1.83
8	0.001	2.1	S_y	116	166	225	275	332	415	254.83
			K_y	4.94	6.06	6.56	9.08	16.28		8.58
9	0.005	3.0	S_y	126	153	174	221	285	338	216.17
			K_y	3.76	2.57	6.96	12.14	12.38		7.57
10	0.002	2.5	S_y	130	214	316	437	558	-	331.00
			K_y	12.04	16.89	28.47	37.62	-		23.75

$\bar{u}$ is the wind speed at 570 m. K_y is calculated from Eq. (2) in Viswanadham and Torsani (1982).

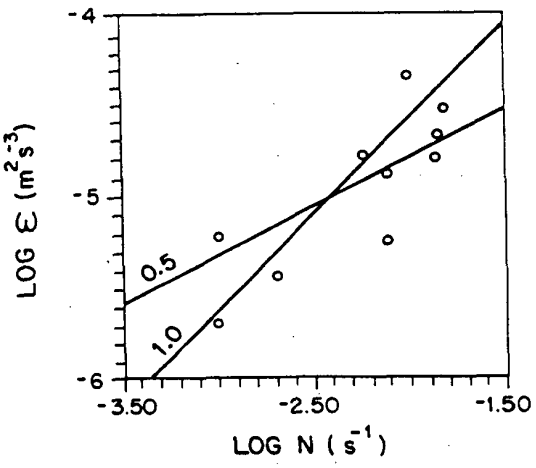


Fig. 2. Average turbulent kinetic energy dissipation rate ε obtained from the Landsat imagery as a function of the Brunt Vaisala frequency N.

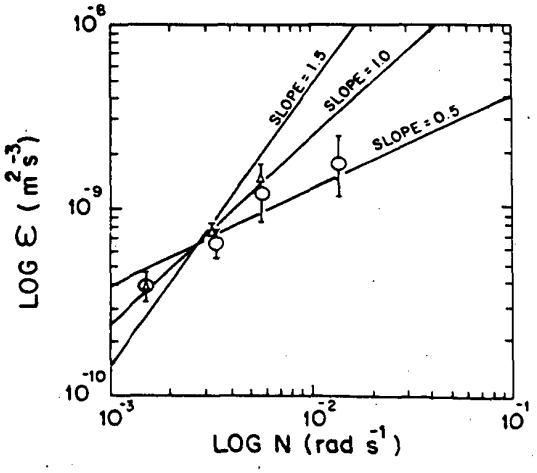


Fig. 3. Turbulent kinetic energy dissipation rate ε as a function of N in the western North Pacific Ocean along 152° E during the Wespac Experiment (after Moum and Osborn, 1986).

positive dependence on N.

Gargett and Holloway (1984) have presented a case for a positive dependence of ε on N based on the hypothesis that the turbulent energy is drawn from the internal wave field. They suggested a relation $\varepsilon \alpha N^{\gamma}$, where y has a value of 1 or 1.5, depending on the appropriate scaling for the vertical velocity variance in the internal wave field. Moun and Osborn (1986) have also presented similar results from the mixing in the main thermocline (Fig. 3). They suggested that if the turbulence is a direct result of energy transformed from the internal wave field, the positive dependence of ε and N results from a greather breaking rate of internal waves where N is large (i. e., in the main thermocline). Furthemore, a comparison of Figs. 2 and 3 implies that, while ε and N obey a similar form, the dissipation rate is not a single valued function of N. Unfortunately, the various aspects of the problem are so diverse that a good synthesis is not yet possible. Yet it was felt that in the long run, it would be beneficial to delve into some of the details surrounding the various case studies.

Acknowledgements. The authors are thankful to the authorities of INPE for their encouragement, Miss A. Y. Harada for data computation and preparing the manuscript, and Mr. C. R. Santos and his group for drafting the figures. They would like to acknowledge the assistance of Mr. J. C. Moreira in the Landsat imagery analysis. This work was partially supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brazil.

REFERENCES

Batchelor, G. K., 1952: Diffusion in a field of homogeneous turbulence, 2 , The relative motion of particles. Proc. Cambridge Philos. Soc., 48, 345-362.

Bunker, A. F., 1977: Structure, turbulence, fluxes and transformations of a maritime cold front during AMTEX. J. Meteor. Soc. Jpn., 55, 586-604.

Ernst, J. A., 1975: A different perspective reveals air pollution. Weatherwise, 28, 215-216

Gargett, A. E., and G. Holloway, 1984: Dissipation and diffusion by internal wave breaking. J. Mar. Res., 42, 15-27.

Lyons, W. A., and S. R. Pease, 1973: Detection of particular air pollution plumes from major point source using ERTS-I imagery. Bull. Am. Meteor. Soc., 54, 1163-1170.

Lyons, W. A., 1975: Satellite detection of air pollutants. In: Remote Sensing Energy Related Studies (ed., by T. N. Vaziroglu), John Wiley, New York, 263-290.

Lyons, W. A., and R. B. Husar, 1976: SMS/GOES visible images detect a synoptic-scale air pollution episode. Mon. Wea. Rev., 103, 1623-1626.

Machado, L. H. R., 1991: Estudos da camada limite atmosférica por imagens do satélite Landsat. Tese de mestrado em Meteorologia, INPE, São José dos Campos, Brazil, (INPE-5351-TDI/460), 125 pp.

Mekler, Y., H.Quenzel, G. Ohring and I. Marcels, 1977: Relative atmospheric aerosol content from ERTS observations. J. Geophys. Res., 82, 967-970.

Moum, J. N., and T. R. Osborn, 1986: Mixing in the main thermocline. J. Phys. Ocea., 16, 1250-1259.

Pond, S., and K. Bryan, 1976: Numerical models of the ocean circulations. Rev. Geophys. Space Phys., 14, 243-263.

Viswanadham, Y., and J. A. Torsani, 1982: A study of atmospheric diffusion from the Landsat imagery. J. Geophys. Res., 87, No. C12, 9621-9635.

The Influence of Change of Soil Temperature Field on the Precipitation in Summer

Ma Zhuguo Tang Maocang

(Lanzhou Institute of Plateau Atmospheric Physics, Chinese Academy of Science, Lanzhou, Gansu, P. R. China)

Abstract

We have discovered that there is a good corresponding relationship between the axes of soil temperature anomalies at the depth of 1. 6m in winter(Dec. to Feb.) and precipitations in the following flood season (Tang etc. ,1982), the axes of maxima and minima of soil temperature anomalies usually lay close to the axes of precipitation maxima and minima. These similarities were used in the long—range prediction of precipitation issued in March of each year. Practical forecasts have been done since 1975 (Xin, 1985, Tang etc. ,1989). The accuracy of the method is higher than that of other method.

Recently, we have also discovered that the changes of soil temperature field significantly influence the precipitation in the following summer. The results of which will be intruduced in the following.

By using the soil temperature data observed in 180 stations in China and the departure distributions of seasonal mean soil temperature at the depth of 3. 2m from 1980—1990, we have analysed slow variation (the sign of the departure of soil temperature is keeps for more than three months), the fast variation (the sign of the departure of soil temperature changes in three months) of soil temperature at the depth of 3. 2m, and the relationships between either of them and the precipitation in the following summer. The results show that when the durations of the departure of soil temperature at the depth of 3. 2m and 1. 6m are longer than six months, there will be a significant influence on precipitation in the following summer, and when the durations are shorter than three months, there will be little influence on precipitation in the following summer. Meanwhile, we have also analysed the changes of the departure of soil temperature during 1990—1991. The results indicate that the long—duration of high soil temperature in the area between Yangtse river and Huaihe river before summer is an important reason of the floods at the area in summer of 1991. The mean duration of the positive soil temperature departure was about 12 months at the depth of 1. 6m and about 14 months at the depth of 3. 2m in this area. Finally, an improved idea of predicting precipitation in summer has been suggested.

Some Aspects of Boundary Layer Processes in Monitoring Climate Change in the Indian Monsoon Region.

Malti Goel
Department of Science and Technology,
New Delhi-110 016, INDIA

Abstract

A typical event of heavy rainfall in the semi-arid area of Jodhpur occured on July 3-5, 1990 during the Southwest (SW) monsoon season. The observations of the synoptic system and the associated changes in the boundary layer processes have been presented. The rapid horizontal advection of moist air mass gives rise to forced convection, vertical advection of moisture, condensation and rainfall.

1. INTRODUCTION

Fig. 1. shows the normal position of Indian Monsoon Trough (MT) which extends from Northwestern India to Western Bengal. The position of trough line varies day to day. No other semipermanent features of the Indian summer monsoon has such control on Southwest (SW) monsoon activity than the MT. The Western end of the trough extends over the desert region of West Rajsthan, which acts as a heat sink. A decrease in the surface albedo near this region (Charney et al, 1977) is expected to intensify MT. However, the stastistical analysis of monsoonal rainfall for 1871 to 1988 does not indicate any significant trend in long term changes (1992).

Boundary layer studies are essential to understand the climate response of the desert. Energy mass exchanges in this region give rise to nondivergent flows and form an important link in the evolution of monsoon (1982). However, the role of planetary boundary layer in maintaining the MT and consequent changes in the monsoon rainfall have not been adequately modelled because of lack of data.

The motivation of this study is to understand the role of boundary layer associated with a typical event of heavy rainfall in MT region.

2. THE CASE STUDY

The case study chosen is that of Jodhpur (26 N, 73 E) at the western periphery of MT, when an active synoptic system in the first week of July' 90, with a copius rainfall passed over it during MONTBLEX experiment (Goel & Srivastava, 1990).

For the first time a detail study of land-atmospheric interactions in MT zone during life cycle of SW monsoon was made possible through MONTBLEX' 90 platforms. Coverage from the micrometeorological towers, acoustic sounding, India Meteorological Department's radiosonde network and satellite enabled this event to be observed in greater detail.

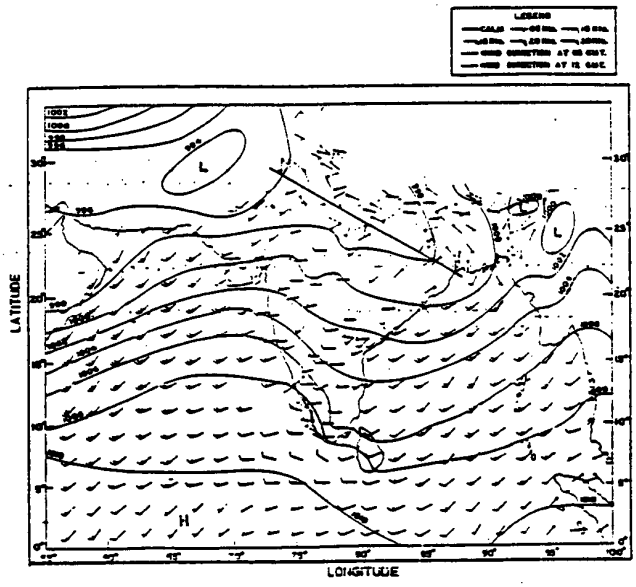


Fig. 1. Mean position of Monsoon Trough during SW monsoon season

3. SYNOPTIC SYSTEM

1990 was a good monsoon year with an excellent spatial and temperal distribution of rainfall over whole of Indian continent. Monsoon depressions developed in the Bay of Bengal moved West-north-westwards in the MT region and the entire country was covered by the monsoon current by July' 90. It was seen that a low pressure area formed southwest of Uttar Pradesh moved over to adjoining West Rajsthan and neighbourhood areas by July 2nd, 1990. Dense organised clouds developed in the night with an abrupt fall in sea level pressure over Jodhpur, which recorded a heavy rainfall of about 68 cm in the following week.

4. DATA ANALYSED

Data on temperature, humidity and wind measurements carried out on a 30 m micrometeorological tower having sensors at 6 levels have been examined. The surface and radiosonde meteorology data of India Meteorological Department enabled tracking of the synoptic system. Fortunately, observations of thermal echo soundings from monostatic sodar installed in the same compound have also been correlated.

5. DISCUSSIONS AND CONCLUSIONS

This region of West Rajsthan is semi-arid with a low humidity at all levels, high cloud base and a few showery clouds, from which downdrafts descend all the way to the surface, giving rise to gusty, dustladen winds. The annual mean rainfall for Jodhpur is approximately 36 cm. The MT activity is expected to bring 80% of it, during the entire SW monsoon season.

The general picture of diurnal temperature distribution in the summer is shows in Fig. 2. The intense surface heating builts up a unstable boundary layer where buoyancy forces arise mainly from sensible heat flux. The surface fluxes are considered to respond much more quickly to the external influences.

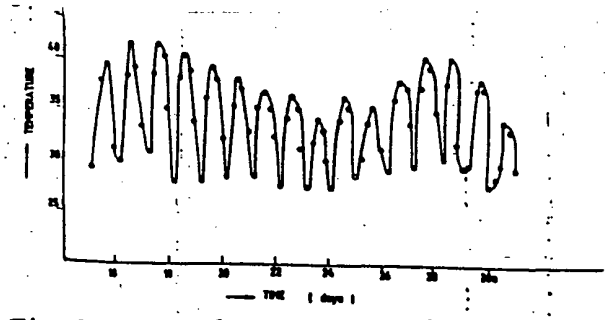


Fig.2. Diurnal variation of temperature in Jodhpur during summer season

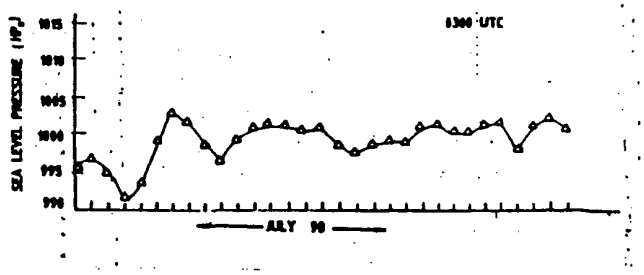


Fig.3 Sea level pressure at Jodhpur during July ' 90.

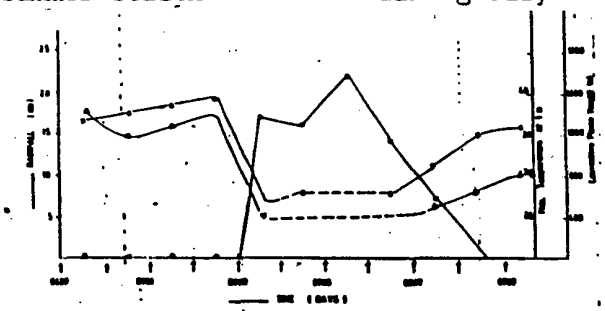


Fig.4. Time variation on daily rainfall (--o--), Maximum daily temperature (--Δ--) and Convective plume heights (--x--) as measured by sodar.

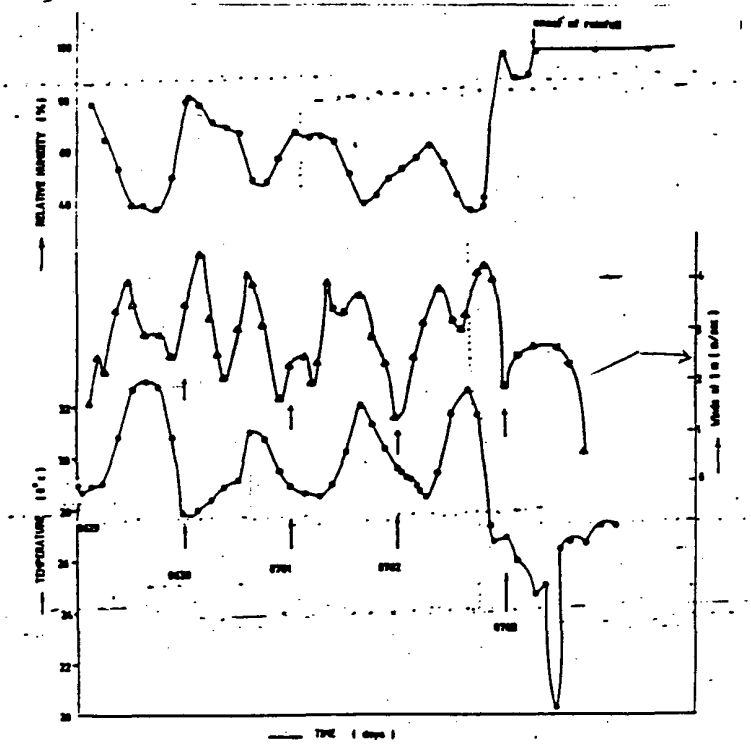


Fig.5. Variation of temperature, relative humidity and wind speeds at 1 m height.

In the case study choosen, the arrival of low pressure wave at Jodhpur (Fig. 3), enabled the moist air mass from south Uttar Pradesh to reach in depth the western periphery of MT, giving rise to large rainfall (Fig. 4). Surface heating is then not an immediate factor and the situation is far from mean picture. Fig. 5 shows the temperature, relative humidity and winds speeds at 1 m height.The horizontal advection of surface moisture approaching from North-northeast direction, as seen from the abrupt change in wind direction on the night of 2nd July, leads to sudden release of energy and cooling of the layer. The sensible heat flux becomes much reduced and the wind speeds increase. Buoyant convection is replaced by forced convection in the boundary layer. The presence of strong, positive vertical wind shear in the environment of disturbances of small scale storms results in the generation of horizontal vorticity, much greater than vertical vorticity. In severe local disturbances which produce flash floods and enormous rainfall, this horizontal vorticity is tilted up, to produce vertical vorticity whose time rate of change is associated with convergence of vertical motion, vertical advection of moisture, condensation and rainfall.

Rajsthan has witnessed similar heavy rainfalls in the earlier years too, some of these being heaviest in the country. Prediction of this phenomena, and its exploitation in irrigation has potential of changing climatology for this region.

Acknoledgements: The Author is much indebted to Prof. Y. Ramanathan for valuable suggestions and discussions. She thanks Prof. A Prabhu, Indian Institute of Science, Dr. S.P. Singal, National Physical Laboratory, Dr. U.C. Mohanty Indian Institute of Technology and India Meteorological Department for observations and processing of data during MONTBLEX Experiment.

She is grateful to Dr. P. Rama Rao, Secretary DST and Dr. Harsh K. Gupta, Adviser for the encouragement.

REFERENCES

1. Charney, I.G., Quirk W.J., Shu-Hsien Chow and Kornfield, J., 1977: A Comparative Study of the Effects of Albedo in semi-arid Regions, J. Atmos. Sci., 34, 1366-85.

2. Goel M. and Srivastava, H.N. 1990: Monsoon Trough Boundary Layer Experiment (MONTBLEX), Bull. Amer. Met. Soc., 71, 1594-1600.

3. Krishnamurthy T.N. and Ramanathan Y., 1982: Sensitivity of the Monsoon Onset to Differential Heating, J. Atmos. Sci., 39, 1290-1306.

4. Srivastava, H.N., Dewan S.N., Dikshit, S.K. Prakash Rao, G.S. Singh, S.S. and Rao, K.R., 1992: Decadal Trends in Climate over India, Mausam, 43, 7-20.

A STATISTIC-DYNAMICAL MODEL FOR CLIMATE FORECASTING

Huang Jianping
Department of Geophysics
Peking University, Beijing 100871, PRC

ABSTRACT

Considering from point of view of the dynamics, it is convenient to regard the field to be predicted as a small disturbance superposed on the historical analogous field, and thus the statistical technique can be used in combining with the dynamics. Along this line, a coupled atmosphere-earth surface analogy-dynamical model is formulated and applied to making seasonal prediction.

The forecast are obtained by using analogy-dynamical model to predict deviation from a seasonally-varying basic state. The basic state is obtained as follows: for a forecast initialised monthly anomaly field, the best historical analogy to the month is selected from all the same month in the historical record. The subsequent evolution of the observed general circulation in the year selected then forms the basic state for the forecast.

This approach facilitated the utility of the useful information contained both in the historical data set and in the initial field to improve the dynamic model based solo on the latter and show better skill in prediction.

STATISTICAL EXAMINATION OF CLIMATOLOGICAL DATA RELEVANT TO GLOBAL TEMPERATURE VARIATION

H.L. Gray, Richard F. Gunst, and Wayne A. Woodward
Department of Statistical Science
Southern Methodist University

Debate continues to rage within the scientific community regarding the possibility of a global "warming trend" due to the build-up of CO_2 and other greenhouse gases. As part of a multi-faceted state-of-the-art statistical analysis of relevant data, research has been conducted on the use of statistical techniques to model climatological data for purposes of assessing whether a warming trend exists, identifying and understanding the explanatory variables, and establishing "best" forecast functions where appropriate. Concurrent with these efforts is a critical examination of the quality of existing data sets upon which conclusions are being drawn. Fundamental issues related to defining and estimating mean global temperature change are also being investigated. This research is of importance because defensible conclusions concerning the predictability of global temperature must be based on valid statistical estimates of global temperature change.

1. *Testing for Trend in the Global Temperature Data*

In recent years a number of statistical tests have been proposed for testing the hypothesis that global warming is occurring. The standard approach is to examine one or two of the more prominent global temperature data sets by letting

$$Y_t = a + bt + E_t,$$

where Y_t represents the temperature at time t and E_t represents error from the trend line, and testing the hypothesis that $b = 0$. Several authors (e.g. Bloomfield and Nychka, 1991; Bloomfield, 1991; Kuo, Lindberg, and Thomson, 1990) have applied these tests for trend to determine whether a significant trend exists in the widely referenced global temperature data sets, and have generally concluded that there is a significant trend in the data of the form which should be considered in forecasts. Indeed, if such tests are applied to the Hansen and Lebedeff (1987,1988) temperature data and the IPCC data (Folland, Karl, and Vinnikov, 1990) the hypothesis that $b = 0$ is rejected. The question which then arises is, *"Should we consider this to be strong evidence in support of the position that there is a warming trend?"* If the question is, *"Is there a trend in the temperature data over the time interval for which the data were observed?"* then these tests certainly strongly support the positive response. However, if the question is, "(*Assuming that the future is driven by the same mechanisms driving the past*), *do the data support the proposition that this trend will continue into the future?*" then the response should be much more guarded.

In addition to a linear trend model, certain autoregressive-moving average (ARMA) models are also very reasonable models for these data due to the random trends present in their realizations. An ARMA model is a model of the form

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(Y_t - \mu) = (1 - \theta_1 B - \cdots - \theta_q B^q)a_t$$

where a_t is white (uncorrelated) noise, the ϕ_j's and θ_j's are real constants, and $B^k Y_t = Y_{t-k}$. The values p and q are called the model orders. The model specifies that

the current value of the process at time t, i.e Y_t, is a linear combination of the previous p values of the process plus a linear combination of noise components. It is shown that the correlation structure for these models is not nearly strong enough for the ARMA-based forecasts to predict any continued increase. However, when data were generated from these ARMA models, the tests for trend quite often incorrectly predict a trend to continue, i.e. the short term random trends in these realizations cause the tests to erroneously conclude $b \neq 0$ (and thus treat them as if they were long-term trends). These statistical tests for testing H_0: $b = 0$ have little or no ability to distinguish between realizations from ARMA models with a high correlation between successive values and those from models of the form $Y_t = a + bt + E_t$. Based solely on the available temperature deviation series, it is difficult to conclude that these temperature data suggest that the trend will continue over any extended length of time. Further, trend tests based on the model $Y_t = \mu_t + E_t$ for the purpose of prediction or inference concerning future behavior should be used with caution. For example, application of the trend tests to the yearly sunspot averages over the 108 years spanned by the Hansen and Lebedeff data results in a significant upward trend, but it is obvious that no one would consider incorporating this trend into forecasts of future sunspot behavior. Rather the observed trend is interpreted as simply a part of the natural cycle of increasing and decreasing amplitudes in these data. A more complete discussion of these findings is given in Woodward and Gray (1991).

2. *Defining and Estimating Mean Global Temperature Change*

Mean global temperature change is a concept whose definition usually must be inferred by how it is estimated. Explicitly lacking in all the major articles on global warming is any definition of a fixed (theoretical) quantity referred to as "mean global temperature change." Without such a definition, the assessment of estimation methodologies relies on the "reasonableness" of the respective approaches, an assessment that is very subjective. On the other hand, a definition of a theoretical quantity such as mean global temperature change permits a rigorous evaluation of estimation strategies, evaluations that can be either theoretical or empirical.

At any fixed time t, each location x on the earth has an associated temperature $T_t(x)$. Current estimation procedures focus on temperature deviations or "anomalies" $a_t(x)$, where $a_t(x) = T_t(x) - \tau(x)$ is the temperature deviation at time t for location x, and $\tau(x)$ represents the temperature mean for location x over a stated period of time. The definition of mean temperature change over a region R used in this research is

$$\delta_t = \int_R a_t(x)\, w(x)\, dx \;/\; \int_R w(x)\, dx \;,$$

where the region R can represent any geographic area of interest, including the entire earth. If the region R is defined by a discrete set of locations $\{x_i\}$, the integrals can be approximated by summations :

$$\delta_t = \sum_i a_t(x_i)\, w(x_i)\, A_i \;/\; \sum_i w(x_i)\, A_i \;,$$

where the subscripts i denote a partitioning of the region R, and A_i is the area of the ith subregion in the partitioning.

Based on the above integral definition of a mean global temperature anomaly (i.e. R is the entire earth), commonly used distance weighting of anomalies, i.e. linear distance weighting used by Hansen and Lebedeff (1987) and inverse distance weighting used by Jones, et al. (1986a, 1986b, 1986c) and Jones (1988), are shown to have inferior statistical properties to ordinary (uniform weighting) and robust (m-estimator weighting) averages for the estimation of regional means. On the other hand, gridding (typically applied by researchers in this area without any justification in the specific context of estimating regional temperature anomaly means) is shown to improve the performance of all the averages studied when station locations are

not evenly dispersed throughout a region, the usual situation in practice. Data reuse in the context of gridding, such as that employed by Hansen and Lebedeff (1987), is shown to often lead to more serious biases than those resulting from the choice of a less than optimal estimator.

As an application of the above findings, mean global temperature anomalies are estimated for the years 1880-1988 based on the NCAR monthly station data and using m-estimator weighting, gridding, and no station reuse. Results obtained with uniform weighting and m-estimator weighting were very similar. Our results suggest that temperature anomalies in the late 1800's might not be as small as heretofore reported. This difference is not so great that previous conclusions made from the temperature records are negated; however, they clearly demonstrate the need for the research that is being conducted. Moreover, these findings leave open the question of whether further investigations will suggest additional modifications of the estimates.

3. *Analysis of Data Using ARMA and GARMA Models*

For data such as the mean global temperature deviations, CO_2 data, or other data related to climate, the correlation structure between observations may persist over long intervals of time. In this case, it is often true that neither a signal + noise model nor an ARMA model are appropriate. A possible solution to this dilemma is the Gegenbauer ARMA (GARMA) model which is a long memory model introduced by Gray, Zhang, and Woodward (1989). This model is an extension of the ARMA model which allows for long term dependency in stationary models, i.e. allows for stationary models with slowly damping autocorrelations. The GARMA model is given by

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(1 - 2uB + B^2)^{\lambda} \times (Y_t - \mu) = (1 - \theta_1 B - \cdots - \theta_q B^q) a_t$$

where u specifies the frequency at which the long memory behavior occurs and λ essentially indicates how slowly the autocorrelations damp. The GARMA model appears to be particularly appropriate for data related to climate change. Current investigations are examining the validity of this observation through the fitting of a number of different climate related data sets. Initial investigations indicate that application of GARMA models to the series of yearly sunspot averages results in models which produce better forecasts than those obtained with ARMA models, resulting in forecast error reduction in the area of 50%.

4. *Kriging Approach to Estimating Mean Global Temperature*

Kriging is a method for analyzing spatial data, and in our case has application to the estimation of mean global temperature from station data. This technique depends on a covariance structure which is essentially a spatial counterpart to the autocovariance structure important in the analysis of time series. Critical components in the fitting of kriging models are the assumptions about the form of the spatial covariance matrix and the procedures used to fit the assumed matrix.

In the published literature on mean global temperature estimation, there is little structured use made of the information provided by the locations of temperature stations. Informally, neighboring stations are used as quality control checks on temperature trends. For example, an apparent discrepant temperature measurement from one station may be corrected if neighboring stations do not have a similar value. In most of the published literature (e.g., Hansen and Lebedeff, 1987; Jones et al., 1986a,b,c, 1988) the proximity of stations is used in the calculation of local averages through the use of distance weighting from the center of a region or a grid point. No formal statistical model fitting in which the use of the actual station locations forms an integral part of the statistical model has yet been reported.

Kriging is a spatial modeling technique in which station location information is used through the correlation structure assumed for the measurements. The covariance matrix or the variogram for a set of measurements is modeled as a function of station location. This fitted covariance matrix or variogram is then used in a generalized least squares estimator of a regression model for temperature. The regression model can include terms for trend as a function of location as well as other concomitant measurements. Kriging models have been used predominantly in the mining and geosciences literature. They have significant potential for use in modeling global temperature variation. Of particular importance in this regard is the ability to calculate standard errors of the predicted temperatures.

Asymptotic properties for kriging estimators using increasing domain asymptotics are being investigated. Mardia and Marshall (1984) establish the consistency and asymptotic normality of a general class of models which includes ones under consideration for use in fitting station temperature data. Their results may be able to be applied to the mean global temperature estimation methods proposed by Gunst et al. (1991), which use unweighted or m-estimator weighted temperature averages for a lattice of grid locations. Of particular interest and complexity is the generalization of the theory to competing estimators that have been proposed in lieu of maximum likelihood.

REFERENCES

Bloomfield, Peter (1991), "Trends in Global Temperature," *Climatic Change* (in press).

Bloomfield, P. and Nychka, D. (1991). "Climate Spectra and Detecting Climate Change," *Climatic Change* (in press).

Folland, C.K., Karl, T.R., and Vinnikov, K.Y. (1990), "Observed Climatic Variations and Change." In *Climate Change: the IPPC Scientific Assessment,* ed. Houghton, J.T., Jenkins, G.J., and Ephraums, J.J., 195-238. Cambridge: Cambridge University Press.

Gray, H.L., Zhang, N. and Woodward, W.A. (1989). "On Generalized Fractional Processes," *Journal of Time Series Analysis* 10, 233-257.

Gunst, R. F., Basu, S., and Brunell, R. (1991), "Defining and Estimating Mean Global Temperature Change," Technical Report, Center for the Study of the Evolution of the Environment, Southern Methodist University, Dallas, Texas.

Hansen, J. and Lebedeff, S. (1987), "Global Trends of Measured Surface Air Temperature," *Journal of Geophysical Research* 92, 13345-13372.

Hansen, J. and Lebedeff, S. (1988), "Global Surface Air Temperatures: Update Through 1987," *Geophysical Research Letters,* 15, 323-326.

Jones, P.D., S.C.B. Raper, R.S. Bradley, H.F. Diaz, P.M. Kelly, and T.M.L. Wigley, (1986a), "Northern Hemisphere Surface Air Temperature Variations: 1851-1984," *J. Clim. Appl. Meterol.,* 25, 161-179.

Jones, P.D., S.C.B. Raper, and T.M.L. Wigley (1986b), "Southern Hemisphere Surface Air Temperature Variations: 1851-1984.," *J. Clim. Appl. Meterol.,* 25, 1213-1230.

Jones, P.D., T.M.L. Wigley, and P.B. Wright, (1986c), "Global Temperature Variations Between 1861 and 1984," *Nature* 322, 430-434.

Jones, P.D., (1988), "Hemispheric Surface Air Temperature Variations: Recent Trends and an Update to 1987," *Journal of Climate* 1, 654-660.

Kuo, C., Lindberg, C. and Thomson, D.J. (1990), "Coherence Established Between Atmospheric Carbon Dioxide and Global Temperature," *Nature* 343, 709-714.

Mardia, K.V. and Marshall, R.J. (1984). "Maximum Likelihood Estimation of Models for Residual Covariance in Spatial Regression," *Biometrika* 71, 135-146.

Woodward, W. A. and Gray, H. L. (1991), "Global Warming and the Problem of Testing for Trend in Time Series Data," Technical Report, Center for the Study of the Evolution of the Environment, Southern Methodist University, Dallas, Texas.

AN ANALYSIS OF REGIONAL CLIMATE DATA

Knut Solna and Paul Switzer

Statistics Department
Stanford University
Stanford, CA 94305 USA

1. INTRODUCTION

Time series of global surface temperatures have been statistically analyzed with a view toward detection of a warming trend. It has been usual to assume that the globally averaged temperature data are error-free, for practical purposes. Whatever errors there may be resulting from spatial interpolation of the monitoring data have generally been assumed to have a small effect on the uncertainty of estimates of temporal change.

Justification for the assumption of a practically error-free global time series derives from one or another of the following two arguments. In one argument, the spatial correlation length of monitoring stations is estimated to be about 1200 km's; circular regions of radius 1200 km's are then shown to cover most of the land surface of the globe. In the second argument, output from General Circulation Models [GCMs] which can simulate the spatial-temporal surface temperature field of the globe is sampled at the actual monitoring network; estimates of the global average temperature derived from the sample data are then compared with "true" global averages.

The foregoing approaches to assessing the importance of the incomplete nature of the data may well have reached a reasonable conclusion. However, the arguments are not founded on a systematic statistical characterization of the observed variability. Essentially, our goal is to replace arguments based on the foregoing crude "parameterizations" by a coherent model of variability from which errors due to interpolation and aggregation can be mathematically described.

Errors in the global time series derive from several sources. There are instrumental errors which can be both systematic and random. The random instrumental errors are likely to be small in relation to other sources of variability discussed here. The systematic instrumental errors arise from calibration problems and placement. Systematic errors induced by placement can include errors associated with changes in the monitoring network over time, or errors associated with geographically localized evolutionary changes such as increasing heat island effects. Corrections made for heat island effects appear to indicate that this source of systematic error is relatively small. However, the systematic error induced by the evolution of the monitoring network warrants more careful attention, in our view.

In addition to systematic errors such as those described above, there are errors which may properly be regarded as random. For example, the discrepancy between the "true" globally and annually averaged surface temperature and the corresponding estimates derived from the available data is sure to be different in different years. This random error derives from the spatial interpolation of monitoring data to unmonitored locations.

One approach which addresses both the systematic error introduced by the evolution of the monitoring network as well as the random error introduced through the spatial interpolation of the monitoring data relies on building a statistical model of spatial-temporal variability. This is the approach taken by us. In the analysis of global time trends one can sometimes justify incorporation of the interpolation random error into the statistical modeling of temporal variability, thus avoiding the need for spatial modeling of error. However, the evolution of the monitoring network still calls for spatial modeling in order to assess and explain the role of changing uncertainty in global temperature estimates.

Of course, a-priori physical information may be available which can narrow the burden of the statistical modeling effort. The dilemma is that the physical information would be most influential where the empirical data are most sparse, while at the same time the physical information can be less trustworthy where the empirical data are sparse.

The salient feature of the error introduced through spatial interpolation is that it is patently heterogeneous on a global scale. Error heterogeneity derives both from the heterogeneity of variability of the underlying temperature field and from the varying density and geographic configuration of the monitoring network. Furthermore, the spatial interpolation errors may be seasonally heterogeneous.

What we shall aim for is a statistical description of the spatial variability of the underlying surface temperature field which recognizes gross zonal and regional differences as well as seasonal differences. Such a description should enable us to assess the magnitude of errors associated with data-derived estimates of spatially aggregated temperatures. Further, the statistical model, when completed, will provide a coherent basis for understanding and quantifying how errors are affected by the level of spatial aggregation, how errors are affected by the location and spatial density of the monitoring data, how errors vary from region to region, how errors vary from season to season, and how errors are affected by temporal aggregation.

While such a detailed statistical model for spatial-temporal variability may not be strictly required to address questions of global change spanning a century, it will provide a single framework for simultaneously assessing both global and regional change estimates, the effects of the evolving monitoring network, and estimates of rates of change for various epochs. Such a model should also prove useful for rational incorporation of spatially aggregated remote-sensing data, since the model will provide a calibration between variability of spatial aggregates and point data represented by the monitoring network.

Having said all this, practically the whole project lies ahead of us. An important preliminary step is to partition the globe into regions with relatively homogeneous variability. To some extent this step could rely on the monitoring data themselves. However, a simple stratification based on geographic-climatic zonation should capture the important distinctions.

2. ILLUSTRATION

As a start we have defined a "steppe" region of Eastern Europe to serve as the initial test region for development of a statistical model of spatial-temporal variability. We were aiming to have zones with some 30-50 monitoring stations which would provide the data for building the variability model. This particular region contains 39 monitoring stations, not all of which were simultaneously active during the period of the historical data. The climatic classification and the surface temperature data for these 39 stations were provided by the Goddard Institute for Space Studies. The locations of these stations are shown in Figure 1.

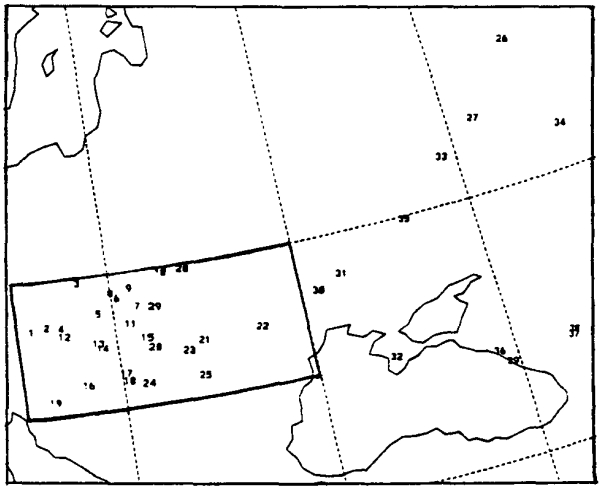


FIGURE 1: MAP OF THE AREA.

To further simplify the initial task of variability modeling we used data only for the period 1951-1990 because of the very spotty record before 1951. Still, a number of stations ceased to provide information after 1970 so that the reality of a changing geographic distribution of data is certainly present in our test region. The data were provided as monthly average temperatures. The pattern of missing data is represented in Figure 2 for the month of January. This missing data pattern is essentially the same for all other months.

Figure 3 shows five time series corresponding to data from a single station for the months January, April, July, and October, together with the average for all months. The substantially larger winter inter-annual variability is clearly seen. Time series of this kind can be analyzed for the presence of trends, either singly or in combination with similar series from other stations.

However, we decided to initially convert all time series to their annually differenced form, i.e., the annual increments, which has some advantages for the analysis. First, the question of time trend reduces immediately to one of estimation of a

mean level. Second, heterogeneity among different stations within the region is partially mitigated. Third, we remove the need to choose a baseline period which can be problematic where monitoring consistency across the network is lacking. Of course, the differenced form of the data can create its own problems in the analysis due to a possibly more complex pattern in temporal autocorrelation, but this additional complexity is not difficult to handle.

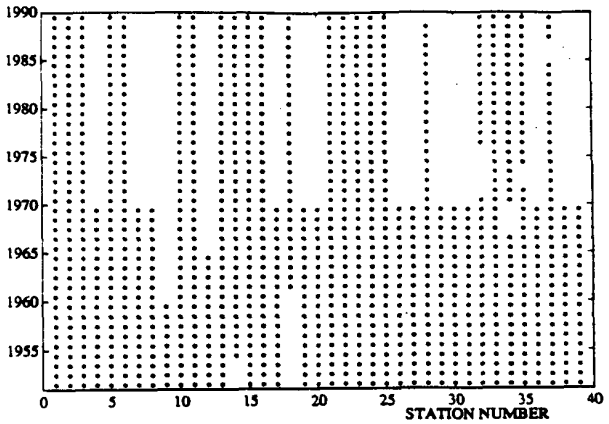


FIGURE 2: PATTERN OF AVAILABLE DATA.

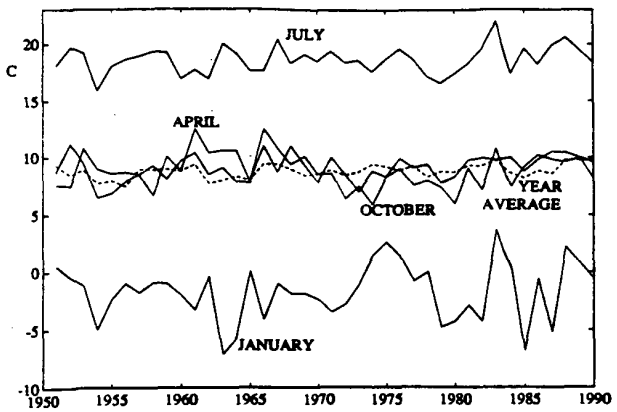


FIGURE 3: TEMPERATURE TIME SERIES FOR STATION #1.

The eight time series of Figure 4 show the inter-annual fluctuations of annual increment for each of four stations [numbered 1,2,34,37 in Figure 1] for both January and July. Note the very close match between the station pairs 1,2 and 34,37 for both winter and summer. Indeed, in winter all four stations exhibit similar patterns over time and will show strong cross-correlation. However, the amplitudes of the inter-annual fluctuations are not all alike. Furthermore, the summer cross-correlations between the station pairs like 1,34 is substantially negative.

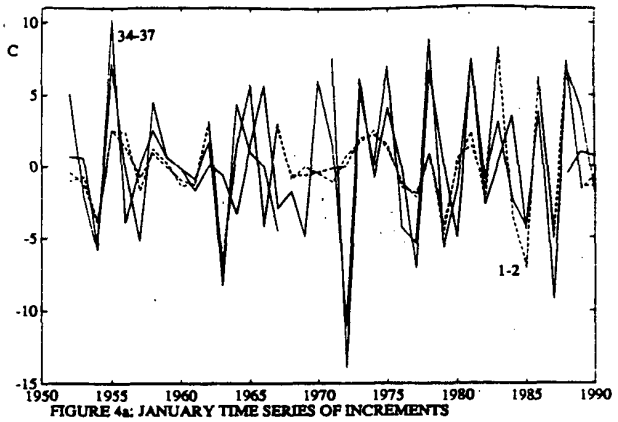


FIGURE 4a: JANUARY TIME SERIES OF INCREMENTS AT STATIONS #1,2,34 & 37.

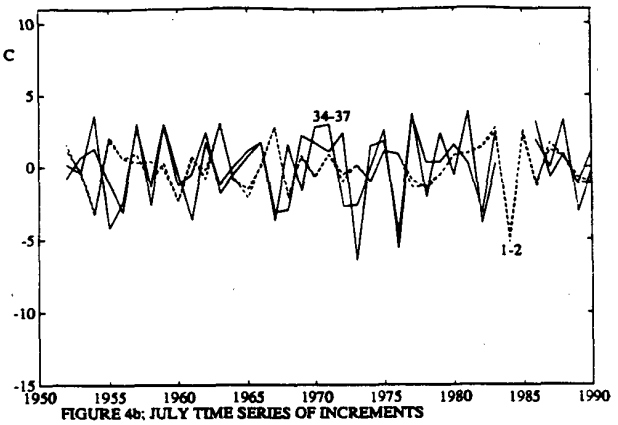


FIGURE 4b: JULY TIME SERIES OF INCREMENTS AT STATIONS #1,2,34 & 37.

The time series of Figure 4 are illustrative of some of the features of the data as seen in individual time series of annual temperature increment. A more comprehensive view of both inter-station spatial variability, seasonal variability, and inter-annual variability is graphically portrayed in the next series of plots.

A rough idea of the regional inter-annual variability is shown by the time series in Figure 5. These time series are obtained by averaging the annual temperature increments across all stations for which data were available. Separate time series are shown for January and July. What is remarka-

ble is that the amplitude of the inter-annual variability is not much less than the corresponding amplitude for individual station time series as seen, for example, in Figure 4.

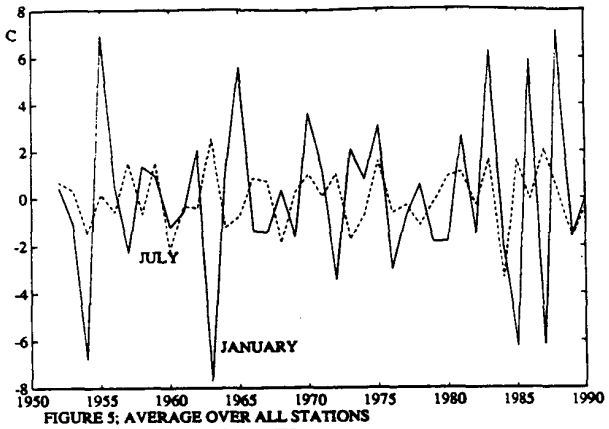


FIGURE 5: AVERAGE OVER ALL STATIONS OF TEMPERATURE INCREMENTS.

Year-to-year autocorrelation between annual temperature increments is shown in Figure 6 for time lags up to six years. Separate autocorrelation plots are shown for summer, winter, and the annual average. These scale-free year-to-year autocorrelations appear to be weak overall, except for the 1-year lag which shows an autocorrelation of -1/2, approximately. This particular -1/2 autocorrelation for the annual temperature increment is exactly what one expects from a stationary and purely uncorrelated time series of temperatures. Roughly speaking, this negative autocorrelation is the embodiment of climatic memory on a decadal time scale, as opposed to independent increments, say, where only the most recent datum is needed to predict the next time step.

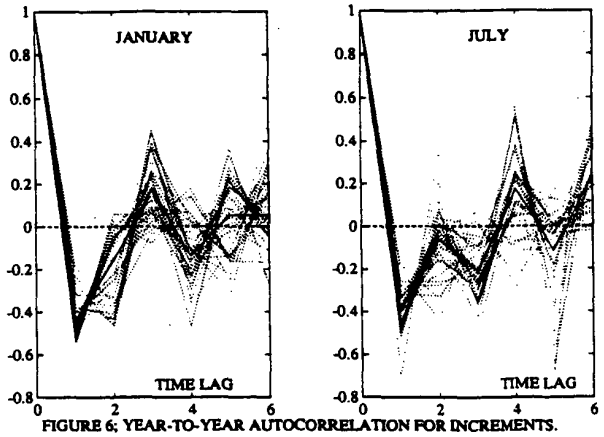


FIGURE 6: YEAR-TO-YEAR AUTOCORRELATION FOR INCREMENTS. EACH LINE CORRESPONDS TO A DIFFERENT MONTH.

Inter-station variability, as expressed through spatial differences is displayed in Figure 7. For each pair of stations we computed the average squared difference between the two time series of annual temperature increments, using data for years in which both stations reported. These averaged squared differences were plotted against the geographic distance [spatial lag] between the two stations. Thus each plotted point in Figure 7 represents a station pair and the ensemble of points is usually called a variogram scatterplot.

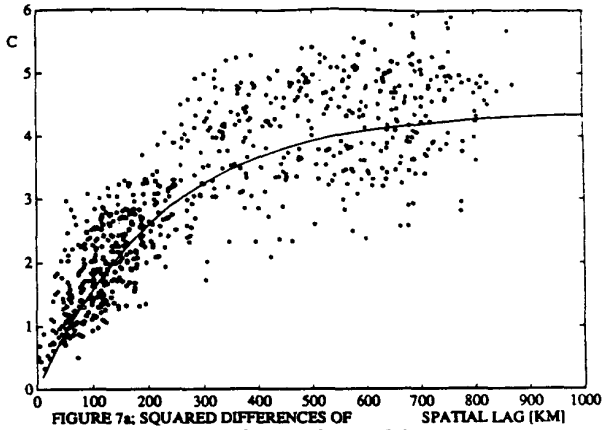


FIGURE 7a: SQUARED DIFFERENCES OF CONTEMPORANEOUS INCREMENTS, JANUARY.

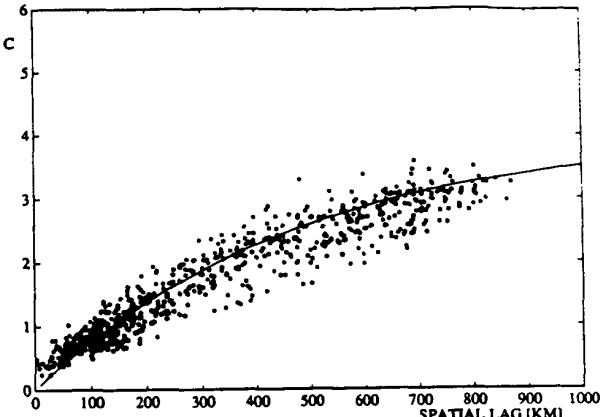


FIGURE 7b: SEE 7a.

Separate spatial variogram plots are shown for both January and July. It is noteworthy that the variogram for July shows a decidedly shorter correlation length, while both summer and winter months show the same inter-station variability at long distances.

Another way to exhibit the seasonal differences in inter-station variability is shown in Figure 8. The average annual increment over the 1950-1990 period was calculated separately for each month of the year and each station. The salient feature of this plot is, again, the difference between summer and winter inter-station variability.

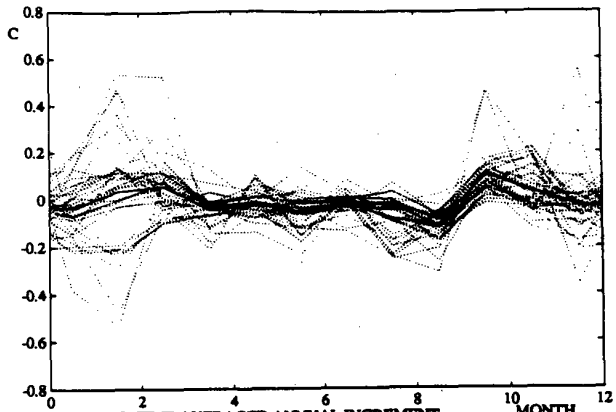


FIGURE 8: TIME AVERAGED ANNUAL INCREMENT AS A FUNCTION OF MONTH. EACH LINE IS A DIFFERENT STATION.

3. MODELING

The foregoing graphical representations of the variability of the data for the illustrative steppe region of Eastern Europe can serve as a modeling tool. The salient features are the inter-annual variability of the region average, rapid decay of the inter-annual autocorrelation, inter-station variability for fixed time, slowly decaying spatial autocorrelation for fixed time, seasonal differences in both inter-annual variability and inter-station variability, and modest spatial variability for time-averaged data.

Because of substantial seasonal differences in variability statistics, we have chosen to describe and model variability as a separate exercise for each of the 12 months of the year. Later the separate statistical descriptions will be combined using the inter-seasonal correlations. In the following notation the seasonal dependence of the quantities described is suppressed:

z[x,t] = temperature at location x
for years t = 1951, .. 1990

x is an arbitrary location in the region of interest

x = x1, x2, .. , xi, .. are the station locations

d[x,t] = z[x,t] - z[x,t-1]
the one-year temperature increment at x.

d[t] = spatially averaged value of d[x,t]
over all locations x in the region.

We decompose the one-year temperature increments for year t and location x as follows:

d[x,t] = d + a[t] + r[x,t], where

d = "true" annual rate of change [climate forcing]

a[t] = regional inter-annual modulation,
regarded as a realization of a stationary
zero-mean time series

r[x,t] = residual space-time field.

The residual field, r, can be endowed with complex structure which would reflect spatial differences in the time-averaged residuals as well as spatial differences in the inter-annual variability. However, in order not to introduce too much complexity into the initial analysis, we have chosen to homogenize certain aspects of temporal and spatial specificity. In particular, r will be regarded as a time series of repeated realizations from the same isotropic and stationary spatial random field. Furthermore, realizations separated by more than one year will be regarded as temporally uncorrelated. There is a reasonable basis in the data for adopting such a homogenization in the analysis.

The goal of the analysis is to elucidate the patterns of variability and to provide an encompassing statistical model. One immediate application is to provide an error assessment for estimates of the regionally averaged temperature increment, d[t], for each year t. The estimate of d(t) which we use is an optimized linear combination of the observed annual increment data available for year t.

Table 1 below illustrates estimates of the optimized regional annual temperature increment separately for January and July for each of the two years 1952 and 1980. The optimization uses a fitted exponential variogram function of spatial lag

as shown in Figure 7. The boundaries of the region are shown in Figure 1 are were chosen to simplify computation for this illustration. For comparison, the simple averages of the available station data are also shown in parentheses.

TABLE 1. Regional Estimates

	1952 estimate	error	1980 estimate	error
January	-0.46 [-1.24]	0.17	-1.93 [-1.54]	0.18
July	0.92 [1.36]	0.12	1.08 [0.85]	0.14

Provided the coefficients in the linear combination of station data sum to unity, it is possible to express the root mean squared error of the estimated regional average in terms of the variogram function. These error estimates are also shown in parantheses in Table 1 and were derived using the variograms of Figure 7. It is instructive to note how the estimation error changes so little with time even though the 1952 estimates are based on 25 stations and the 1980 estimates are based on 17 stations. The July regional estimates show noticeably better precision because of the shorter spatial correlation length in summer months.

The same basic decomposition and simplification of spatial and temporal variability of the annual increment data can be used also to optimize estimates of time-averaged regional increment and to provide corresponding error assessments. Also, the "true" annual rate of change, d, can be estimated from all the available data, and estimates of the monitoring period needed to reach required levels of precision can be derived. The effects of an increased or decreased monitoring density can be quantitatively evaluated.

Finally, there are questions which are not taken up here at all but which are an important part of our research effort. For example, the knitting together of separate seasonal models and separate regional models will be needed to assess the uncertainty of estimates of global change.

THE JOINT NORMAL MODES OF THE COUPLED ATMOSPHERE-OCEAN SYSTEM OBSERVED FROM 1967 TO 1986

Jin-Song Xu
Max-Planck-Institut für Meteorologie

Two aspects of the Principal Oscillation Pattern (POP) analysis are used to study the joint normal modes of the coupled atmosphere-ocean system from a combined dataset including both atmospheric (sea level pressure, 700-mb and 200-mb zonal wind) and oceanic (sea surface temperature, Pacific sea level and Pacific subsurface temperature) parameters. The first aspect is that the Principal Oscillation Patterns (POPs) can be considered as estimated normal modes of the system. We assume that the considered system can be described by a linear equation:

$$\mathbf{x}(t+1)=\mathcal{A}\mathbf{x}(t) \qquad (1)$$

corresponding to a linearized dynamical model where $\mathcal{A}$ describes the linear dynamics of the system. The only difference between this and a theoretical dynamical approach is that $\mathcal{A}$ is not derived from dynamical theory but estimated from the data. POPs are the eigenvectors of the matrix $\mathcal{A}$. The correspondence between the estimated normal modes (the POPs) and theoretically derived normal modes has been demonstrated by Schnur et al. (in the same proceedings) for the atmospheric baroclinic waves. In this study, the atmosphere-ocean system is considered. A straightforward way to study the linear dynamics of such complex coupled system is to estimate the system matrix $\mathcal{A}$ from the data. After having estimated the POPs, $\mathbf{x}(t)$ can be represented in the space spanned by the POPs ($\mathbf{P}_j$):

$$\mathbf{x}(t)=\sum z_j(t)\mathbf{P}_j \qquad (2)$$

where $z_j(t)$ are the coefficient time series of $\mathbf{P}_j$.

The second aspect of the POP analysis is that it can be considered to be a multivariate spectral analysis. Following the linear assumption (1), the time evolution of the POP coefficient time series $z_j(t)$ is described by:

$$z_j(t+1)=\lambda_j z_j(t)+n_j(t) \qquad (3)$$

where λ_j is the eigenvalue corresponding to pop $\mathbf{P}_j$. If (3) is restated in the frequency domain, the autospectrum Γ of $z_j(t)$ can be calculated:

$$\Gamma_{z_j}(\omega) = \frac{\Gamma_{n_j}(\omega)}{(e^{i\omega} - \lambda_j)\ (e^{-i\omega} - \lambda_j^*)} \qquad (4)$$

It is expected that the autospectrum Γ of the noise $n_j(t)$ is white or at least very smooth. In this case, the spectral characteristics of the modes are determined by the eigenvalue λ_j. For complex λ_j, i.e. $\lambda_j = |\lambda_j| e^{i\phi_j}$, (4) describes a peak spectrum which indicates an oscillation with oscillation period around $2\pi/\phi_j$. For real λ_j, the problem reduces to a first order auto-regressive process, i.e. (4) describes either a white or red noise spectrum depending on the amplitude of λ_j.

Five real modes and one complex mode are found in the coupled atmosphere-ocean system. Fig.1 shows the eigenvalues and the POP coefficient time series of the six modes. There is a clear conrrespondence between the spectral features suggested by λ_j and those given by a conventional spectral analysis of the time series.

Among the six modes, mode 4 describes a decadal time scale variation which is not yet well documented. It is characterized by the simultaneous development of tropical SST anomalies which have greatest amplitude over the west Pacific (Fig.2a), and anomalous convection over the same region (Fig.2b). According to (2), the signal given by the time series shown in Fig.1d and the patterns shown in Fig.2a, b describe a decrease of the tropical SST and a strengthening of the ascending branch of the Walker Circulation over the west Pacific from the late-sixties to the mid-seventies, and an increase of the tropical SST and the west Pacific convection since the mid-seventies. The results indicate an active air-sea interaction process in the tropics on the decadal time scale. Further investigation of the patterns allows us to speculate that this tropical variation could induce changes in the extratropical tropospheric circulation with largest wind anomalies in the North Pacific (Fig.2b) and that these might cause changes in the subtropical and subpolar gyres in the North Pacific. The latter are indicated by the negative (positive) sea level changes along the east coast of South (North) Japan (Fig.2c) and by the zonally distributed North Pacific SST anomalies (Fig.2a).

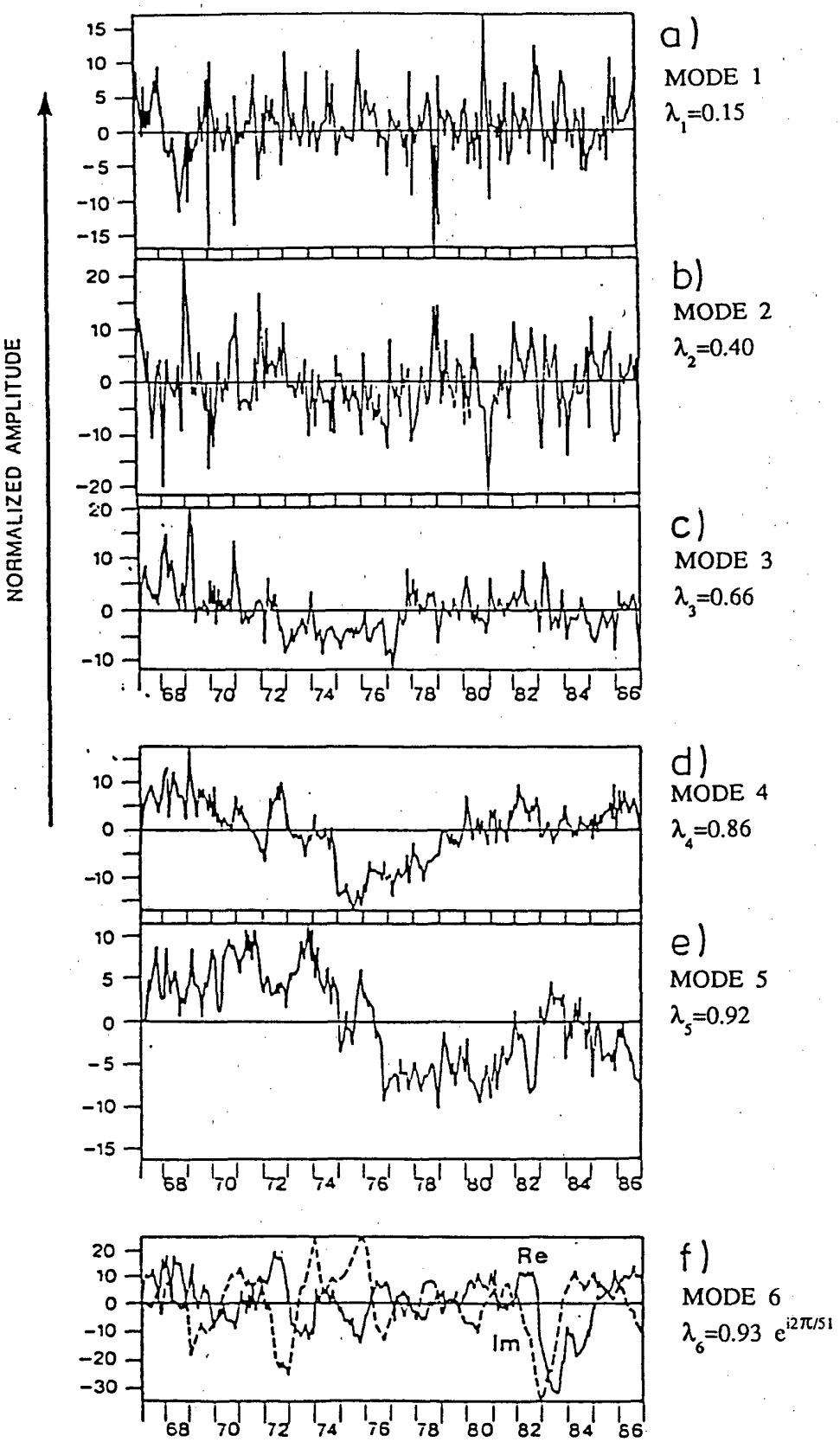


Figure 1: Eigenvalues and POP coefficient time series of the six modes found in the coupled atmosphere-ocean system.

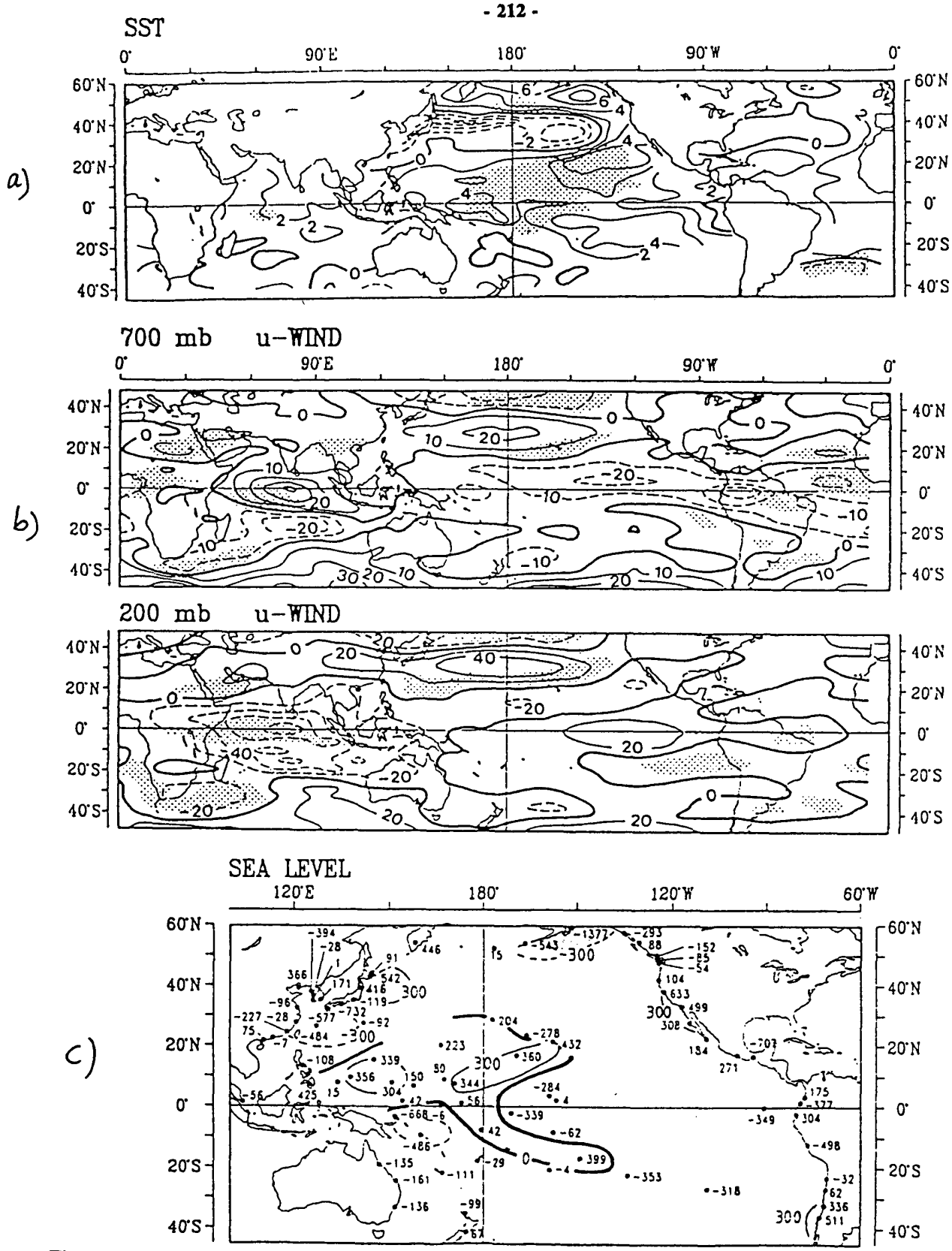


Figure 2: POP pattern P of the mode 4 for a) SST (in °C), b) 700-mb and 200-mb zonal wind (in m/s), and c) sea level (in mm). Shaded areas indicate explained variance larger than 10% for sea level and larger than 5% for SST and zonal wind.

Principal Oscillation Patterns and Stability Analysis: Estimated and Quasi-Geostrophic Normal Modes of the Atmosphere*

Reiner Schnur
Max-Planck-Institut für Meteorologie, Hamburg, Germany

Abstract

The Principal Oscillation Pattern (POP-) Analysis is a technique to empirically identify time-dependant spatial patterns in a multivariate time series of geophysical or other data. POPs can be seen as normal modes of a linear approximation to a system the dynamics of which are unknown or too complex. The system matrix is estimated from observations of nature.
A conceptually different approach to the derivation of normal modes is to consider a linearization of a possibly complex set of dynamical equations and to compute the corresponding eigenmodes. This requires of course a thorough understanding of the dominating dynamics.
These two concepts have been compared in the context of baroclinic unstable waves which are responsible for much of the high-frequency variability in midlatitudes. For the statistical part, a POP-Analysis has been applied to tropospheric geopotential height fields of ECMWF analyses from 1984 to 1987. The data have been subjected to zonal Fourier decomposition and to time filtering so that variations with periods between 3 and 25 days were retained. Analyses have been performed separately for each zonal wave number 5 to 9 on both hemispheres in winter. On the other hand, a conventional stability analysis has been performed where the system matrix of the linearized system is derived by quasi-geostrophic reasoning from the potential-vorticity equation for a stratified fluid on a hemisphere. Only the basic state depends on time- and space-averaged fields of observed wind and temperature from the ECMWF data. By analyzing the observations in a way that is analogous to the way the theoretical dynamics are analyzed (normal modes) the comparison between observations and theory is facilitated.
It turns out that the most significant POPs are very similar in time and spatial structure to the

---

* Available as "Schnur, R., G. Schmitz, N. Grieger and H.v. Storch, 1992: *Normal Modes of the Atmosphere as Estimated by Principal Oscillation Patterns and Derived from Quasi-Geostrophic Theory*. Report No. **81**, Max-Planck-Institut für Meteorologie, Bundesstr. 55, 2000 Hamburg 13, Germany".

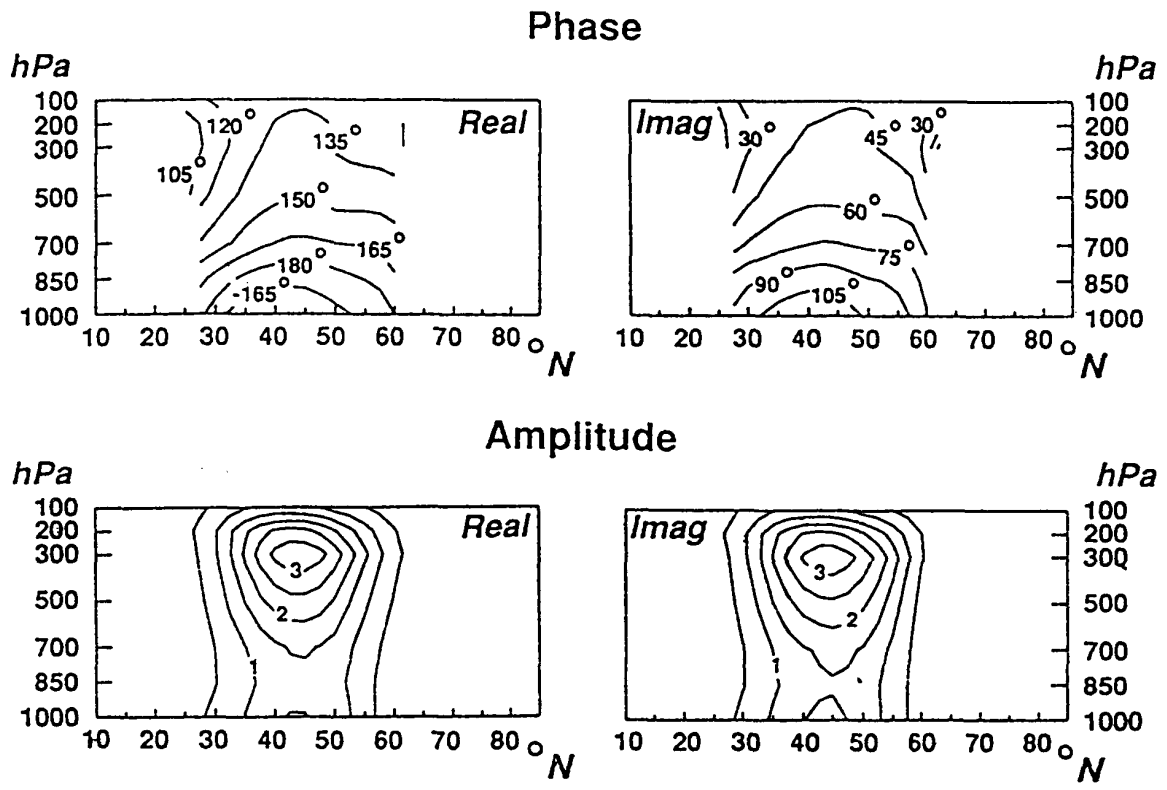


Figure 1:
POP analysis of twice-daily geopotential height in the Northern Hemisphere for the three winters (DJF) 84/85 to 86/87 and for zonal wave number 8: the POP which describes the linear growing phase of an unstable baroclinic wave. The POP explains 35% of the total wave number 8 variance, its oscillation period T is 4 days.
Since the analysis is done in wave-number space the analyzed time series consists of the sine- and cosine coefficients for a wave number 8 wave at all levels and latitudes. Thus both the real and imaginary part of the POP have to be interpreted as patterns of sine- and cosine coefficients, too. They can be represented by phase (top) and amplitude patterns (bottom) in latitude-height diagrams.
The evolution of the pattern is defined by the sequence

... → imaginary part → real part → - imaginary → - real → imaginary

where the transition from one pattern to another is performed in a quarter of the period T. Since the phases of the real (left) and imaginary (right) part are out of phase by about 90° (i.e. a quarter of a period) this gives the picture of an eastward propagating wave pattern. The absolute values of the amplitudes are arbitrary.

most unstable waves in the stability analysis. They describe the linear growth phase of baroclinic unstable waves which propagate eastward with periods of 3 to 7 days. Figure 1 shows, as an example, the spatial pattern of the Northern Hemisphere POP for zonal wave number 8 by means of amplitude/phase diagrams. The amplitude of the POP is characterized by a geopotential maximum in the upper troposphere and at 45°N. The phases of the real and imaginary part define an eastward propagating wave pattern which tilts slightly westward between the surface and 300 hPa. Above 300 hPa the phase is nearly constant. There is only a small dependence of the phase with latitude indicating no meridional momentum transport. Eliassen-Palm cross sections show that northward heat flux at low levels and vertically propagating wave activity from the surface into the middle troposhpere are the primary mechanisms connected with this mode. These features are in very good coincidence with the corresponding mode from the linear stability analysis as well as with other theoretical studies (e.g. Simmons and Hoskins, 1976). The interpretation of the POP as an unstable baroclinic wave in its growing phase is thus confirmed.

Moreover, the POP analyses reveal patterns which are not found in the linear stability analysis. Figure 2 shows this POP pattern for wave number 8 of the Northern Hemisphere analysis. The amplitude pattern has a broad maximum in the upper troposphere with peak values at 35°N. In the vertical the wave has only a slight westward tilt, but there is now a strong meridional phase gradient resulting in a southwest-northeast tilted wave. This POP is reminiscent of the latitudinal structure of baroclinic waves in their nonlinear decay phase which where e.g. identified by Hoskins and Simmons (1978) with a primitive equation model. They are characterized by a strong northward transport of momentum which corresponds to a distinct Rossby wave dispersion. Eliassen-Palm cross-sections for this mode also show that the direction of wave activity propagation turns equatorwards in this stage.

Summarizing, the results show a nice equivalence between the two approaches, nameley, the empirical method which uses atmospheric observations and the theoretical method which uses first-order dynamical reasoning. Since the POPs are purely derived from observations the results also indicate the appropriateness of the assumptions usually made in linear stability analysis of zonally symmetric flows to explain high-frequency atmospheric fluctuations. The POP analyses also reveal waves which are not found in linear stability analysis and which describe the nonlinear decay phase of baroclinic waves. Thus the results of the POP analyses can be interpreted in terms of the life cycle of unstable baroclinic waves (cf. Randel and Stanford, 1985).

References:

Randel, W.J. and J.L. Stanford, 1985: The Observed Life Cycle of a Baroclinic Instability. *J. Atmos. Sci.*, 42, 1364-1373.

Simmons, A.J. and B.J. Hoskins, 1976: Baroclinic Instability on the Sphere: Normal Modes of the Primitive and Quasi-Geostrophic Equations. *J. Atmos. Sci.*, 33, 1454-1477.

—— and ——, 1978: The Life Cycles of Some Nonlinear Baroclinic Waves. *J. Atmos. Sci.*, 35, 414-432.

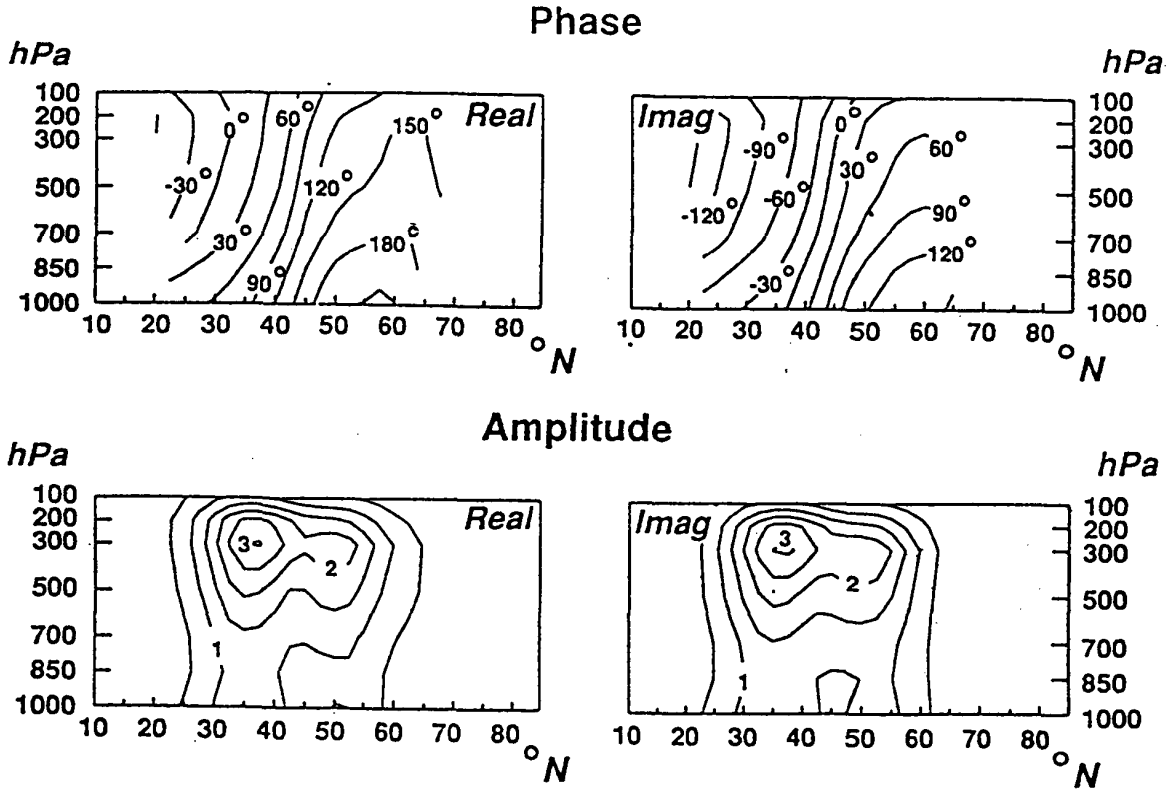


Figure 2:
As in Figure 1 but for the POP which is attributed to the decay phase of an unstable baroclinic wave.

COMPLEX POP ANALYSIS

Gerd Bürger, Hamburg, Germany

The Complex Principal Oscillation Pattern (CPOP) analysis is an extension of the conventional POP analysis which was invented by HASSELMANN [1988] and STORCH et.al. [1988]. As will be shown, the latter analysis technique is unable to resolve standing oscillations, a fact which holds for *any linear system of first order*. To overcome this problem we introduce, just like in Hamiltonian mechanics, for a stationary process $\mathbf{q}(t)$ the *conjugate process* $\mathbf{p}(t)$. $\mathbf{p}$ results from $\mathbf{q}$ by shifting the process at each frequency by an amount of $\pi/2$. So $\mathbf{p}$ can be considered as the momentum of $\mathbf{q}$. The transition $\mathbf{q} \longrightarrow \mathbf{p}$ is known as the Hilbert transform and it was already introduced into the geophysical literature by WALLACE & DICKINSON [1972]. We now consider the *complex process*

$$\mathbf{z}(t) = \mathbf{q}(t) + i\cdot\mathbf{p}(t)$$

and assume, analogously to the POP analysis, that it is the output of a stochastically driven linear system of first order. It is of the form

$$\mathbf{z}(t+1) = \mathfrak{C} \cdot \mathbf{z}(t) + \textit{noise} \qquad (*)$$

with a complex system matrix $\mathfrak{C}$. $\mathfrak{C}$ is to be estimated from the data according to

$$\mathfrak{C} = \mathrm{COV}_1 \cdot \mathrm{COV}_0^{-1}$$

with COV_i denoting the covariance matrix of lag i. In this way *state and momentum* of the process are modeled *simultaneously*.

The modes of this system are the *CPOPs* of the process $\mathbf{q}$. So for an n-dimensional process $\mathbf{q}$ one finds n (usually) different complex patterns $\mathbf{C}_j = \mathbf{C}_j^R + i\cdot\mathbf{C}_j^I$, the eigenvectors of $\mathfrak{C}$, supplied with eigenvalues λ_j. With the CPOPs the complex state $\mathbf{z}$ of the system can be rewritten in the following form:

$$\mathbf{z}(t) = \gamma_1(t) \cdot \mathbf{C}_1 + \ldots + \gamma_n(t) \cdot \mathbf{C}_n \qquad (\dagger)$$

with $\gamma_j(t)$ being CPOP coefficients. Utilizing the CPOP coefficients and writing $\lambda = \rho \cdot e^{i\omega}$, the coupled equation (*) can easily be solved:

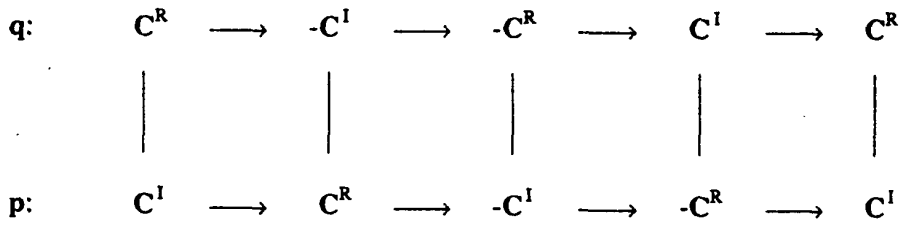
$$\gamma_j(t) = \rho^t \cdot e^{i\omega t} \quad . \qquad (\ddagger)$$

Since both state and momentum are modeled the evolution $C(t) = C \cdot \gamma(t)$ of a mode C with corresponding eigenvalue λ is twofold:

in the *state* space: $\quad C^R(t) = \rho^t \cdot [\cos(\omega t) \cdot C^R - \sin(\omega t) \cdot C^I]$

in the *momentum* space: $\quad C^I(t) = \rho^t \cdot [\cos(\omega t) \cdot C^I + \sin(\omega t) \cdot C^R]$

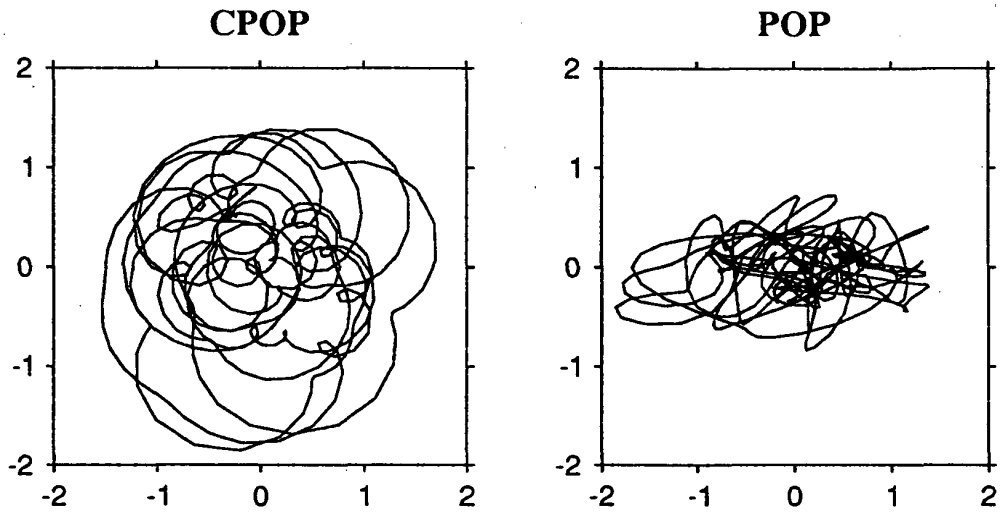
Hence, the state as well as the momentum process perform a damped oscillation between the two patterns C^R and C^I; the only difference between them is that the state leads the momentum by a quarter of a period.



In order to test whether the CPOP analysis works correctly we produced a series of low dimensional complex autoregressive processes of first order with prescribed modes. The real part of this complex process was then subjected to a CPOP analysis. The modes are chosen in such a way that small as well as large travelling rates (a term to be defined) are present. We shall show that in fact there is no significant difference between the 6 input modes and the resolved CPOPs. A POP analysis of the same process, however, can resolve at most half of the maximum number of oscillations, and it resolves only those modes which have a large travelling rate.

To test the CPOP model further and to compare the results with the POP model we analyzed a time series of equatorial monthly SST anomalies. This time series is a composition of two independent series which span the indic and pacific ocean between 45E and 85W. The first series, which ranges from 1951 to 1986, is an updated form of the data set described in BARNETT [1983]. The second series ranges from 1982 to 1991 and is an updated form of the data set described in REYNOLDS [1988].

We subjected the 24-dimensional SST process to a POP and a CPOP analysis. The POP analysis was able to resolve a signal only after a time-filtering which retained periods between 10 and 100 months. The only POP which explained a substantial amount of variance (20%) did not develop very coherently and had a vague spectral peak near 30 months. The CPOP analysis resolved a 38 months standing oscillation which explaines about 38% of the variance (this is nearly the same amount as the first EOF!) and which develops coherently at the 99%-level at that period. The figure shows the complex POP and CPOP coefficient, the POP coefficient being calculated by evoking (†). For a completely undisturbed (and undamped) mode $\gamma(t)$ would perform a clean cycle according to (‡). For visibility, we used a smoothing.



References:

BARNETT [1983]. - *Interaction of the Monsoon and Pacific Trade Wind System at Interannual Time Scales I.* Mon. Wea. Rev. 111 (756-73)

HASSELMANN [1988]. - *PIPs and POPs. The reduction of complex dynamical systems using Principal Interaction and Oscillation Patterns.* J. Geophys. Res. 93 (11015-21)

STORCH,BRUNS,FISCHER-BRUNS,HASSELMANN [1988]. - *Principal Oscillation Patterns analysis of the 30-60-day oscillation in General Circulation Model Equatorial Troposhere.* J. Geoph. Res. 93,D9 (11022-36)

REYNOLDS [1988]. - J Climate 1 (75-86).

WALLACE,DICKINSON [1972]. - *Empirical Orthogonal Representation of Time Series in the Frequency Domain I.* J. App. Met. 11,6 (887-92)

A neural network approach to the estimation of state space models

B. Grieger
Max-Planck-Institut für Meteorologie, Hamburg, FRG

Abstract. State space models are used to describe dynamical systems with many degrees of freedom. In the principal interaction pattern (PIP) formulation the state space is an arbitrary sub vector space of the observation space and the system function is a prescribed nonlinear function which may depend on various *a priori* unknown parameters. The problem is to obtain these function parameters and a set of patterns which put up the state space from a given set of observations. The neural network approach described in this paper provides a suitable class of nonlinear system functions accompanied by a matching optimization procedure. Applications of this algorithm on artificial test data and observational data of the El Niño/Southern Oscillation (EN/SO) phenomenon demonstrate that it is possible to recover limit cycle attractors.

1 Introduction

State space models of dynamical systems with many degrees of freedom can be constructed with the principal interaction pattern (PIP) method due to Hasselmann (1988). However, in most applications (e. g. Storch 1988) a simplified linear version is used, the so-called principal oscillation pattern (POP) analysis. The POP method is able to disentangle various damped oscillations of a stochastically driven system, but the estimated model is restricted to exactly one fixpoint attractor. There are only a few applications of the nonlinear PIP analysis (e. g. Storch *et al.* 1990), because it is difficult to find a suitable class of system functions, and depending on that function class the estimation of the model can be quite cumbersome.

During the last years neural networks have successfully been applied on a variety of problems. It has been shown by Elsner and Tsonis (1992) that neural networks can be used for the nonlinear prediction of a univariate chaotic time series.

In this paper a state space model is described in terms of neural networks. The optimization of a suitable structured neural network with the backpropagation learning algorithm is comparable with fitting a PIP model by error minimization using steepest descent methods. The neural network formulation provides a large class of model functions accompanied by analytical derivatives, which considerably hastens the optimization procedure.

2 Theory

In a layered feed-forward network the output of neuron i in layer k is given by

$$y_i^{(k)} = f_{\text{out}}^{(ki)}\left(\sum_{j=0}^{N_{k-1}} W_{ij}^{(k)} y_j^{(k-1)}\right), \qquad (1)$$

where N_{k-1} is the number of neurons in layer $k-1$, $W_{ij}^{(k)}$ is the weight for input j of neuron i in layer k, and $f_{\text{out}}^{(ki)}$ is the output function of neuron i in layer k. Below a three-layered network with $f_{\text{out}}^{(1i)} = f_{\text{out}}^{(2i)} = \tanh$ and $f_{\text{out}}^{(3i)} = \text{id}$ will be used. $y_j^{(0)}$ denotes the input vector. Given a training set of input vectors accompanied by desired output vectors, the neural network can be optimized by backpropagation of the output errors through the whole network to adjust all weights (e. g. Widrow and Lehr 1990).

Consider a system represented by a high-dimensional state vector $\mathbf{\Phi} = (\Phi_1, \ldots, \Phi_n)$ which varies with time. We are looking for a small set of n-dimensional patterns $(\mathbf{p}_\nu)_{\nu=1,\ldots,m}$, so that every state the system achieves can be approximated by a linear superposition of this patterns. Let $\mathbf{x} = (x_1, \ldots, x_m)$ be the coefficients of this superposition, then at every time t the state of the system can be written

$$\begin{aligned} \mathbf{\Phi}(t) &= \sum_{\nu=1}^{m} x_\nu(t)\mathbf{p}_\nu + \text{Noise} \\ &= \mathbf{P}\mathbf{x}(t) + \text{Noise}, \end{aligned} \qquad (2)$$

where $\mathbf{P}$ is a matrix with columns $\mathbf{p}_\nu$.

The temporal evolution of the system is now described in the low-dimensional coefficient space, i. e.

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{F}[\mathbf{x}(t), \mathbf{x}(t-1), \ldots, \mathbf{x}(t-(\tau-1)); \\ &\quad \alpha_1, \ldots, \alpha_k] \\ &=: \mathbf{F}_\alpha(\mathbf{x}^{(\tau)}(t)), \end{aligned} \qquad (3)$$

where $\mathbf{F}$ is a function which has to be prescribed but may depend on an arbitrary number of free parameters $(\alpha_j)_{j=1,\ldots,k}$ and τ is the order of the system. For convenience we define $\mathbf{F}_\alpha$ to be the system function for the parameter set $(\alpha_1, \ldots, \alpha_k)$ and $\mathbf{x}^{(\tau)}(t)$ to be the sequence of the latest τ state space vectors.

The patterns $(\mathbf{p}_\nu)_{\nu=1,\ldots,m}$ and the function parameters $(\alpha_j)_{j=1,\ldots,k}$ can be obtained from a given set of observations $(\mathbf{\Phi}(t))_{t=1,\ldots,T}$ by the minimization

$$\sum_{t=1}^{T-1} \|\mathbf{\Phi}(t+1) - \mathbf{P}\mathbf{F}_\alpha(\mathbf{x}^{(\tau)}(t))\|^2 = \min. \quad (4)$$

If we describe the problem in terms of neural networks, the system function $\mathbf{F}$ can be represented by a two-layered nonlinear network ($f_{\text{out}} = \tanh$) where a sequence of state space vectors $\mathbf{x}^{(\tau)}(t)$ up to a certain time is the input vector and the state space vector $\mathbf{x}(t+1)$ one time step later is the corresponding output vector, i. e.

$$\mathbf{x}(t+1) = \tanh[\mathbf{W}^{(2)} \tanh(\mathbf{W}^{(1)}\mathbf{x}^{(\tau)}(t))]. \quad (5)$$

Then the weights $\mathbf{W}^{(1)}, \mathbf{W}^{(2)}$ in this network are the function parameters $(\alpha_j)_{j=1,\ldots,(\tau+1)mN_1}$.

The patterns $(\mathbf{p}_\nu)_{\nu=1,\ldots,m}$ can be taken as the weights of a one-layered linear network ($f_{\text{out}} = \text{id}$), where a state space vector $\mathbf{x}(t)$ is the input vector and the corresponding observation vector $\mathbf{\Phi}(t)$ is the output vector.

The two networks are tied together by taking the output of the former as input of the later. Fig. 1 shows the structure of a neural network which is applied on a three-dimensional system in section 3.

The weights of the complete network could be estimated with the backpropagation learning algorithm if the state space vectors $\mathbf{x}(t)$ were provided. These are also unkown, but they can be optimized simultaniously with the weights by extending the training procedure to the following two operations:

1. A sequence of coefficient vectors $\mathbf{x}^{(\tau)}(t)$ is choosen from the training set and presented to the network. The desired output vector $\mathbf{\Phi}(t+1)$ is used to adjust the weights of the network—the patterns $(\mathbf{p}_\nu)_{\nu=1,\ldots,m}$ and the parameters $(\alpha_j)_{j=1,\ldots,(\tau+1)mN_1}$—with the backpropagation algorithm.

2. The vectors $\mathbf{x}(t), \mathbf{x}(t-1), \ldots, \mathbf{x}(t-(\tau-1))$ are presented one by one only to the last layer of the network. The corresponding desired output vectors $\mathbf{\Phi}(t), \mathbf{\Phi}(t-1), \ldots, \mathbf{\Phi}(t-(\tau-1))$ are used to adjust the weigths of the last layer—the patterns $(\mathbf{p}_\nu)_{\nu=1,\ldots,m}$—a second time, and the input vectors $\mathbf{x}^{(\tau)}(t)$ by extending the backpropagation algorithm one step further back.

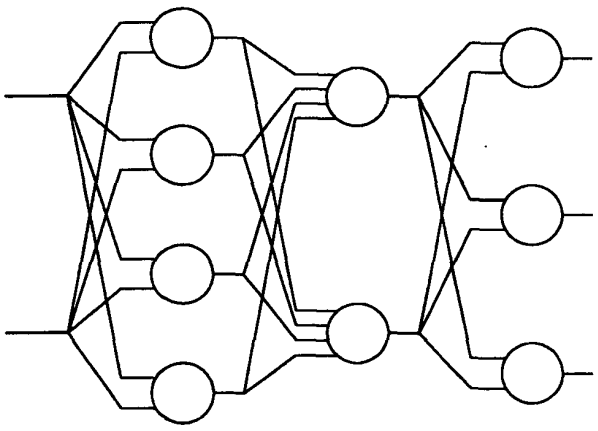


Figure 1: Structure of the neural network used for the estimation of a two-dimensional state space model for a three-dimensional system. Input (left side) is an actual state space vector, output of the second layer is the predicted state space vector for the next time step, and output of the last layer is the corresponding observation space vector.

The second adjustment of the patterns is necessary to achieve convergence in this situation, where the input vectors of the training set are changing their values during the training. In the beginning the state space vectors are set to zero and the network weights to small random values.

3 Applications

To test the described algorithm it was applied on simulated data of a three-dimensional artificial system. The system is of first order and has a limit cycle attractor. Two phase space trajectories of the undriven system are shown in Fig. 2. The simulated data are generated by driving the system stochastically and adding an observational error of about 5 per cent. The resultant phase space trajectory is also shown in Fig. 2.

The neural network used to estimate a state space model of this system is shown in Fig. 1. The system is three-dimensional, so the network has three units in the third (the last) layer. We prescribe the state space to be two-dimensional, so there are two units in the second layer. Because we assume the system to be of first order, we have the same number (two) of input values. The number of units in the first layer determines the complexity of the system function and is chosen to be four.

Optimization of this network yields two patterns which put up the state space as weights of the last

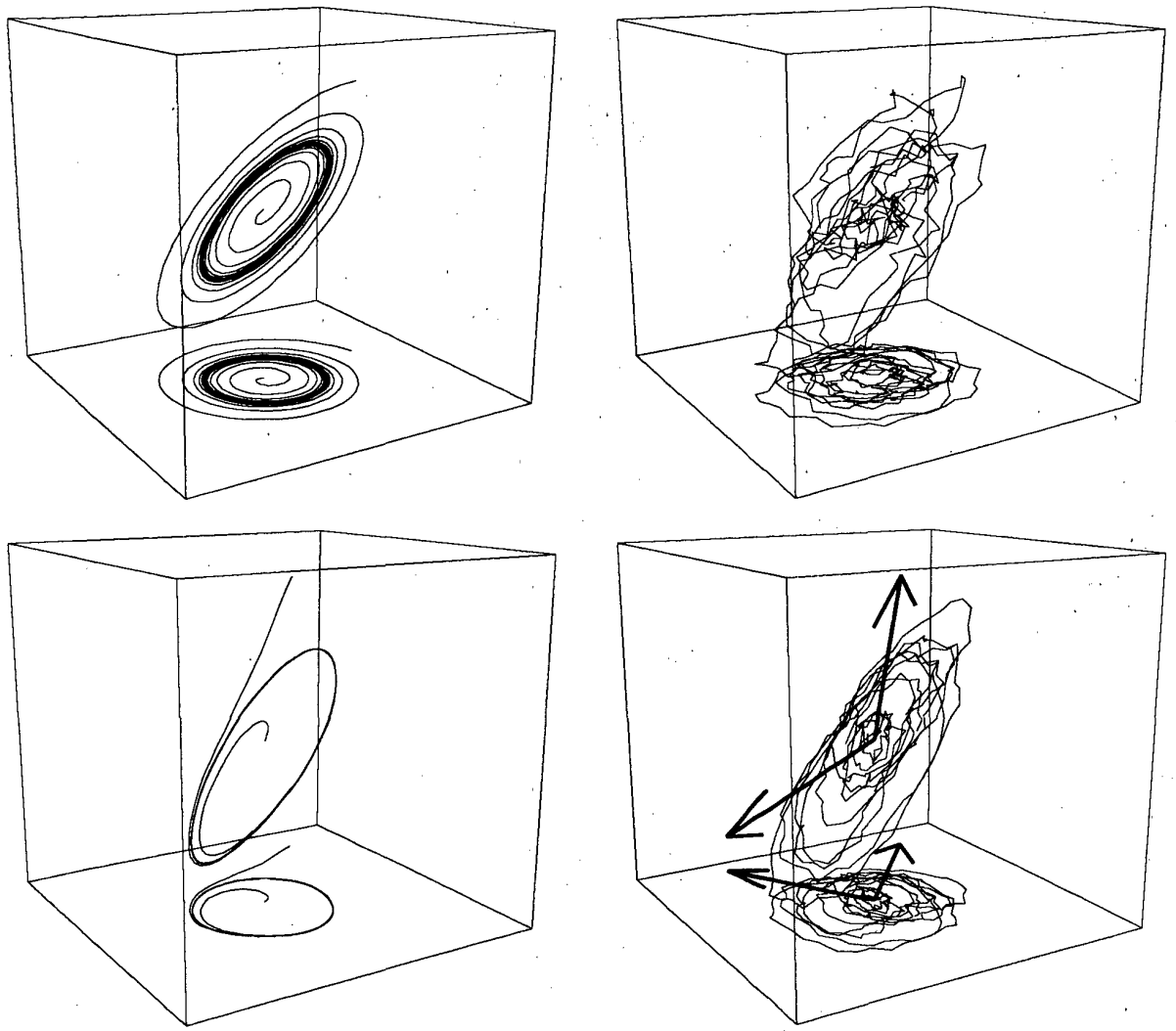


Figure 2: Application on an artificial system. *Upper left:* Two phase space trajectories of the undriven system, one staring outside and spiralling inward, the other starting in the middle and spiralling outward. Both end up on an elliptical limit cycle attractor. *Upper right:* "Observed" phase space trajectory of the stochastically driven system. *Lower right:* Observed trajectory projected on the estimated state space, which is spanned up by the two indicated vectors. *Lower left:* Two undriven trajectories as predicted by the estimated system function. The position and orientation of the attractor is recoverd with passable accuracy, but the strength of the attractor in the state space model is larger than the original. (In all figures the curve at the bottom of the box is just a projection—a "shadow"—to ilucidate the three dimensional structure.)

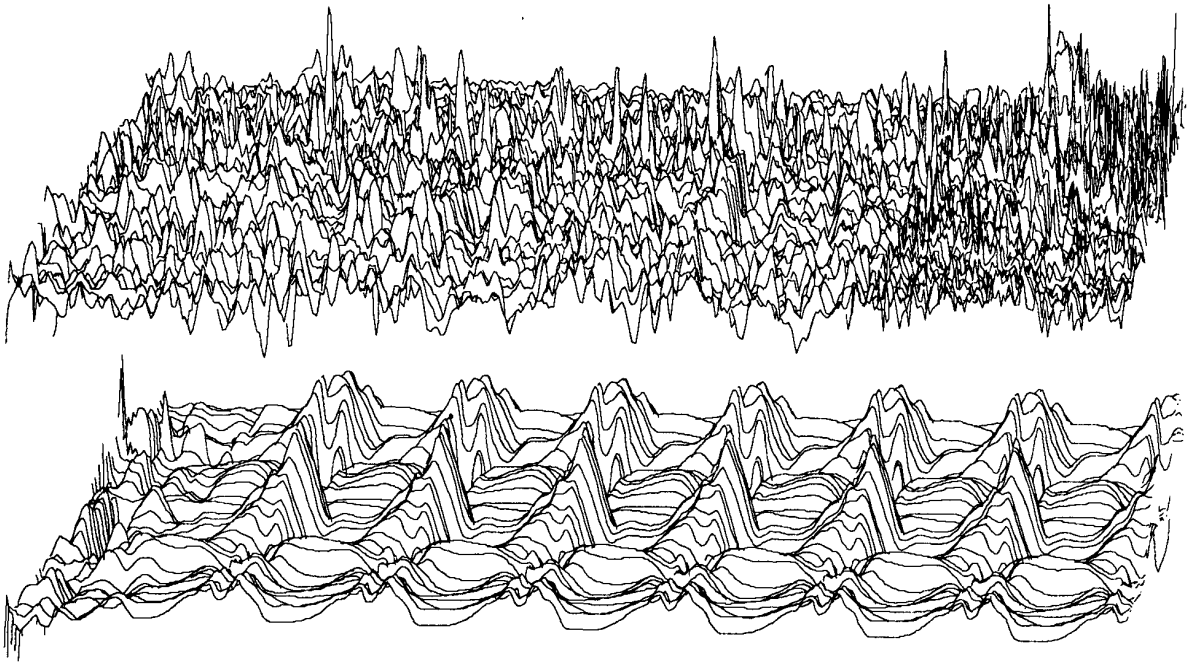


Figure 3: Application on the EN/SO phenomenon. In each of the two graphs the temporal variation of 48 observed values (monthly means, normalized variances) is shown, with time running from left to right. The nearer 24 curves are sea survace temperatures on even spaced points along the equater from 45°O to 85°W, the farer 24 curves are the corresponding zonal winds. *Top:* The observed data (January 1951 – December 1986). *Bottom:* An attractor found in the estimated ten-dimensional state space model.

layer and a system function which is represented by the first two layers. The patterns and two phase space trajectories as predicted by the system function are also shown in Fig. 2. The limit cycle attractor of the original system is recoverd in the model.

A second application was performed with observed data—sea suface temperatures and zonal winds—of the EN/SO phenomenon, see Fig. 3. The observation space is 48-dimensional, and the state space is chosen to be 10-dimensional. Again we assume the system to be of first order, and the number of neurons in the first layer is set to 20. The system function of the resultant state space model has a limit cycle attractor which is also shown in Fig. 3.

4 Conclusions

It has been demonstrated that neural networks are a useful tool for PIP analysis. A limit cycle attractor in an artificial system was recovered from simulated observations, and such an attractor was found in observed EN/SO data. To critisize the evidence of this result, more extensive tests with simulated data should be performed.

References

Elsner, J. B., Tsonis, A. A., 1992: "Nonlinear Prediction, Chaos, and Noise", *Bull. Am. met. Soc.* **73**, 49–60

Hasselmann, K., 1988: "PIPs and POPs: The Reduction of Complex Dynamical Systems Using Principal Interaction and Oscillation Patterns", *J. geophys. Res.* **93**, 11,015–11,021

Storch, H. von, Bruns, T., Fischer-Bruns, I., Hasselmann, K., 1988: "Principal Oscillation Pattern Analysis of the 30- to 60-Day Oscillation in General Circulation Model Equatorial Troposphere", *J. geophys. Res.* **93**, 11,022–11,036

Storch, H. von, Weese, U., Xu, J. S., 1990: "Simultaneous Analysis of Space-Time Variability: Principal Oscillation and Principal Interaction Patterns with Applications to the Southern Oscillation", *Z. Meteorol.* **40**, 99–103

Widrow, B., Lehr, M. A., 1990: "30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation", *Proc. IEEE* **78**, 1415–1441

Principal Modes of Variation of Rain Rate Probability Distributions

Thomas L. Bell and R. Suhasini*†

Laboratory for Atmospheres
Goddard Space Flight Center
Greenbelt, MD 20771, U.S.A.

1. Introduction

Areas with rain in them can often be partitioned into regions containing convective or stratiform rain, based on relative intensities in the regions and vertical and horizontal structure of the precipitating region. Houze (1981) has argued for the widespread applicability of these two descriptive categories. There are a number of other instances where it is informative to discuss the distribution of rain rates over an area and how it changes with time. The distribution of rain rates varies considerably with the time of day over many areas, due no doubt to variations in the strength of convective activity (see, for example, Griffin 1990). The success of methods of estimating area-averaged rain rate from the fraction of the area with rain rates above a well-chosen threshold has been explained in terms of choosing the threshold so that the parameter in the estimator is insensitive to variations of the rain rate distribution (Short et al. 1992).

We could describe the changes in the distribution of rain rates with time in terms of the changes in the parameter values of fits to the distributions, if we could find a parametric distribution—perhaps the log-normal or the gamma distribution—that adequately captures the changes in the rain distributions. We shall instead explore a method of extracting from the time-varying histograms of rain rates a set of underlying empirical probability distributions (p.d.'s) whose linear combination is able to describe the changes in the histograms. This approach has proven remarkably successful in capturing much of the variability of the GATE (Global-Atmospheric-Research-Program Atlantic Tropical Experiment) radar-derived rain rate distributions, with just two modes. One mode is dominated by relatively light rain rates, the other by heavier rain rates. It is tempting to label them "stratiform" and "convective," though more detailed analysis will be necessary to see if this association is justified, since these terms carry with them dynamical associations that may not be present in every respect in the modes we isolate.

2. Principal Mode Decomposition

We shall describe the method assuming that we have a sequence of gridded radar maps of rain rate over an area several hundred kilometers in diameter. The method is easily adapted to data sets consisting of sequences of rain gage records or satellite instrument scans covering an area.

The GATE data set we use was derived by Hudlow and Patterson (1979) from radar measurements taken during the summer of 1974. They provide gridded images of rain rates $R(\mathbf{x}, t)$ on a 4 km grid covering

* University Space Research Association Visiting Scientist

† Permanent address: Physical Research Laboratory, Navrangpura, Ahmedabad 380009, India.

an area 400 km in diameter, at roughly 15 min. time intervals. We use data in the circumscribed square area 280 km on a side. Each square therefore contains 4900 grid boxes. Histograms of rain rates $n(r_i, t)$ are obtained, where $n(r_i, t)$ is the count of grid boxes with rain rates falling in bin r_i, with the ith bin delimited by

$$r_i \leq R < r_{i+1} \,. \tag{1}$$

We shall use logarithmically spaced bins here, and exclude zero rain rates, so that

$$N(t) = \sum_i n(r_i, t) \tag{2}$$

will generally be less than 4900.

Guided in part by the phenomenological observation that areas of rain can often be typed as convective or stratiform, we propose to describe the time variation of $n(r_i, t)$ by an expression of the form

$$n(r_i, t) = \sum_\alpha n_\alpha(t) p_\alpha(r_i) \tag{3}$$

where each "mode" p_α corresponds to the p.d. of a type of rain, and n_α is the number of grid boxes with rain rates of type α. Our hope is that if the area is large enough (or long enough time segments are used, in the case of rain gage data), only a few "modes" will be needed to capture the behavior of $n(r_i, t)$.

Constraints on n_α, p_α are based on our physical interpretation of them. We require

$$n_\alpha \geq 0 \,, \tag{4a}$$

$$\sum_\alpha n_\alpha(t) = N(t) \,; \tag{4b}$$

$$p_\alpha(r_i) \geq 0 \,, \quad i = 1, \ldots, I \,, \tag{4c}$$

$$\sum_i p_\alpha(r_i) = 1 \,. \tag{4d}$$

It can be shown that approximate maximum likelihood estimates of the modes p_α can be obtained from the eigenfunctions of the matrix

$$C_{ij} = \frac{1}{\sqrt{p(r_i)}} \Big[\sum_t n(r_i, t) n(r_j, t) / N(t) \Big] \times \frac{1}{\sqrt{p(r_j)}} \,, \tag{5}$$

a kind of weighted covariance of the vector of histogram counts $n(r_i, t)$. The distribution $p(r_i)$ is the climatological p.d. of rain (conditional on $R \geq r_1$),

$$p(r_i) = \sum_t n(r_i, t) / \sum_t N(t) \,.$$

There is not enough space to give the details of the procedure here. They will be described in a forthcoming paper.

3. Principal Modes of GATE rainfall

Histograms of the gridded data from the 280 km square area centered in the GATE region described earlier were decomposed into principal modes, using Phase I data (June 28–July 16). The first two eigenfunctions accounted for 91% of the trace of the matrix (5). They could be readily combined into two distinct modes p_α, $\alpha = 1, 2$. They are shown in Fig. 1. We have labeled them "s" and "c", because their characteristics resemble so closely those ascribed to stratiform and convective rainfall—but it must be emphasized that our analysis deals only with histograms and has lost much of the information about the spatial structure of the radar data.

The mean rain rates associated with each mode,

$$\overline{r}_\alpha = \int_0^\infty r p_\alpha(r) dr \,,$$

are $\overline{r}_s = 2.6$ mm/h and $\overline{r}_c = 8.8$ mm/h. Figure 2 shows a time series of rain amount $[n_\alpha(t)/4900]\overline{r}_\alpha$ due to modes s and c for 4 days of GATE data. We have chosen to relax condition (4a), and allow $n_\alpha < 0$. The extent of this anomaly is a measure of the deviations of the radar rain p.d. from the two mode expansion. In general, 98% of the variance of area-averaged rain rate is explained by just two modes. In GATE Phase I, 65% of the rain volume is accounted for by c-type rain, and 35% by s-type.

4. Conclusions

Variations in the distribution of GATE rain seem to a remarkable extent to be described as variations in the relative amounts of two underlying distributions. These are at least qualitatively similar to the division into stratiform and convective types that studies of the structure of rain-producing systems in GATE and other regions have suggested (Houze and Betts 1981). The analysis of rain distributions in terms of principal modes seems to provide a natural and informative tool for describing the variability of rain distributions with time. They can describe diurnal variability efficiently, provide an alternative route to understanding the success of the area-time-integral (ATI) estimates of area-averaged rain rate, and suggest improvements in remote-sensing algorithms for rain estimation.

REFERENCES

Gibbins, C. J., 1990: Rainfall rates, raindrop size distributions and millimetre-wave attenuations. Proceedings of URSI Commission F, Open Symposium on Regional Factors in Predicting Radiowave Attenuation Due to Rain, Rio de Janeiro, Brazil, 29–34.

Houze, R. A., 1981: Structure of atmospheric precipitation systems: A global survey. *Radio Science*, **16**, 671–689.

Houze, R. A., and A. K. Betts, 1981: Convection in GATE. *Rev. Geophys. Space Res.*, **19**, 541–576.

Hudlow, M. D. and V. L. Patterson, 1979: GATE Radar Rainfall Atlas. NOAA Special Report, available from U.S. Government Printing Office, Washington, DC 20402. 158 pp.

Short, D. A., D. B. Wolff, D. Rosenfeld, and D. Atlas, 1992: A study of the threshold method utilizing rain gauge data. To appear in *J. Appl. Meteor.*

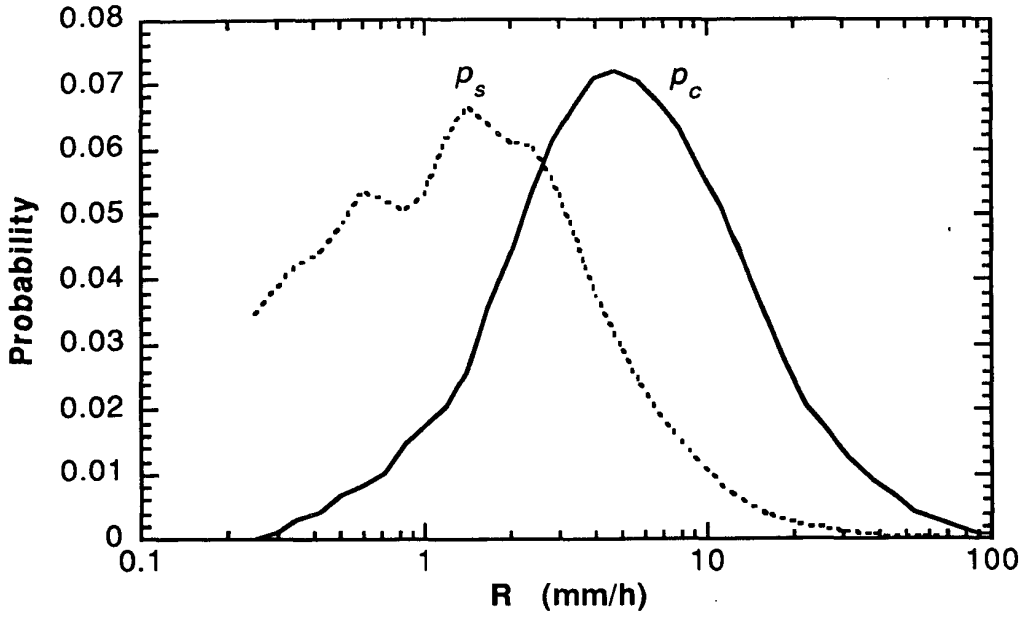


Figure 1. Principal modes of GATE Phase I rain rate distributions, based on 18 days of radar data at 4 km resolution. Histograms of rain rates from a radar image are generally well-described by a linear combination of these two modes.

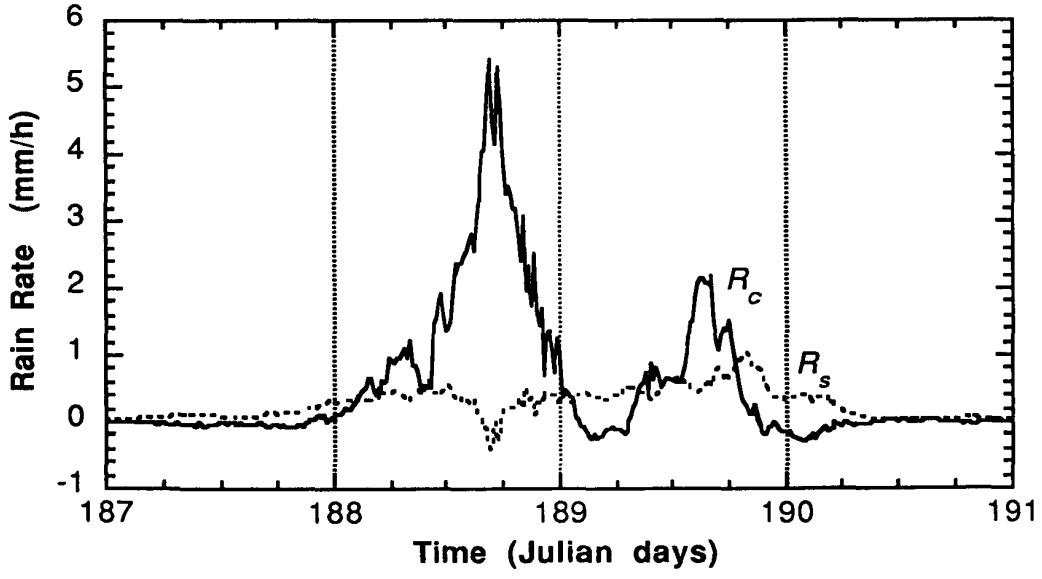


Figure 2. Area-averaged rain rate of two modes "c" and "s" for four days of GATE Phase I. Condition (4a) was relaxed, so that fits can sometimes produce "negative" rain rates.

DETERMINATION OF DIMENSIONALITY IN EIGENANALYSIS

Michael B. Richman
Cooperative Institute for Mesoscale Meteorological Studies
The University of Oklahoma
Norman, OK 73019

James R. Angel
Illinois State Water Survey
Champaign, IL 61820

and

Xiofeng Gong
Weather Bureau of Liaoning Province
Shenyang, People's Republic of China

1. INTRODUCTION

One important feature of eigentechniques such as Principal Component Analysis (PCA) or Empirical Orthogonal Functions (EOF) which has attracted application in the geosciences is the possibility of reducing dimensionality to represent most of a data set's variability. This, in turn, can suggest ways to simplify the data, highlight major features of variability, truncate noise from the analysis, and make subsequent statistical manipulation tractable. A survey of geoscience research using PCA/EOF indicates that the number of applications has been increasing steadily over the past several decades. Moreover, it also indicates that virtually all of the published research use a *reduced subspace* to describe data. Curiously, there is no agreed upon method by which to accomplish this reduction, yet the specific point at which PCs/EOFs are truncated can play an important role in the physical interpretation of these types of methods (North et al., 1982; Richman, 1986).

The goal of this work is to summarize the state of the science with respect to tests or rules proposed to select the most appropriate number of eigenvectors. We will compare each test's ability to separate signal from noise on real data. This is not an easy task as Preisendorfer (1988) mentions that a veritable cottage industry for these rules developed in other sciences and the social sciences over the past forty years. Nevertheless, we attempt to summarize a large number of rules and document their biases under varying conditions. This paper will present a small fraction of our overall research design.

2. DATA

This phase of our research uses a data set which has been heavily scrutinized by several analysis techniques such as common factor analysis, PCA and cluster analysis. Richman et al. (1991) describe the data set which measures daily precipitation totals at 766 locations over the eastern two-thirds of North America for the period of January 1, 1949 through December 31, 1988. The advantages of using these data includes the additional quality control applied and the availability of extensive analyses to provide a guide to the nature of the signal. The daily measurements were totalled to both three and seven-day values. The data were further stratified into the four traditional seasons since the scale and coherence morphology of precipitation is seasonally dependent.

3. METHODOLOGY

3.1 *Data preparation for eignanalysis*

The manner in which a data set is manipulated and specific analysis decisions applied are important to properly executing a PCA or EOF analysis. In this part of the research, the new data (ie., 3-day precipitation totals) were square root transformed to minimize positive skewness, following the extensive discussion in Richman and Lamb (1985, p. 1327-1328). The goal of the analysis is to determine the number of physically interpretable precipitation modes in space over the station network. Accordingly, the data were related in a S-mode analysis, where an interstation correlation matrix was derived and then decomposed into 766 eigenvalues and associated PC loadings as documented in Richman and Lamb (1985;1987). The eigenvalues (Tables 1 lists the first 100 for each season) and the PC loadings for the three-day totals formed the basis for all subsequent testing. These tests were applied to both the unrotated PCs and after a linear transformation using the Varimax or Harris-Kaiser B^TB algorithms (Richman, 1986).

3.2 *Tests and rules used to determine dimensionality*

A mentioned at the onset of this work, there were a large number of available methods proposed in the past forty years literature in fields as diverse as psychology, biology, chemistry and atmospheric sciences. A brief overview of a large subset of those we tested will now be presented.

3.2.1. *Tests which examine eigenvalue magnitude*

a. Eigenvalue 1.0 rule /K-G criterion (Guttman, 1954)

This rule was originally proposed by Guttman as a weakest lower bound to the number of components problem. Guttman concluded that for a Gramian correlation matrix, the dimensionality could by bounded by ordering the eigenvalues in descending magnitude, $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots \geq \lambda_n$ and selection that λ_i slightly larger than or equal to 1.0 as a lower bound. This means that the number of dimensions was thought to be no less than this number. This point is frequently lost by researchers in the scores of applications selecting the dimensionality at that $\lambda_i \leq 1.0$ (Richman, 1988). This is at least partly attributable to Kaiser's (1960) work with 'Little Jiffy' -- a computer program which selected only those components with eigenvalues exceeding 1.0. Other researchers, such as Jolliffe (1972, 1973), have suggested eigenvalue cutoffs at prescribed values other than 1.0.

b. Parallel Analysis Test (Horn, 1965; Montenelli and Humphreys, 1975)/Rule N (Preisendorfer and Barnett, 1977)

The parallel analysis test compare and ordered set of

Table 1. The leading 100 eigenvalues of three day totals of square root transformed precipitation data for each season.

Root Number	Eigenvalues			
	Winter	Spring	Summer	Fall
1	122.337	89.472	52.077	105.017
2	63.636	77.038	46.757	79.950
3	56.741	53.876	34.307	46.711
4	38.160	38.982	26.695	37.929
5	27.484	28.975	23.732	29.370
6	22.893	24.851	19.773	24.686
7	22.435	21.383	17.133	20.413
8	16.965	16.490	15.158	16.880
9	14.677	15.923	13.049	13.851
10	12.601	13.301	11.619	13.195
11	11.615	11.493	10.915	10.950
12	10.776	10.955	9.080	9.661
13	9.280	9.225	8.994	8.784
14	8.363	8.297	8.220	8.397
15	7.417	7.778	7.436	7.696
16	7.012	7.549	6.571	7.180
17	6.273	7.335	6.544	6.929
18	6.093	5.995	6.215	6.716
19	5.750	5.799	5.757	6.281
20	5.551	5.511	5.441	5.493
21	5.234	5.308	5.361	5.107
22	4.985	5.194	5.047	4.803
23	4.625	4.737	4.728	4.701
24	4.316	4.502	4.525	4.494
25	4.196	4.362	4.272	4.276
26	3.780	4.133	4.151	3.873
27	3.514	3.836	4.115	3.753
28	3.372	3.527	3.840	3.712
29	3.280	3.486	3.737	3.449
30	3.142	3.236	3.640	3.255
31	3.065	3.127	3.583	3.030
32	2.936	3.079	3.456	2.985
33	2.832	2.928	3.422	2.868
34	2.593	2.796	3.274	2.804
35	2.526	2.728	3.190	2.692
36	2.430	2.546	3.020	2.593
37	2.264	2.542	2.979	2.501
38	2.224	2.482	2.861	2.401
39	2.195	2.372	2.792	2.356
40	2.090	2.310	2.718	2.299
41	2.068	2.190	2.710	2.270
42	1.970	2.188	2.665	2.229
43	1.938	2.124	2.632	2.144
44	1.893	2.055	2.504	2.132
45	1.835	1.973	2.467	2.003
46	1.769	1.942	2.380	1.916
47	1.750	1.893	2.360	1.906
48	1.732	1.863	2.324	1.893
49	1.709	1.815	2.259	1.849
50	1.681	1.770	2.232	1.813
51	1.620	1.764	2.186	1.767
52	1.604	1.713	2.159	1.699
53	1.544	1.676	2.133	1.653
54	1.505	1.647	2.066	1.632
55	1.503	1.591	2.057	1.616
56	1.464	1.563	2.033	1.598
57	1.439	1.544	2.013	1.558
58	1.432	1.521	1.971	1.525
59	1.389	1.515	1.922	1.477
60	1.372	1.478	1.877	1.464
61	1.366	1.443	1.865	1.421
62	1.334	1.413	1.847	1.404
63	1.320	1.352	1.833	1.381
64	1.298	1.335	1.811	1.361
65	1.291	1.325	1.777	1.350
66	1.236	1.310	1.757	1.335
67	1.233	1.292	1.726	1.297
68	1.218	1.286	1.709	1.273
69	1.204	1.273	1.700	1.259
70	1.189	1.232	1.677	1.229
71	1.177	1.228	1.661	1.219
72	1.167	1.205	1.645	1.209
73	1.162	1.188	1.619	1.197
74	1.129	1.182	1.607	1.187
75	1.107	1.157	1.591	1.176
76	1.091	1.131	1.563	1.146
77	1.081	1.125	1.555	1.139
78	1.067	1.116	1.551	1.133
79	1.051	1.096	1.536	1.115
80	1.047	1.085	1.517	1.107
81	1.032	1.082	1.493	1.081
82	1.016	1.077	1.477	1.072
83	1.014	1.067	1.470	1.061
84	1.001	1.053	1.460	1.046
85	0.992	1.044	1.451	1.038
86	0.980	1.021	1.427	1.027
87	0.973	1.018	1.426	1.018
88	0.956	1.012	1.407	1.008
89	0.953	0.990	1.388	1.001
90	0.942	0.983	1.377	0.997
91	0.941	0.978	1.371	0.979
92	0.924	0.967	1.359	0.970
93	0.917	0.959	1.355	0.967
94	0.911	0.952	1.354	0.951
95	0.898	0.941	1.331	0.942
96	0.894	0.934	1.325	0.934
97	0.889	0.929	1.310	0.924
98	0.884	0.919	1.303	0.918
99	0.867	0.910	1.287	0.911
100	0.862	0.901	1.280	0.900

eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots \geq \lambda_n$ from a Wishart distribution (random normal deviates) to those calculated from a data set. At the location i where the random (R_i) eigenvalue exceeds the data eigenvalue (V_i) no further eigenvectors are retained. Rule N is based on a similar principle but adds a confidence bandwidth at σ_i (95) and σ_i (05). Preisendorfer (1988, p. 199-201) summarizes Rule N as selecting the number of dimensions, p' as the greatest i for which the data eigenvalue $V_i > \sigma_i$ (95) as long as at least one eigenvalue exceeds σ_i (95). Zwick and Velicer (1986) report that the parallel analysis test was the most accurate of five methods applied to psychometric literature.

c. PVP test (Guiot, 1981)

This rule of thumb was introduced in Guiot's PhD thesis. Briffa et al. (1983) reported on its use for PC regression work. The rule states that for an ordered set of eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots \geq \lambda n$, the λ_i for which cumulative product of the eigenvalues drops below 1.0 is where no further PCs are retained. Briffa et al. mention the PVP test tends to eliminate about one third of the total number of dimensions. This test has not gained wide acceptance in the literature.

d. Condition Number (Jochum et al., 1981)

The use of the condition number, which tracks error propagation was suggested by Jochum et al. (1981) as test for dimensionality. Originally, condition number was created as a tool in numerical analysis to estimate the probability of inverting a matrix. For a square matrix, Jochum et al. demonstrate

$$\mathrm{COND}(A) = \frac{\max\lambda}{\min\lambda} = \frac{\lambda_1}{\lambda_i}$$

The values of COND (A) is proportional to the error and small as long as the eigenvalues are different from zero. Once λ_i becomes close to zero, COND (A) grows quickly. This property has been exploited in the chemistry literature, where plots of COND (A) versus root number are examined for a sharp discontinuity at the truncation point.

3.2.2 *Tests based upon eigenvalue spacing*

a. The Scree test (Cattell, 1966)/LEV test (Craddock and Flood, 1969)

Despite some controversy surrounding its origin, Cattell (1966) is credited with devising this test which plots up the ordered root or component number against the proportion of variance it extracts. A curve fitted to such a plot generally exhibits a decreasing negative slope until random errors are reached, where it levels off. The LEV test is based on the same premise and plots logarithmic eigenvalue (LEV) against ordered root number. The key to effectively utilizing the scree test is to locate the point on the curve where the tail levels off. Problems have been reported with subjective interpretation and multiple points where the curve levels-off in "steps" (Rummel, 1970, p. 364). In the case of multiple points, Cattell (1966) recommended choosing the last step. O'Lenic and Livezey (1988) discuss the implications of choosing various steps for filtering the 700 mb level.

b. Eigenvalue separation criterion (Kendall, 1980; North et al., 1982)

The procedure was developed to help insure that adjacent ordered eigenvalues were not too "close" in magnitude. Kendall (1980) discusses some computational instabilities which may

arise under conditions of nearly equal eigenvalues. North et al., (1982) used Girschick's (1939, p. 217) work, which shows that the standard deviation of the eigenvalues for large samples are of the form

$$\delta\lambda_i = \lambda_i \left(\frac{2}{N}\right)^{1/2} \quad i = 1,...,n$$

to arrive at their conclusion that if the sampling error $\delta\lambda_i$ of the ith population eigenvalue λ_i is comparable in size to the spacing between λ_i and a neighboring eigenvalue λ_k sampling error of the ith eigenvalue, V_i, will be comparable to that of the nearby V_k. North et a. (1982) point out this is a form of degeneracy which can result in determining of sample eigenvector patterns. The test has gained increased popularity in recent years to reject sections or ordered eigenvalues which do not pass the $\delta\lambda_i$ sampling error test.

c. Alex and Savoie Test 1 (Alex and Savoie, 1989)

This test was created to sharpen the gap between the last significant eigenvalue and the first non-zero eigenvalue. It has seen limited use in the field of chemistry. Dimensionality can be deduced from a plot of root number versus test value as the authors state (p. 30) "this pseudo-function exhibits a maximum when the index i becomes equal to the value of the rank of m". The function is defined as:

$$f_1(i) = \left(\frac{\lambda_i - \lambda_{i+1}}{\lambda_{i+1} - \lambda_{i+2}}\right) .$$

d. Alex and Savoie Test 2 (Alex and Savoie, 1989)

Not fully pleased with the results of f_1, Alex and Savoie (1989, p. 30) created a variation of it "as the normalization and square root operator improve the definition of the extracted minimum". As in f_1, the rank was thought to be equal to the minimum value of f_2. The function is defined as:

$$f_2(i) = \frac{\sqrt{\tau_i} - \sqrt{\tau_{i+1}}}{\sqrt{\tau_{i+1}} - \sqrt{\tau_{i+2}}}$$

where

$$\tau_i = \frac{\lambda_i}{\sum_{i=1}^{n} \lambda_i}$$

3.2.3 *Model fit tests*

a) Minimum Average Partial (MAP) test (Velicer, 1976)

Velicer (1976) proposed a test based upon calculating the average of the squared partial correlations. When the average squared partial correlation reaches a minimum, the number of eigenvectors partialled out is considered the proper number for retention. The MAP formula is given by:

$$F_k = \sum_{i=1}^{n} \sum_{j=1}^{n} (z_{ij}^2)/n(n-1)$$

where z_{ij} is an element of the matrix of partial correlations after the kth eigenvector/component is removed. The plot of F typically shows a steep drop followed by a broad valley and a subsequent rise. Zwick and Velicer (1982) report that the test was quite accurate when compared with the KG test and the scree test.

b) The Indicator Function (IND) test (Malinowski, 1977)

Malinowski (1977) put forth an empirical function, IND, which can be claimed (p. 614) "...can be used to deduce to dimensionality of the factor space". It has received considerable use in the chemistry literature. The function is defined as:

$$IND = \left[\frac{\sum_{i=k+1}^{n} (\lambda_i)}{c(n-k)}\right]^{1/2} / (n-k)^2$$

where c is the number of columns in the PC loading matrix, n the number of rows in the PC loading matrix, k the number of eigenvectors retained and λ_i is the ith eigenvalue. Malinowski (1977, p. 613) claims the quantity in the square brackets is representative of the "real error" between the represented and original data. A plot of IND against root number usually exhibits a minimum when it is thought to coincide with the correct number of non-zero eigenvalues, though some situations were noted when IND failed to yield a single unambiguous minimum.

c) The Exner Function (Exner, 1966)

Another test arising from the chemistry literature is Exner's empirical function (Exner, 1966, p. 3227). Exner's equation compares the mean square value difference between the data and modeled values. It is given by:

$$EXNER = \left[\frac{\sum_{i=1}^{n} \sum_{k=1}^{c} (d_{ik} - d_{ik}^R)^2}{\sum_{i=1}^{n} \sum_{k=1}^{c} (d_{ik} - \bar{d})^2} \times \frac{nc}{nc-k}\right]^{1/2}$$

where i indexes the number of rows, k indexes the number of columns in the PC loading matrix, k represents the number of PCs retained, d_{ik}^R is the PC loading for the reproduced covariance matrix, d_{ik} the PC loading matrix. The function exhibits a minimum at the location thought to represent the correct dimensionality.

d) Minimization of the mean absolute difference between the eigenvector/PC loading pattern and the dispersion matrix pattern (Richman and Lamb, 1987)

This test, introduced by Richman and Lamb (1987, Appendix) applies a different rationale to previous tests for selecting the optimum number of eigenvectors/PCs to retain. The idea is to minimize the distance between the dispersion matrix (covariance or correlation) and the eigenvector/PC patterns. The PC loading with the largest absolute magnitude is identified b_j for each column i of the matrix. The jth row of the correlation (or covariance) matrix is then compared to the appropriate PC loading vector. Richman and Lamb (1987) use the following formulation to select a stopping point at the minimum value of d_{AB}:

$$d_1AB = \sum_{i=1}^{M} \left[\sum_{j=1}^{N} (|a_j - b_j|) / N\right] / M$$

where A is a PC loading vector with individual variable values a, B is the interstation correlation vector with individual variable values b for the variable with the highest absolute loading for a given PC. N indexes variables and M indexes PCs retained.

e) Minimization of the root mean squared difference between the eigenvector/PC loading pattern and the dispersion matrix pattern

This test is identical to the proceeding one except the test statistic d is computed for root mean square:

$$d_2AB = \sum_{i=1}^{M}\left[\sum_{j=1}^{N}(a_j^2 - b_j^2)^{1/2} / N\right] / M$$

f) Minimum largest PC loading test

Generally, only those PC loading values in excess of some arbitrary value are used for physical interpretation. If one orders the eigenvectors with associated eigenvalues $\lambda \geq \lambda_2 \lambda_3 \geq \ldots \geq \lambda_n$, the decreasing variance is directly related to decreasing PC loading values. This test takes advantage of this by charting the largest absolute value of the PC loadings b_{ij} tends to be related to fragmentation of coherent variables as shown by Lamb and Richman (1987, Appendix). Such a discontinuity can serve an indicator of a solution to reject.

g) Number of PC loadings in excess of |0.40|

The motivation for this test was to the previous one. The number of occurrence of PC loadings $b_{ij} \geq |0.40|$ were calculated for each successive truncation point. The PC with the fewest values $\geq |0.40|$ had the number of values noted and graphed. Such a plot tends to indicate a sharp decline in an area closet to where one might wish not to retain additional PCs. Furthermore, if the value drops below 3, the hyperplane position will not be stable (Richman, 1986) and the PC subspace will not be reliable. Hence, this test if a useful bound for the maximum number of stable PCs.

h) Minimum correlation on worst matching PC

This is another proposal test which examined the correlation between the PC vector and the corresponding base correlation (which is chosen at the variable with the highest PC loading for each PC). The one worst correlation match per solution is recorded and plotted. A sharp decrease in the minimum correlation might indicate an appropriate spot to stop retaining further PCs.

3.2.4 Consistency

This was a subjective test where each PC loading vector was plotted and isoplethed on base map. As each additional PC is admitted to a rotated solution, the investigator searched for abrupt spatial changes. Typically, large coherent regions would subdivide into smaller coherent regions in a predictable fashion (see Richman and Lamb, 1987, Appendix, for examples) until such a point where the spatial distribution of the loadings would fragment into numerous smaller areas. The point at which this occurred was where the PC solution was rejected. The goal was to admit as many coherent PCs as possible. This was performed independently for orthogonal and oblique solutions.

4. RESULTS

The consistency test was used as the arbiter of success for this data set. Numerical tests which differ markedly from the "selected value" are associated with either spatial degradation or fragmentation. This will be noted for each test's bias (Table 2). The eigenvalue tests, such as KG, PVP parallel analysis and condition number all had a bias to overestimate the number of physically interpretable signals. The condition number and PVP test were the most flagrant in this regard, being over an order of magnitude too generous. The parallel analysis/Rule N tests were the best of the group, overestimating by about a factor of two (Fig. 1).

Tests based on the spacing between adjacent eigenvalue magnitudes were not uniform in their truncation bias. The two tests presented by Alex and Savoie (1989) usually overestimated the number of interpretable signals by at least a factor of two (Fig. 2). However, in summer, when the leading eigenvalues had a lower magnitude (Table 1), they were very close to the chosen

Table 2. Results of various dimensionality tests in separating precipitation signal from noise for each season.

Test/Rule	Number of Principal Components Retained			
	Winter	Spring	Summer	Fall
Eigenvalue 1.0/KG	84	88	142	89
Parallel Analysis/Rule N	33	35	43	35
PVP Test	247	257	366	254
Condition Number	329	377	383	311
Scree/LEV Test	33	35	43	30
Eigenvalue Separation Test	(12,13,14, 16,25,33)	(10,12,13, 17,22,27)	(11,13, 14,15)	(10,11,12, 14,19,25)
Alex and Savoie Test 1	53	40	15	45
Alex and Sovoie Test 2	65	35	15	45
MAP Test	30	30	30	31
Indicator Function	60	61	46	64
Exner Function	7	7	8	8
Mean Absolute Difference Test	9	8	21	7
Root Mean Square Difference Test	9	8	21	7
Minimum Largest PC Loading Test	24	21	16	22
Number of Loadings $\geq$ |0.40| Test	24	21	16	22
Minimum Correlation of Worst Matching PC Test	22	21	16	24
Consistency	17	17	16	18

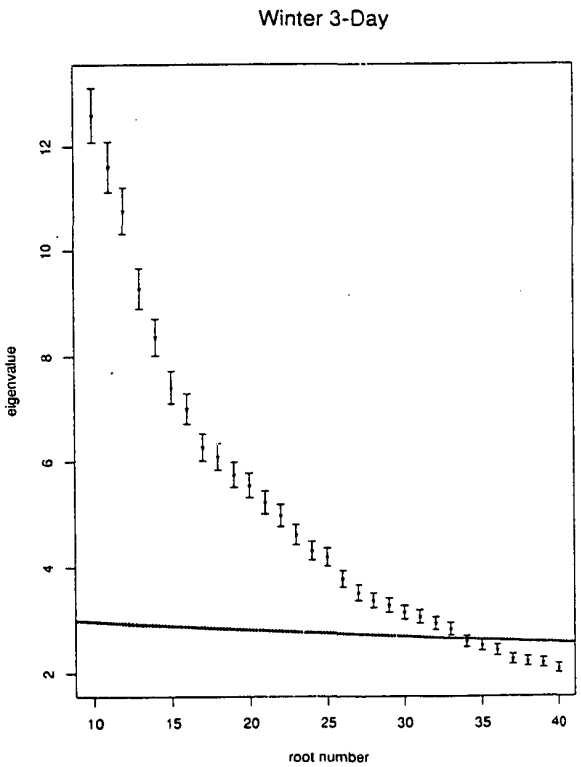


Figure 1. Combined plot of scree test, Rule N and eigenvalue separation test for winter 3 day precipitation analysis.

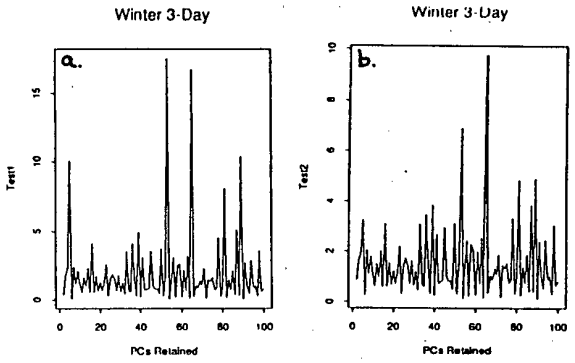


Figure 2. Plots of Alex and Savoie (1989) truncation tests for winter analysis. Panel a is for Test 1, panel b is for Test 2.

number. It is unknown whether this is related to the power of the test or just a chance finding. Given the poor performance in other seasons and the multiple large spikes in the values (Fig. 2), we can not recommend either test for general application. The eigenvalue separation test is more problematic to judge as it doesn't tell one where to truncate, rather it tests for closely spaced values where the conventional wisdom is that one should not break up a degenerate multiplet. The results shown in Table 2 and Fig. 1 indicate those roots which do not overlap with the next root between ordered eigenvalues 10 and 40 (the largest reasonable range). One could take the last break larger than the standard errors as a truncation point. In such a case, the test would overestimate the number of physically significant eigenvalues in winter, spring and fall. The summer results are quite accurate, reflecting the lower leading eigenvalues for that season. Overall, the test did not yield the desired results in most cases. This will be elaborated upon below. The scree test also overestimated the selected number of PCs in all seasons by a factor of close to two. The main problem with using the scree test on a real data set with a large number of dimensions is the multiple "shelves" or points where on could declare the appropriate spot to stop drawing new dimensions (Fig. 1). Deciding which of these spots is "the correct one" is subjective (O'Lenic and Livezey, 1988) -- which would lower the reliability of the test.

The final group of tests are loosely called model fit tests. Using different techniques, they all try to somehow relate the PC loadings either to the correlation/covariance dispersion matrix or to see how the modeled error increases. The MAP test overestimated the dimensionality by a factor of close to two (Fig. 3). It was amazingly uniform in assigning ~30 as the correct number of components for every season. The IND test uses the remaining n-r eigenvalues as a measure of error. The denominator is created in such a manner as to create a penalty for retaining many PCs. Despite this the IND test overestimated the dimensionality by a large number (about a factor of three) (Fig. 4). Conversely, the other fit test to come out of the chemistry literature, the EXNER function, tested the PC loadings directly. This underestimated the dimensionality in all seasons by approximately a factor to two (Fig. 5).

The remaining model fit tests are either motivated by Richman and Lamb (1987) or are new ones. The mean absolute difference test is identical to that shown in Richman and Lamb (1987) for their Canadian precipitation study. Both the mean absolute difference and RMS difference tests yielded identical results. Both retained too few PCs, keying in on the largest patterns for winter, spring, and fall (Fig. 6 and 7). In summer, when the patterns were all small (similar in spatial scale) these tests slightly overestimated the number of PCs. The largest PC loading minimum test and number of loadings ≥ |0.40| on the smallest PC yielded the same answers (Figs. 8 and 9). These slightly overestimated the optimal dimensionality in winter, spring, and fall and correctly picked the correct value in summer.

Finally, the minimum correlation on the worst matching PC was in the same range as the previous two tests -- it slightly overestimated winter, spring and fall and correctly estimated summer (Fig. 10).

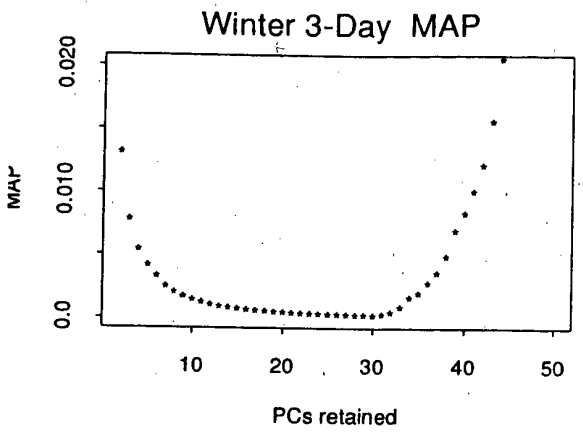


Figure 3. Plot of Velicer's (1976) MAP test for winter analysis.

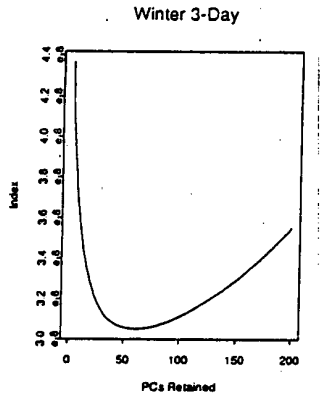


Figure 4. Plot of Malinowski's (1977) Indicatitor Function for winter analysis.

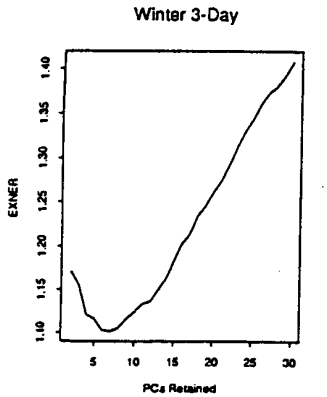


Figure 5. Plot of Exner's (1966) Function for winter analysis.

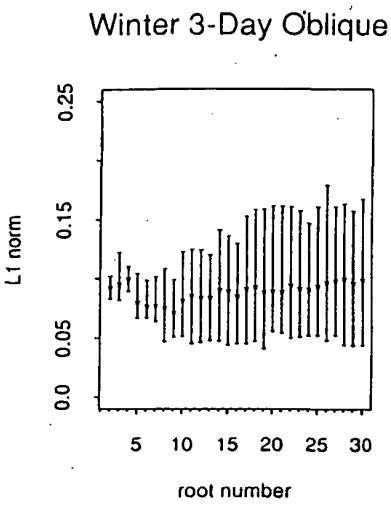


Figure 6. Plot of Richman and Lamb's (1987) Test for winter analysis.

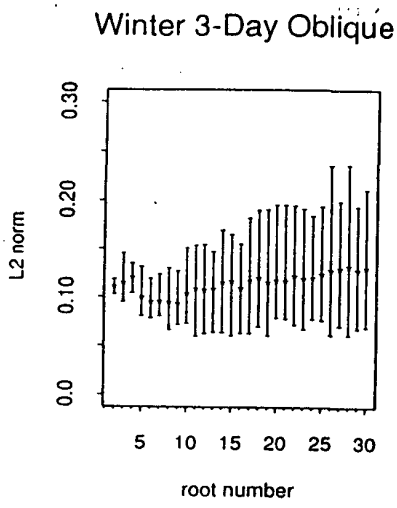


Figure 7. Plot of RMS Difference Test for winter analysis.

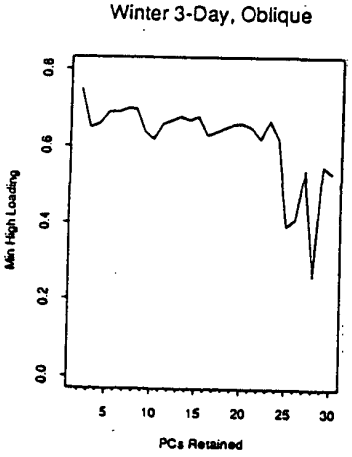


Figure 8. Plot of Minimum Largest PC Loading Test for winter analysis.

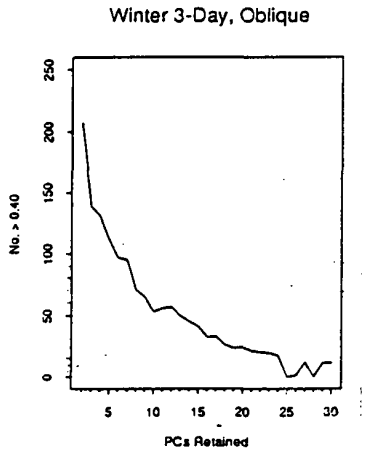


Figure 9. Plot of number of variables ≥|0.40| on the PC with the smallest absolute maximum loading for winter analysis.

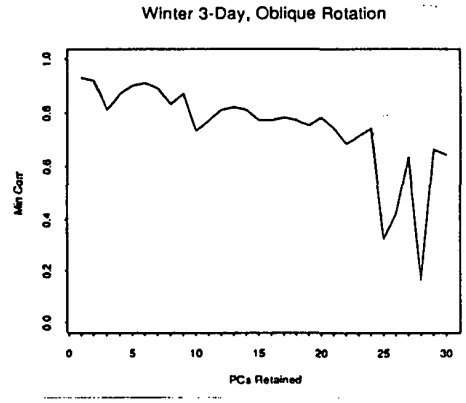


Figure 10. Plot of correlation between worst matching PC loading vector and input correlations for winter analysis.

5. CONCLUSIONS

Some conclusions were very easy to draw, whereas others were more difficult at this phase of the full research design. What this portion of the research indicates is that for a large (1200 x 766) data set with mixed scales of variability, the traditional tests which are often used in the geosciences, fared poorly. For example, the most popular rule, the KG test, severely overestimated the number of dimensions with interpretable signal. Combined with studies such as Zwick and Velicer (1982, 1986) we recommend that any stopping rule with a fixed eigenvalue such as 1.0 or Jolliffe's (1972, p. 171) suggestion of 0.7 be avoided as these have been found to be biased over a broad range of conditions. The second most popular rule, the scree test, was somewhat more accurate but biased high when the last scree was sought. The subjectivity inherent in its application is also problematic -- there is a non-trivial probability that five scientists would find five different screes with the same data tested. The test outlined in North et al. had some usefulness. For those investigators using unrotated PCs, they should insure not to stop retaining roots inside a degenerate multiplet. The final root to clear the standard error criterion was also biased high in three of four seasons. However, an important finding was that we found coherent rotated modes showing no signs of intermixing for those roots where unrotated degenerate multiplets existed.

This combined with questionable results of the eigenvalue magnitude based tests lead us to our first major conclusion: *in the absence of a dynamical basis, eigenvalue tests are not sufficient to reliably determine the dimensionality of the real data PC space, particularly when the PCs are to be linearly transformed.*

The model fit tests fell into two regimes, those that were not found accurate for this data set (ie., MAP, IND, EXNER, d_1, d_2) and those that held promise (largest PC loading minimum, number of PC loadings $\geq|0.40|$ and the minimum correlation on the worst matching PC). These last ones seemed to choose the truncation point within 7 roots of the optimum number. Admittedly, the present data set is a rigorous test for the diagnostic procedures tested herein to correctly distinguish signal from noise due to the large number of variables, the mixed scales and the noise level. However, it is at the low end of the number of variables, observations, and spatial extent which will be used for high resolution GCMs in the near future. Such results suggest that the mixed scale lead to a situation analogous to that outlined by Tucker et al. (1969) where the modes of variation can be subdivided into the more reliable major modes and the more tenuous signal known as minor modes. Our visual inspection of thousands of PC loading maps further promoted this view as the larger scale precipitation modes (major ones) were more stable.

In other experiments (not shown) the various truncation tests performed much better. If there was pure signal, all tests worked adequately. Unfortunately, read world data set do not tend to be obliging in that respect.

A final note regards our subjective consistency test for orthogonal versus obliquely rotated PCs. We were able to consistently retain several additional oblique patterns prior to fragmentation, due to relaxation of the orthogonality constraint. It therefore appears that rotation choice may be partly a function of the number of PCs retained. Fitting 15 or more orthogonal dimensions to a non-orthogonal PC subspace was found to be difficult and led to inferior matches. Results presented in Gong and Richman (1992) for minimum matches on orthogoanal versus oblique PCs appear to confirm this finding. *We therefore conclude that the use of any of these tests or rules of thumb be tailored to the specific solution(s) being considered.* What this implies is that if one plans to perform a Varimax transformation, several such solutions should be computed, varying the number of PCs retained and tests should then be applied to the rotated variance retained (rather than the traditional use of the unrotated variance) or to the loading distributions. The comparisons presented in Table 2 offer a guide to selecting some of the more reasonable algorithms.

ACKNOWLEDGEMENTS

This research was partly supported by NSF Grant ATM 89-08545 and USEPA Cooperative Agreement No. CR-816318-02-0. Some of the computations were performed on the National Center for Supercomputing Applications CRAY2. We wish to thank Debbie Rush and Ginger Rollins for typing the manuscript.

REFERENCES

Alex, S. and R. Savoie, 1989: Evaluation of nonzero eigenvalues in factor analysis. *Canadian J. Spectroscopy, 34,* 27-33.

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 from England. *J. Climatology, 3,* 233-242.

Cattell, R.B., 1966: The scree test for the number of factors. *Multivariate Behav. Res., 1,* 245-276.

Craddock, J.R. and C.R. Flood, 1969: Eigenvectors for representing the 500 mb geopotential surface over the Northern Hemisphere. *Quart. J. Royal Meteor. Soc., 95,* 576-593.

Exner, O., 1966: Additive physical properties. I. General relationship and problems of statistical nature. *Collection Czechoslov. Chem. Commun., 31,* 3222-3251.

Girschick, M.A., 1939: On the sampling theory of roots of determinantal equations. *Ann. Math. Statist., 10,* 201-224.

Gong, X-F and M.B. Richman, 1992: An examination of cluster analysis methodology in geophysical research. Part I: precipitation data. Submitted to *J. Climate.*

Guiot, J., 1981: Analyse Methematique de Donnees Geophysiques, Applications a la Dendroclimatologie. Unpublished Ph.D. Dissertation, Universite' Catholique da Louvain, Faculte des Sciences, Louvain-la-Neuve, Belgium.

Guttman, L., 1954: Some necessary conditions for common factor analysis. *Psychometrika, 19,* 149-162.

Horn, J.L., 1965: A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179-185.

Jochum, C., P. Jochum and B. Kowalski, 1981: Error propagation and optimal performance in multicomponent analysis. *Analyt. Chem., 53,* 85-92.

Jolliffe, I.T., 1972: Discarding variables in a principal component analysis. Part I: Artificial data. *J. Roy. Statist. Soc., C22,* 160-173.

Jolliffe, I.T., 1973: Discarding variables in a principal component analysis. Part II: Real data, *J. Roy. Statist. Soc., C23,* 21-31.

Kaiser, H.F., 1960: The application of electronic computers to factor analysis. *Educ. and Psychol. Measurement, 20,* 141-151.

Kendall, M., 1980: *Multivariate Analysis.* Charles Griffen, High Wycombe, 211pp.

Malinowski, E.R., 1977: Determination of the number of factors and the experimental error in a data matrix. *Analyt. Chem., 49,* 612-617.

Montenelli, R.G. and L.G. Humphreys, 1976: Latent roots of random data correlation matrices with squared multiple correlations on the diagonal: A Monte Carlo Study. *Psychometrika, 41,* 341-348.

North, G.R., 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.

O'Lenic, E.A. and R.E. Livezey, 1988: Practical considerations in the use of rotated principal component analysis (RPCA) in diagnostic studies of upper-air height fields. *Mon. Wea. Rev., 116,* 1682-1689.

Preisendorfer, R.W., 1988: *Principal Component Analysis in Meteorology and Oceanography.* Elsevier, New York, 425pp.

Preisendorfer, R.W. and T.P. Barnett, 1977: Significance tests for empirical orthogonal functions. *Fifth Conf. on Probability and Statistics in Atmos. Sci.,* Las Vegas, NV. American Meteorol. Soc., 169-172.

Richman, M.B., 1988: A cautionary note concerning a commonly applied eigenanalysis procedure. *Tellus, 40B,* 50-58.

Richman, M.B., 1986: Rotation of principal components. *J. Climatology, 6,* 293-335.

Richman M. B. and P.J. Lamb, 1987: Pattern analysis of growing season precipitation in southern Canada. *Atmosphere-Ocean, 25,* 137-158.

Richman, M.B. and P.J. Lamb, 1985: Climate pattern analysis of three-and seven day summer rainfall in the central United States: Some methodological considerations and a regionalization. *J. Climate Applied Meteor.,* 24, 1325-1343.

Richman, M.B. P.J. Lamb and J.R. Angel, 1991: Relationships between monthly precipitation over central and eastern North America and the Southern Oscillation. *Proceedings of the Fifteenth Annual Climate Diagnostic Workshop,* U.S. Dept. of Commerce, Washington, D.C., 373-383.

Rummel, R.J., 1970: *Applied Factor Analysis,* Northwestern Univ. Press, Evanston, 617pp.

Tucker, L.R., R.F. Koopman and R.L. Linn, 1969: Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika, 34,* 421-459.

Velicer, W.F., 1976: Determining the number of components from partial correlations. *Psychometrika, 41,* 321-327.

Zwick, W.R. and W.F. Velicer, 1986: Comparison of five rules for determining the number of components to retain. *Psychol. Bull., 99,* 432-442.

Zwick, W.R. and W.F. Velicer, 1982: Factors influencing four rules for determining the number of components to retain. *Multivariate Behav. Res., 17,* 253-269.

ON THE INTERPRETATION OF EXTENDED EMPIRICAL ORTHOGONAL FUNCTION ANALYSIS

J.-M. Chen and P. A. Harr

Naval Postgraduate School
Monterey, California

1. Introdution

The capability of including heterorgenous data in an Empirical Orthogonal Function (EOF) analysis greatly enhances the use of EOF analysis as a data reduction technique. But extra care must be taken when physically interpreting EOFs in such cases. The interpretation of relationships between substructures within one EOF may be misleading.

Two types of heterogeneous data have been used in EOF analysis. The first type contains different field variables and has been labeled EOF complexes (Kutzbach 1967) or multi-variate EOF (MV-EOF) (Wang 1992). In MV-EOF analysis, patterns or substructures of different variables within one EOF are interpreted as the interrelated patterns between those variables. The second type contains the same variable but at different times. This has been referred as extended EOF (EEOF) analysis (Weare and Nasstrom 1982). In EEOF analysis, successive patterns or substructures of the same variable within one EEOF are interpreted as the propagation or evolution in time of the first substructure of the function. These interpretations may be misleading since the mathematical derivation of EOF analysis is based on the optimization of the variance of each EOF as an entity and has no reference to the correlations among partial structures within one EOF.

2. Limitation on EEOF Interpretation

The limitation of EEOF interpretation on correlations between substructures within EEOFs can be explained with the following schematic example.

Define X and Y as the variables, S and W as two principal components (PCs) obtained from X and Y, and a and b are loadings (EOFs). With the orthonormality in space between EOFs, the relationships between S, W, X and Y can be written as

$$S = aX + bY \tag{1}$$

$$W = bX - aY \tag{2}$$

with $a^2+b^2 = 1$. By definition, there is no correlation in time between S and W

$$a^2\overline{XY} - ab(\overline{X^2}-\overline{Y^2}) - b^2\overline{XY} = 0 \tag{3}$$

here an overbar represents an average over time. Definition of a loading partition ratio, $r = a/b$, and a normalization factor, $m = 1+r^2$ allows

$$S = m^{-1/2}(rX+Y) \tag{4}$$

$$W = m^{-1/2}(X-rY) \tag{5}$$

$$r^2(\overline{XY}) - r(\overline{X^2}-\overline{Y^2}) - (\overline{XY}) = 0 \tag{6}$$

or

$$r^2cq - r(q^2-1) - cq = 0 \tag{7}$$

Here,

$$q \equiv (\overline{X^2}/\overline{Y^2})^{1/2} \tag{8}$$

is a variation ratio of X to Y, and c

$$c \equiv \frac{\overline{XY}}{[(\overline{X^2})(\overline{Y^2})]^{1/2}} \tag{9}$$

is the time correlation coefficient between X and Y. The solution, r (loading partition), versus the correlation coefficient, c, and the

variation ratio, q, is presented in Fig. 1a and 1b.

Comparison of these two figures indicates that the partition of the loadings is much more sensitive to the variation ratio of X and Y than to the correlation between X and Y for reasonable values of c (say, $c > 40\%$). The loading partition only shows excessive sensitivity to c for values less than 20%. Unfortunately, these are cases that are not desirable when a researcher is investigating relationships between fields.

It is interesting to note from Fig.1b that for variation ratio near 1, the loading patition ratio exhibits large variation when correlation coefficient is small (say $< 10\%$). This represents the problems caused by sampling errors in EOF analysis as defined by North et al. (1982). Ideally, EOFs are obtained with no correlation between any two PCs, but sampling errors may induce small correlations between PCs. When the variances of the PCs are similar, new components formed as linear combinations of the original are highly sensitive to the induced correlation.

The mathematics for a two-field (heterogeneous data matrix) case is similar to the two-variable case. Eq. (3) can now be written as

$$\overline{(\vec{A}\vec{X})(\vec{A}\vec{Y})} - [\overline{(\vec{A}\vec{X})(\vec{B}\vec{X})} - \overline{(\vec{A}\vec{Y})(\vec{B}\vec{Y})}] - \overline{(\vec{B}\vec{X})(\vec{B}\vec{Y})} = 0 \qquad (10)$$

In this case, the correlation coefficient, c, can vary upon different substructures within the loadings $\vec{A}$, $\vec{B}$. As in the two-variable case,

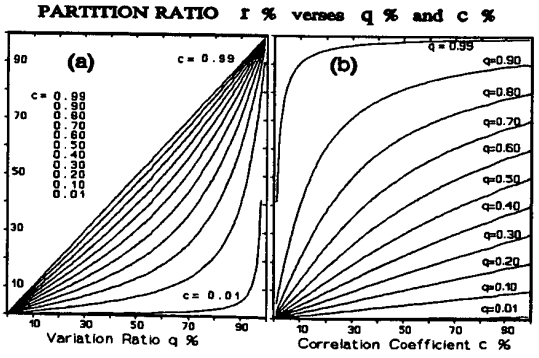


Figure 1. Partition ratio verses (a) the variation ratio and (b) the correlation coefficient.

the solution for the loading partition r exhibits greater sensitivity to the variation ratio, q. Therefore, the resulting EEOFs represent maximum variance solutions with little regard to correlations between substructures within individual functions. It is possible that significant correlation does exist between substructures when the spatially most significant features are also the most persistent in time. In such a case, similar conclusions could also be drawn from examining correlations between PCs obtained from a regular EOF analysis.

Finally, a substructure in one EEOF is not uncorrelated with its counterpart or with other substructures in other EEOFs. There is not a unique solution for projection of a variable field onto the substructure hyper-space. Correlations between substructures and variance partitions of each substructure cannot be computed directly. Therefore, inclusion of heterogenous data in EOF analysis can make physical interpretation of the resulting functions very difficult.

3. Example

An EEOF analysis is performed on a hyperthetical test dataset to show possible misleading interpretation of correlations between substructures in EEOFs. An EEOF analysis usually is applied using an augmented dispersion matrix with one component, $\vec{X}$, presenting the data field at a base time and a second component, $\vec{Y}$, representing the data field lagged in time with the base time. A test dataset is constructed as

$$\begin{aligned} X &\equiv (X_1, X_2, X_3, X_4, X_5, X_6) * [E] \\ &= (1.0P_1+0.4P_2,\ 0.5P_3+0.5P_4, \\ &\qquad 0.2P_5+0.5P_6,\ 0, 0, 0) \\ &\quad * (E_1, E_3, E_5, E_2, E_4, E_6 \end{aligned} \qquad (11)$$

$$\begin{aligned} \vec{Y} &\equiv (Y_1, Y_2, Y_3, Y_4, Y_5, Y_6) * [E] \\ &= (0, 0, 0, 1.0P_2, 0.5P_4, 0.5P_6) \\ &\quad * (E_1, E_3, E_5, E_2, E_4, E_6) \end{aligned} \qquad (12)$$

here P_i and E_i are PCs and EOFs derived from 700 mb zonal wind anomalies between June-

October 1979. Creation of the augmented matrix in this manner allows direct assessment of the interrelationships between EEOF substructures. Table 1 shows the correlation coefficients between substructures X_i and Y_j. This is only possible because we know the exact structure of each data set, which is never known in practice.

Figure 2 shows the substructures (*X*- and *Y*-) of the first EEOF. The *X*-structure is similar in sign and magnitude to the original E_1 (not shown) and the *Y*- structure is similar in sign and has less magnitude than E_2 (not shown). These patterns suggest a positive correlation between the substructures. This interpretation may be extended to include features in one substructure as representing some alteration of the same feature in the other substructure (e.g. propagation of the feature in time and space)

The *X*-structure of EEOF2 (Fig. 3a) represents a negative EOF1 and also has much less amplitude than E_1. The *Y*- structure of EEOF2 represents a positive E_2 with similar magnitude. Similar to EEOF1, EEOF2 is also representing the contributions of EOF1 and EOF2 to the dataset, however, a negative correlation is implied by the EEOF2 substructures. Therefore, the EEOF analysis correctly identified two modes of variability in the dataset, as defined by (11) and (12), but the correlation or relationship between these modes is unclear. The loadings on the *P*'s pertaining to each EEOF are presented in Table 2. Note that the EEOF analysis splits the representation of the dataset's intrinsic modes defined by (11) and (12) into pairs of EEOF, but the correlation implied by the sign of loadings (Table 2) between the pairs is different.

Finally, by definition, the correlation between E_1 and E_2 is the smallest (Table 1), but they are composed of the first two PCs. Although there is little correlation between the

i,j	Corr(X_i,Y_j)
1,4	0.2959
2,5	0.6478
3,6	0.9249

***Table 1.** Correlations Coefficients between the substructures (X_i and Y_j) of the test dataset.*

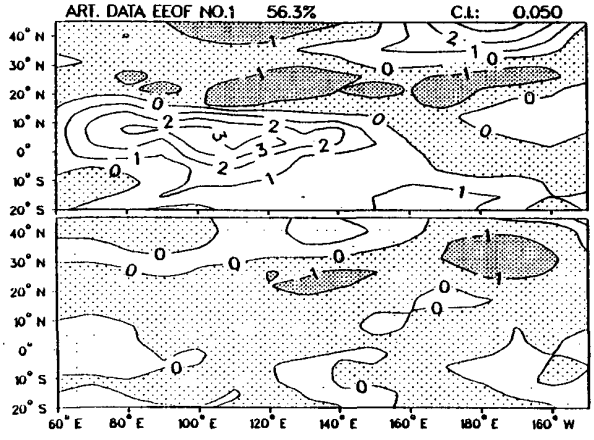


***Figure 2.** (a) X-substructure of the first spatial EEOF pattern of the test dataset. (b) As in (a) except for Y-substructure. Negative values are shaded.*

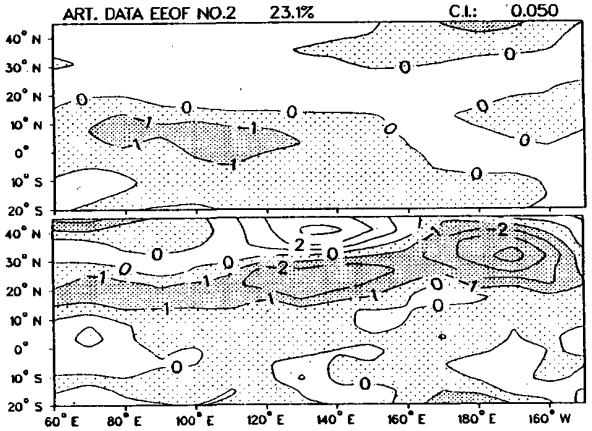


***Figure 3.** As in Figure 2. except for second EEOF.*

#	X-Attribute	Y-Attribute
1	+0.93 E_1	+0.37 E_2
2	-0.37 E_1	+0.93 E_2
3	-0.88 E_3	-0.46 E_4
4	+0.74 E_5	+0.68 E_6
5	+0.46 E_3	-0.88 E_4
6	+0.68 E_5	-0.74 E_6

***Table 2.** Loadings Associated with each EEOF. Empty cells represent zero loading on the the repspective PC.*

substructures, the large variance associated with them dictates that they be represented by the leading EEOFs. The largest substructure correlation is defined by E_5 and E_6 (Table 2) which is composed of the fourth and sixth PCs. Although there is extremely high correlation between the substructures the maximization of variance criterion results in the patterns being placed in low order EEOFs.

4. Summary and Discussions

Including heterogenous data in an EOF (EEOF) analysis may complicate the physical interpretation of resulting EOFs. The direct interpretation of correlation between substructures within one EEOF is misleading as shown by two examples. If it is desirable to study the relationships between substructures, then a canonical correlation analysis (CCA) (Hotelling 1936) which is constrained to maximize correlation rather than variance would be more appropriate. CCA was applied to this dataset (Table 3). The canonical functions are ordered according to maximum correlation rather than maximum variance. Therefore, the mode with highest correlation between substructures is represented by the leading canonical function. Applications of CCA on multiple-set data fields are currently under investigation.

CCA# and Correlation	X-Component	Y-Component
1, 0.9249	$1.0 * X_3$	$1.0 * Y_6$
2, 0.6478	$1.0 * X_2$	$1.0 * Y_5$
3, 0.2959	$1.0 * X_1$	$1.0 * Y_4$

Table 3. Loading associated with each canonical component for the test dataset.

REFERENCES

Hotelling, H., 1936: Relations between two set of variates. Biometrika **28**, 321-277.

Kutzbach, J. E., 1967: Empirical eigenvectors of sea-level pressure, surface temperature and precipitation complexes over North America. J. Appl. Meteor., **6**, 791-802.

North, G. R., T. L. Bell, R. F. Cahalan, F. J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions. Mon. Wea. Rev., **110**, 699-706.

Wang, B., 1992: The vertical structure and development of the ENSO anomaly mode during 1979-1989. J. Atmos. Sci., (in press).

Weare, B. C. and J. S. Nasstrom, 1982: Examples of extended empirical orthogonal function analysis. Mon. Wea. Rev., **110**, 481-485.

ON THE INTERPRETATION OF PRINCIPAL COMPONENT ANALYSIS AS APPLIED TO METEOROLOGICAL DATA

Compagnucci R.H. and N.E. Ruiz

CONICET - Universidad de Buenos Aires F.C.E.Y N. Dto de Cs. de la Atmósfera
Ciudad Universitaria. (1428) Capital Federal, Argentina

1.INTRODUCTION

Principal Component Analysis (PCA) used in multivariate statistics is equivalent to Empirical Orthogonal Functions (EOFs) introduced several years ago by Lorenz (1956) in the research of the atmospheric behavior.

In the following years, EOFs, PCA and generically the eigen analysis were used in order to study spatial variability of meteorological data (Kutzbach,1967; Stidd,1967), in cycles and periodicities analysis of climatological patterns (Kidson,1975), to reduce the dimensionality of basic data and to describe climatological patterns (Craddock and Flood,1963), to assist in analogue forecasting (Grimmer,1963) and to detect homogeneous groups of variables (Dyers,1975).

Since the '80s the eigen analysis has become a frequently used research tool in meteorology and climatology. A lot of applications have been published in the literature during these last years. There exist some articles about theoretical considerations. The extensive discussion given by Richman(1986) refers to the different options in the PCA using rotated and un-rotated solutions. Buell(1975), studied the physical meaning of PCA results varying the shape boundary of the area from which the data come and North et. al.(1982) changed the number of independent realizations and the size of the sample having the same aim. Furthermore, problems with the inferences using EOFs results based on small sample size are shown by Storch and Hannoschöck (1985).

Other advances about physical interpretation of PCA have been made by Wallace and Dickinson (1972), North (1984), Preisendorfer et. al. (1981) , and Preisendorfer (1988), who gives an extended compilation about PCA in meteorology and oceanography.

However, most of these studies refer only to the use of S-Mode input matrix, even when this is not explicity mentioned. But, as is well known, in synoptic climatology problems we often need identified subgroups of observations - namely synoptic maps, thought as snapshots of atmospheric circulation data- with similar spatial patterns, thus simplifying the original time series and corresponding to the case of T-Mode input matrix.

Previously, T-Mode and S-Mode results obtained from PCA of theoretical fields by Vargas and Compagnucci (1983) and Compagnucci (1989) showed the differences between both analysis; Eherendorfer (1987) emphasize that, given the deferens between both solutions, serious thought previor to computer runs should be given to the selection of the decomposition mode bearing in mind the objective of the study.

In order to bring to light the lack of answers to some questions on the interpretation of PCA results obtained by S-Mode and especially by T-Mode input matrix, we used in this work theoretical data sets simulating basic atmospheric flows. We examined different samples chaining the number of fields with a given spatial patterns and on the other hand varying the density and distribution of the grid points in the network.

2.METHODOLOGY

There are various texbooks dealing with the PCA theory, most of them within the context of multivariate analysis as Green(1978), whose notation we adopted and took as reference book.

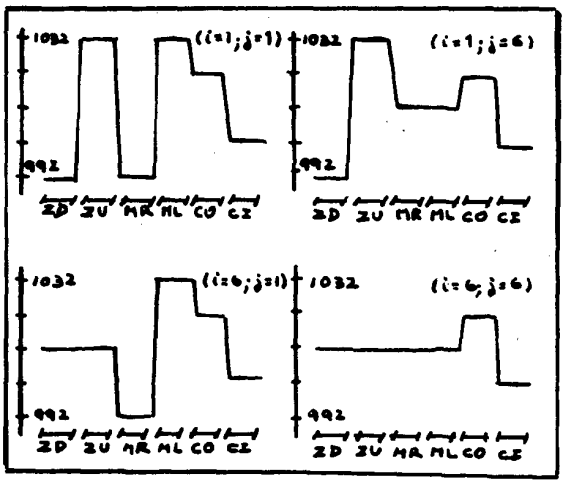


Figure 1: Grid points time series located at NW(i=1;j=1), NE(i=1;j=6), SW(i=6;j=1) and SE(i=6;j=6) in the regular network.

SAMPLES	S-MODE				T-MODE			
ZD. ZUp. MR. ML. CO. CIn.		EIGENVALUE	%VAR.	PATTERN		EIGENVALUE	%VAR.	PATTERN
5 5 10 10 3 3	1	26.06	75.18	A	1	19.97	55.47	F
36 grid point (regular)	2	6.44	16.90	B	2	9.98	27.74	E
	3	2.44	6.79	C	3	5.96	16.55	G
4 8 2 4 6 12	1	26.61	73.92	A	1	17.87	49.65	G
36 grid point (regular)	2	5.50	15.28	B'	2	11.98	33.29	E
	3	3.83	10.65	C'	3	5.99	16.65	F
10 10 5 5 3 3	1	27.09	75.26	A	1	19.97	55.47	E
36 grid point (regular)	2	6.39	17.75	B	2	9.98	27.73	F
	3	2.46	6.85	C	3	5.96	16.56	G
10 10 5 5 3 3	1	35.58	79.06	A	1	20.04	55.67	E
double density in NW sector	2	6.63	14.75	B	2	9.90	27.51	F
(45 grid point)	3	2.73	6.07	C	3	5.96	16.57	G
10 10 5 5 3 3	1	34.88	77.88	A	1	20.03	55.65	E
double density in NE sector	2	7.32	16.27	B	2	9.91	27.55	F
(45 grid point)	3	2.73	6.08	C	3	5.96	16.56	G
10 10 5 5 3 3	1	34.30	76.23	A	1	20.03	55.65	E
double density in SW sector	2	7.86	17.47	B	2	9.91	27.54	F
(45 grid point)	3	2.77	6.15	C	3	5.96	16.57	G
10 10 5 5 3 3	1	34.16	75.91	A	1	20.04	55.67	E
double density in SE sector	2	7.45	16.56	B	2	9.91	27.53	F
(45 grid point)	3	3.31	7.37	C	3	5.96	16.56	G
10 10 5 5 3 3	1	35.28	78.40	A	1	19.96	55.45	E
double density in Central	2	6.83	15.18	B	2	9.98	27.73	F
sector (45 grid point)	3	2.83	6.30	C	3	5.96	16.58	G

Table 1: Eigenvalues, explained variance (%), and patterns for the three first PCs obtained by the S-mode and T-mode. Number of cases in the samples which have zonal flow with gradient down (ZD) and up (ZU); meridional flow with gradient right (MR) and left (ML); circular flow with gradient outside (CO) and inside (CI).

The un-rotated Principal Components (PCs) were calculated using the OCOEF and OFPRI subroutines of the IMSL package. Matrix correlation between variables were chosen as input matrix.

When the S-Mode is run the input matrix is formed by the correlations between stations or grid points time series whose dimension is (nºpoint X nºpoint). In the data matrix the rows are the fields or data snapshots and the columns are the time series grid points for the network. The PCs named Standardized Component Scores (CSs) given the new uncorrelated time series, account for a maximum in residual total variance of the original data set. By means of the PC primary patterns coefficients named Component Loadings (CLs) - corresponding to each CSs - spatial patterns are obtained when attempting to isolate subgroups of grid points which covary similarly.

For the T-Mode the input matrix is constructing using the correlations between fields or data snapshots given from data matrix where the rows are time series of the stations and the columns are the snapshots. The new variables are the PCs or CSs which given the spatial patterns and the CLs are the time series of that corresponding pattern. Therefore, subgroups of snapshots with similar spatial patterns are identified.

The eigenvalues are obtained in descending order and through them we computed the explained variance for each PCs.

3.DATA

In order to analyze the sensivity of PCA solutions to different options in the sample we used sets of variables with previously known internal relations defined by Cattel and Sullivan(1962) as a plasmode.

From the T-Mode point of view, the samples are conformed by 36 elementary basic fields with zonal, meridional and circular flows whose gradients are down (ZD) or up (ZU), right (MR) or left (ML) and outside (CO) or inside (CI) respectively.

On the other hand, from the S-Mode point of view, the fields areas were placed into the regular and quadrangular network with 6X6 grid points (i.e. stations) taken as variables. The number of grid points are changed in some sets of data in order to increase the density of the particular subareas into the network. In fig.(1) we observe only some grid points time series to have an idea of the time series data set analyzed.

3.RESULTS AND CONCLUSIONS

a) Different number of fields:

Having 36 fields in each sample - ZD, ZU, MR, ML, CO and CI - we made several different experiments changing the number of fields for each flow.

When the three basic flows (Z, M and C) have the same number of fields, T-Mode solutions result degenerate like those obtained by Vargas and Compagnucci (1983). In their sample the patterns given by the CSs were lineal combinations of the flow models snaphoots involved in the sample.

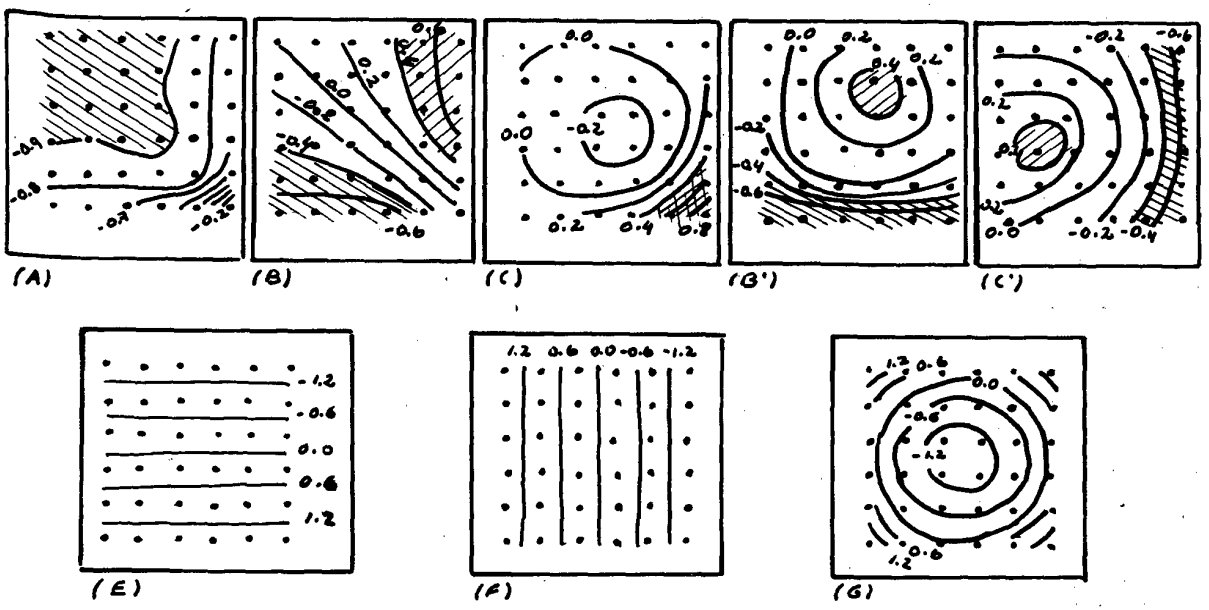


Figure 2: First pattern (A), second and third pattern (B and C or B' and C') for the S-Mode given by the Factor loadings. Patterns (E, F and G) corresponding to the three first PCs in the T-Mode given by the Component Scores.

For the same data Richman (1986) proposed rotated PCA in order to have patterns similar to the actual flows. Here we do not discuss this particular example again.

We show the results for two different groups of samples. In one group we have the same number of fields for both directions of the gradient (i.e. down and up) for each flow; in these cases the mean fields have constant values in all the grid points; in the other group with different number of filds in both directions the mean fields are not constant.

The patterns obtained for the T-Mode are displayed in fig. (2E, 2F and 2G). Showing the same features observed in the actual basic fields involved in the analyzed set. All the results were similar and only the eigenvalues and order of PCs or CSs patterns changed.

Table (1) presents the eigenvalues, the explained variance and their corresponding patterns for some representative examples. As we can see, the eigenvalues are related to the number of snaphots in the sample belonging to a particular CSs pattern.

For each PCs the corresponding CLs time series show values near (1) or (-1) for the snaphots whose features are similar to the CSs pattern and vanish for the other times so that the factor loadings matrix exhibits approximations to a simple structure.

On the other hand, the S-Mode results patterns are given by the CLs (see fig. 2A, 2B and 2C). They changed for the second and third component when the number of circular flows (Co and CI) are higher than half the total number fields. In these cases the first pattern (2A) is the same but the following two have the patterns (B') and (C'); this occurs because the time series in the SE subregion (fig 1) takes greater importance.

Besides, we would understand the strong gradient present in patterns (A) and (C) at SE subregion observing the time series in fig. (1). At (i=6; j=6) values are constant for zonal and meridional flows and only changed for circular flows while at (i=1; j=1) the time series has in opposition the maximum variability.

Shadowed subregions in the patterns are those whose time series grid points have high correlation with the corresponding CSs. In pattern (A) NW subregion points agree with the first CSs time series explained more of half the total variance while time series of points included at SE subareas shadowed in pattern (C) explained the variance account for the third PCs. Pattern (B) shows two different subregion, the NNE in which the time series points (fig.1) display the same values for MR and ML flows in opposition to SSW are

where the point time series have the same values for ZD and ZU flows.

As we can see in the examples of table (1) the eigenvalues in the S-Mode stay nearly constant although the number of fields for same flows changed. Our results agree with those of Buell (1975, 1979). The patterns obtained by the S-Mode seem to be independent to the actual fields topography; while in the T-Mode they were similar to snaphots features in the samples.

b) Modifying the network density:

The last four examples in table(1) related the results obtained when we changed the network density in some subregion. In those samples NW, NE, SW and central sectors the network distributions were respectivetily modified adding nine points.

Eigenvalues results for the S-Mode varied according to the subregion disturbed so that when the points were added to the NW sector we observed the maximum first eigenvalue, in the other side the minimum occur when the nine supplementary points were in SE sector and the third eigenvalue is the higher of all the samples. The second eigenvalue increased when the points were added to SW or NE subareas.

S-Mode patterns were similar to those obtained from the regular network.

The resulting eigenvalués using the T-Mode remain unchanged when the network density is modified. But, CSs patterns corresponding to the zonal and meridional flows show a slight slope while patterns which represent the circular flows have those center biased to the subregion where the points were added.

4.BIBLIOGRAPHY

Buell,C.E., 1975: *The topography of empirical orthogonal functions.* Preprints Fourth Conf. on Prob. and Stats. in Atmos. Sci.,Tallahassee, Fl., Am. Met. Soc., 188-193.

Cattell, R.B. and M. Sullivan, 1962: *The scientific nature of factors: A demostration by cups of coffee.* Beahvioral Sci., 7, 184-193.

Compagnucci, R.H., 1989: *Climatologia sinoptica de las precipitaciones en Cuyo.* Dr. Thesis, Univ. of Bs. As., 238pp.

Craddock, J.M. and Flood, C.R., 1969: *Eigenvectors for representing the 500 mb. geopotential surface over the northern hemisphere.* Quart. J. R. Met. Soc., 95, 576-593.

Dyer, T.G.J., 1975: *The assignment of rainfall stations into homogeneous groups: an application of Principal Component Analysis.* Quart. J. R. Met. Soc., 101, 1005-1013.

Ehrendorfer, M., 1987: *A regionalization of Austria's precipitation climate using Principal Component Analysis.* J. of Clim., 7, 71-89.

Green, P.E. and J.D., Carol, 1978: *Analyzing Multivariate Data.* Dryden Press, U.S.A., 519pp.

Grimmer, M., 1963: *The space-filtering of monthly surface temperature anomaly data in terms of patterns, using Empirical Orthogonal Functions.* Quart. J. R. Met. Soc., 89, 395-408.

Kidson, J.W., 1975: *Tropical eigenvector analysis and the Southern Oscillation.* Mo. Wea. Rev., 103, 187-196.

Kutzbach, J., 1967: *Empirical eigenvectors of sea level pressure, surface temperature, and precipitation complexes over North America.* J. Appl. Met., 6, 791-802.

Lorenz, E.N., 1956: *Empirical orthogonal functions and statistical weather prediction.*Sci. Rep.1, Statis. Forecasting Project, MIT, 48pp.

North, G.R., 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.

North, G.R., 1984: *Empirical Orthogonal Functions and Normal Models.* J. of the Atms. Sci., 41, 790-887.

Preisendorfer, R.W., F.W. Zwiers and T.P. Barnet, 1981: *Foundations of Principal Component selections rules.* SIO Ref. Ser. 81-4. Scripps Inst. of Oceanograpy, 192pp.

_____, 1988: *Principal Component Analysis in Meteorology and Oceanography,* Elsevier, 419 pp.

Richman, M.B., 1986: *Rotation of Principal Components.* J. of Climatology, 6, 293-335.

Stidd, C.K., 1967: *The use of eigenvectors for climatic estimates.* J. Appl. Met., 6, 255-264.

Storch H. v and Hannoschöck, G., 1985: *Statistical aspects of estimated principal vectors(EOFs) based on small sample sizes.* J.Clim. Appl. Meteor. 24, 716-724.

Vargas, W.M. and R.H. Compagnucci, 1983: *Methodological aspects of Principal Component Analysis in meteorological fields.* Procc. "II Int. Meet. on Statistical Climatology", Lisboa, Portugal, Inst. Nac. de Met. y Geof., 5.3.1-5.3.9.

Wallace, J.M. and R. Dickinson, 1972: *Teleconnections in the geopotential height field during the Northern Hemisphere winter.* Mon. Wea. Rev., 109, 785-812.

Analysis of climate anomalies using uninformed patterns

Andreas Hense, Meteorologisches Institut
Universität Bonn, Auf dem Hügel 20, W-5300 Bonn 1

April 30, 1992

1 Introduction

Analysis of observed or simulated climate anomalies in the atmosphere requires statistics to decide whether an observed or simulated anomaly pattern – defined as the deviation of the disturbed case from a suitably chosen control case – is

- due to the naturally occuring internal variability generated by any nonlinear system or
- due to the imposed changes in the boundary conditions (sea ice, sea surface temperatures, trace gas concentration etc.) or whatever is used to define the disturbed case.

In principle the analysis should be done on a highdimensional phase state vector describing the state of the atmosphere or the model as completely as possible. On the other side we have at hand only a limited number of samples of the state vector to draw statistical conclusions on the system. Due to this restriction – limited sample size vs. highdimensional state vector – we can analyse only a (very) small number of degrees of freedom (dof) of the state vector systematically. Therefore we have to choose prior to the statistical analysis out of the large pool of available dof's those which are thought to be relevant. One can think of an a-priori defined filter which separates the relevant dof's from the noise part. The relevant ones are thought to carry the signal imposed eg. by the varied boundary conditions while the noise part describes the internal variability or measurement errors in case of observations. All this ideas have been put foreward by Hasselmann (1979).

There exist several ways to define the filter or the relevant dof's (Storch, 1987). First of all we may think of patterns in grid point space when we talk about dof's. A statistical way to define relevant dof's is the use of optimally chosen patterns with respect to some error measure. The standard meteorological examples are empirical orthogonal functions (EOF's) in any version, which are defined by minimizing the residual variance not represented by themselves. Because these patterns require statistical information namely the estimates of the covariance matrix, we will call such patterns *informed patterns*. They carry information on the second moment structure of the data. Besides this there may be also problems in calculating accurate estimates of the EOF's.

A second point in the analysis of observed or simulated climate anomalies in the atmosphere concerns the interpretation of the anomalies. Either having observed a natural climate anomaly or simulated one with a complicated GCM with a large variety of feedbacks, one may be interested in the questions

- Can we understand the climate anomaly on the basis of a simplified theory?
- Can we interpret (parts of) the signal as arising from only those few feedbacks incorporated in a simple model?

It will be shown that it is possible to treat jointly the statistical signal-to-noise analysis, the data reduction problem and the interpretational part

in an inverse modelling approach. From a simplified model which is chosen a-priori and thought to represent major interactions of the problem under consideration we derive a set of patterns with as least as possible statistical information – at best with no information at all – from the data set to be analysed. This set of patterns will be called *uninformed patterns.* They should specify those dof's of the underlying complex system which should show a response (signal, climate anomaly) in case the simplified theory is valid. Further from these patterns only a few are selected according to some specified rule. Their amplitudes enter the signal-to-noise analysis and allow the systematic evaluation of their statistics.

2 The statistics

We will concentrate on changes of the climatic mean estimated from monthly averages. Let us define the phase state vector which describes the complete (monthly mean) flow as $\vec{x}$. This is a random variable of dimension q. From the simulation or the observation we have at hand two samples of size m_1 and m_2 are drawn from the control and anomaly case. If the flow is affected by the changes in boundary conditions the control and anomaly case constitute different ensembles while in the other case they do not. Based on the given samples we want to decide whether the Nullhypothesis of equal expectation values

$$H_0 : E(\vec{x}_1) = E(\vec{x}_2)$$

or the Alternativhypothesis of different expectations holds

$$H_A : E(\vec{x}_1) \neq E(\vec{x}_2)$$

If we make the assumptions that both underlying ensembles of monthly means are multivariate Normal distributed random variables and that the covariance matrix Σ does not vary between both ensembles the appropriate testvariable for the above pair of hypothesis is the Hotelling T^2 statistics (Morrison, 1976):

$$T^2 = \frac{m_1 m_2}{m_1 + m_2}(\bar{\vec{x}}_1 - \bar{\vec{x}}_2)^t S^{-1}(\bar{\vec{x}}_1 - \bar{\vec{x}}_2)$$

where $\bar{\vec{x}}_i$ are the estimated climatic means und S is the estimate of the covariance matrix Σ. In case that we can compute the inverse of S and that H_0 is true the testvariable T^2 is Fisher-F distributed with$(q, m_1 + m_2 - q - 1)$ degrees of freedom. If we use the standard maximum likelihood estimate for Σ the rank of S is given by $\min(q, m_1 + m_2 - 2)$. Thus for $q > m_1 + m_2 - 2$ the T^2 statistics is not defined and the test is not possible. This is the mathematical expression that a systematic analysis of all q flow components is not possible if only small samples are available. A short calculation even for a low resolution GCM (T21, 10 levels) shows that for data represented by grid point values ($q = 64 * 32 * 10$) any realistic sample size is too small.

3 Data reduction

We have seen that a systematic statistical analysis of the complete state vector with respect to changes of the mean is not possible given a sample of restricted size. We therefore split the state vector into a part which is accessible by statistics and into an unresolvable part. The accessible part is characterized by being spanned by a (low dimensional) set of so called guess vectors $\vec{g}_k, k \in [1, \tilde{q}], \tilde{q} << q$ (Storch, 1987) such that we may write for any $\vec{x}$

$$\vec{x} = \sum_k y_k \vec{g}_k + \vec{r}$$

where $\vec{r}$ gives the unresolvable part of $\vec{x}$. These guess vectors are also q dimensional vectors which can be assumed to represent typical patterns in grid point space. If we arrange the guess vectors as column vectors of a $\tilde{q} \times q$ matrix G and the amplitudes of the guess vectors y_k in a $\tilde{q}$ dimensional column vector we can rewrite the last equation as

$$\vec{x} = G\vec{y} + \vec{r}$$

which may be solved for $\vec{y}$ by requiring

$$\vec{r}^t W \vec{r} = \text{minimum}$$

where W is a yet unspecified, but freely chosen metric of the scalar product. The solution reads

$$\vec{y} = (G^t W G)^{-1} G^t W \vec{x}$$

which is well defined provided that the guesspatterns form a linearly independent set. The dimension of the projected state vector $\vec{y}$ is $\tilde{q}$ which depends only on the numbers of guess vectors specified a-priori. Thus if we specify only $\tilde{q} < (<<)m_1 + m_2 - 2$ guessvectors the covariance matrix of the projected state vector is most probably of full rank and can be inverted. Thus the T^2 statistics in the subspace of the phase space which is spanned by the guessvectors is well defined and can be used to test H_0, H_A that the expectation values of the original state vector in the direction of guess vectors are equal or not.

4 Data

To illustrate the general outline we will show different examples applied to two different data sets. We have at our disposition

- three GCM simulations in permanent January conditions using the T21 (cycle 17) version (Hense et al., 1990) which differ in the SST prescribed in the Atlantic Ocean: *cold* used typical temperature for the decade 1904/13 which are anomalously low, *warm* used SST for the decade 1951/60 where anomalous high SST have been observed. *Control* had prescribed standard climatological SST which are typical for the 70'ies. 30-day means of different variables will be analysed. The sample size is 17 for *cold* and *warm* and 9 for *control*.

- one data set of monthly mean observed free atmospheric temperatures given by the relative topography 500/1000 hPa. The data are calculated from regular analyses of the German Weather Servive (DWD) and cover the forty year period January 1949 to December 1988. The area of interest is the Arctic north of 55°N. The data resolution is 5° in N/S and 10° in E/W direction.

5 The models

We will concentrate on the analysis of differences between climatic means. Therefore we restrict the models to be stationary and – as a first guess – to be linear. A general linear and stationary model ($\mathcal{M}$ is the system matrix, dimension q) of the form

$$\mathcal{M}\vec{\Psi} = \vec{Q}$$

can produce a nontrivial solution $\vec{\Psi}_{stat}$ only in case of a nonzero, stationary forcing $\vec{Q}$. We will assume that such a forcing is present due to the varied boundary conditions but unknown. We can construct the stationary solution with the help of the eigenmodes of the linear model as

$$\vec{\Psi}_{stat} = \sum_k \frac{(\vec{d}_k^t \vec{q})}{\lambda_k \beta_k} \vec{e}_k$$

where $\vec{e}_k$ is the k-th eigenmode with eigenvalue λ_k and $\vec{d}_k$ is the adjoint eigenmode. All eigenmodes and all adjoint modes form a biorthogonal system of the form

$$(\vec{d}_k^t \vec{e}_{k'}) = \beta_k \delta_{k,k'}$$

The stationary solution indicates that $\vec{\Psi}_{stat}$ is similar to some eigenmode $\vec{e}_k$, if the eigenvalue is small (or the projection of the forcing onto the adjoint mode is large but this is unknown). This leads to the conclusion that we may use the eigenmodes sorted according to the modulus of their respective eigenvalue as guess vectors

$$\begin{aligned} \vec{g}_k &= \vec{e}_k \quad , \quad k \in [1, \tilde{q}] \\ |\lambda_1| &< |\lambda_2| \quad \ldots \quad |\lambda_k| < |\lambda_{\tilde{q}}| \end{aligned}$$

We will chose those eigenmodes which are most "resonant" with respect to a stationary forcing. Note that complex eigenmodes define two types of guesspatterns, namely the real and imaginary part seperately. If we perform the statistical test in the subspace of the phase space spanned by those modes and we have to accept the Nullhypothesis we have to conclude that the a-priori chosen linear model does not give an appropiate description for the data. In case that the Alternativhypothesis is to be accepted we have shown that a significant shift in climatic means is present and that the linear model is compatible with the data. We will now give several examples on different linear models which define different sets of eigenmodes or guess patterns. At first we

consider one type of variable (e.g. streamfunction or temperatur alone) at a certain level. One could think of a first step in data reduction.

5.1 Diffusion on the sphere

The most simplest case is to redistribute the forcing by mere diffusion

$$\alpha\nabla^2\Psi = Q$$

If we will analyse data over the sphere or one hemisphere the eigenmodes are simply spherical harmonic Y_l^m. The appropriate guess patterns arise if we sort the spherical harmonics with respect to their eigenvalue $-l(l+1)$. Thus the subspace within which the testing is performed is simply a low order triangular truncation of the total data vector. Note that this type of test is easily implemented for any spectral GCM because the projected data (= amplitudes of the spherical harmonics) are readily available. Applying the test to the streamfunction at 300 hPa and comparing the *cold* and *warm* simulation allows to reject H_0 at an error level of 5% if the gravest 21 modes (T6 truncation) are taken into account. The signal explains about 50% of the total raw difference pattern (*cold* minus *warm*) and is illustrated in Fig.1.

5.2 Advection - diffusion on the sphere

The next step is to add to the redistribution process for Q advective processes with an a-priori specified velocity distribution. The simplest case is a mean zonal advection by a superrotational flow $u = u_0\cos(\varphi)$, then a meridionally varying zonal flow $u(\varphi)$ and finally a twodimensionally varying flow $\vec{v}(\lambda,\varphi)$ where $\vec{v}$ is the mean flow of the control and anomaly sample. Note that this adds statistical information to the guess patterns but this information is orthogonal to the difference pattern between the two samples. Then the general model reads for the anomaly field Ψ'

$$(\vec{v}\vec{\nabla})\Psi' + \alpha\nabla^2\Psi' = Q$$

In case of the superrotional flow the eigenmodes are still spherical harmonics but now sorted according to the eigenvalue $-\frac{\alpha l(l+1)}{a^2} + i\frac{mu_0}{a}$ (a is the radius of the earth). In case of the more complicated flows a spectral representation leads to an algebraic eigenvalue problem which can solved numerically. Examples analysing the differences of the lower tropospheric temperature 850/1000 hPa between the three T21 simulations *cold*, *warm* and *control* are shown in Hense et al. (1990). If one is interested in a restricted area on the sphere – such as the Arctic – one has to discretize the advection - diffusionoperator e.g. by finite difference methods. Selecting appropriate boundary conditions (most probably von Neumann conditions) will lead to an algebraic eigenvalues problem which can be solved by numerical methods. An application of the last example can be used to search for significant temperature trends during the last 40 years over the Arctic or equivalently for significant differences between the first and second 20 year period. Selecting the gravest 25 modes of the advection-diffusion operator (u_0 = 10 m/sec,$\alpha = 10^6$ m^2/sec) discretized over the polar cap yields for January a highly significant trend pattern (error level less than 1%, explained variance 80%) which is depicted in Fig.2.

5.3 Barotropic eigenmodes

The two aforementioned models will apply to any variable at any level. If we are interested in flow changes in the extratropical upper troposphere we can use the barotropic nondivergent vorticity equation for defining a new model. Let Ψ' be the anomalous geostrophic streamfunction und $\bar{\Psi}$ the mean flow streamfunction (mean of control and anomaly, assumed to be fixed and known). Then the stationary, linearized barotropic vorticity equation (Simmons et al.,1983, Branstator, 1985 a,b) reads

$$J(\bar{\Psi},\nabla^2\Psi') + J(\Psi',\nabla^2\bar{\Psi}+f) + \alpha\nabla^2\Psi' = Q_\zeta$$

where $J(a,b)$ is the Jacobian operator in spherical coordinates. Expanding Ψ' in spherical harmonics (T21) and projecting the above equation also on spherical harmonics one obtains again an algebraic eigenvalue problem which determines the eigenmodes (for details see Hense et al., 1990). Accepting the Alternativhypothesis that

the difference field in geostrophic streamfunction between control and anomaly has significant nonzero amplitudes in the subspace spanned by the barotropic eigenmodes then allows the following conclusion: the shift in climatic means is compatible with the notion that a vorticity source Q_ζ embedded within a twodimensional varying, barotropic basic flow has initiated the anomaly Ψ'. Examples can be found in Hense et al. (1990).

5.4 Tropical model

The last model was restricted to the extratropical, upper troposphere. If we are interested in analysing tropical climatic anomalies we have to apply another model. Based on the ideas of Matsuno (1966) and Gill (1980) we can assume that especially SST anomalies in the tropics initiate deep convective heating anomalies. These in turn lead to flow patterns which can be derived from the stationary, linearized shallow water system:

$$\epsilon\vec{v} + \frac{u_0}{a}\frac{\partial}{\partial\lambda}\vec{v} + f\vec{e}_r \times \vec{v} = -\vec{\nabla}\Phi$$

$$\epsilon\Phi + \frac{u_0}{a}\frac{\partial}{\partial\lambda}\Phi + N^2H^2(\vec{\nabla}\vec{v}) = -Q$$

where we have added the advection by a superrotational flow. N^2 is the Brunt Vaisalla frequency of the basic state, ϵ a damping constant and Q the heating distribution. The model can be reformulated in streamfunction and velocity potential and discretized by spherical harmonics which then leads to the eigenvalue problem (Longuett-Higgins, 1964). Each eigenmode has a streamfuction, velocity potential and geopotential component. Therefore it is necessary to define a scalar product metric W which takes into account the different units and absolute values. It turns out that the best way is to choose the energy norm which combines kinetic energy $\frac{1}{2}(u^2+v^2)$ with potential energy $\frac{1}{N^2H^2}\Phi^2$. The acceptance of the Alternativhypothesis allows the conclusion that stationary flow patterns induced by a deep heating Q are compatible with the data. Combining the gravest 12 eastward and westward propagating modes as guess patterns for analysing the equatorial changes between *cold* and *warm* jointly in the all three variables streamfunction, velocity potential at 300 hPa and surface pressure results in a highly significant signal (error level for H_0 less than 1%, explained energy – all three fields together – 14 %). The streamfunction part of the signal is depicted in Fig.3.

6 Conclusion

I have shown several examples how to implement a multivariate, statistical analysis of simulated or observed climate anomalies. However, one is not even restricted to those examples given above. Several other model types for defining uninformed patterns are possible. Also other types of strategies for deriving the guess pattern can be thought of. Extensions of the method to time varying analysis appears also possible. In that case a close relationship to Kalman filtering techniques exists. Lastly the parametric test procedure using the Hotelling T^2 test can be relaxed eg. by using a "jack - knifing" procedure.

7 References

Branstator, G. (1985a,b): *J. Atmos. Sci*,42,2225-2254

Gill, A. (1980): *Quart. J. Roy. Met. Soc.*,106,447-462

Hasselmann, K. (1979): *Meteorolgy over the tropical oceans*, 251-259, Roy. Met. Soc., London

Hense, A., Glowienka-Hense, R. Storch, H.v., Stähler, U., (1990), *Climate Dynamics*,4, 157-174

Longuett-Higgins, (1964): *Proc. Roy. Soc.*,A 279,446-473

Matsuno, (1966): *J. Meteorol. Soc. Japan*,44,25-43

Morrison, D.F. (1976): Multivariate Statistical Analysis, McGraw-Hill, 415pp

Simmons, A.J., Wallace, J.M., Branstator, G. (1983): *J. Atmos. Sci.*,40,1363-1392

Storch, H.v. (1987): *Beitr. Phys. Atmosph.*,60,464-477

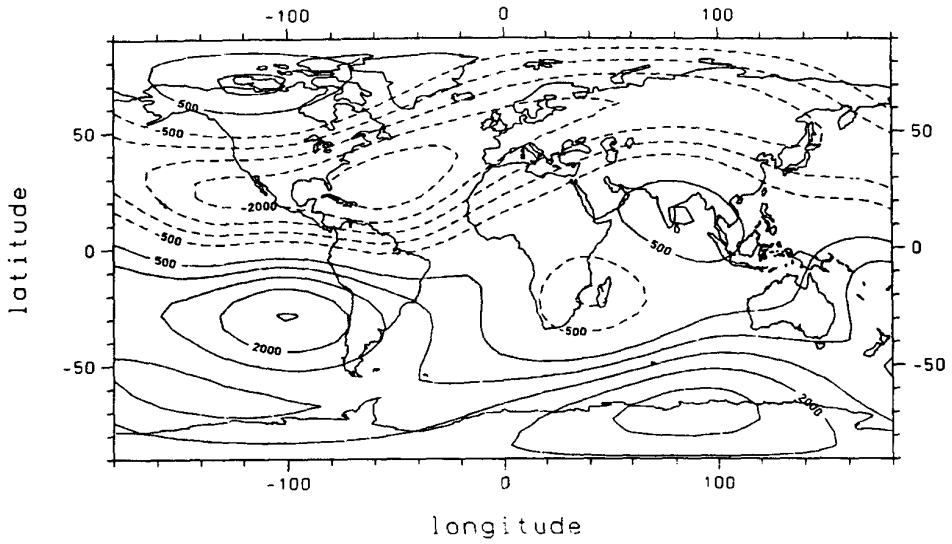


Figure 1: Significant difference pattern of 300 hPa streamfunction between T21 Simulation with cold Atlantic SST and warm Atlantic SST. Spherical harmonics (diffusion on the sphere) are chosen as guess patterns or uninformed patterns. Contour interval is 510^5 m^2/sec. Zero contour is omitted.

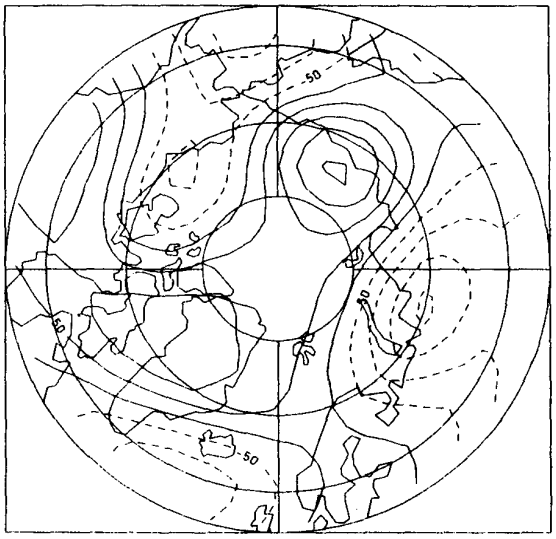


Figure 2: Significant difference pattern of 500/1000 hPa rel. topography over the Arctic between January 1949/67 and January 1968/88. The pattern is equivalent to the linear trend expressed in temperature change per 20 years. The eigenmodes of the discretized advection-diffusion model have been used as guess patterns. Contour interval is 0.5K, dashed contour are negative

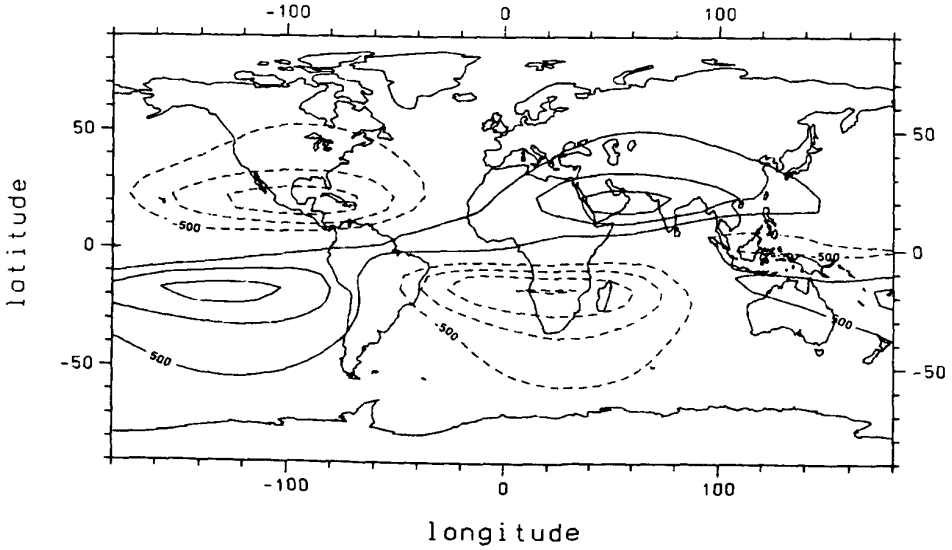


Figure 3: Same as Figure 1, except that the eigenmodes of the simple tropical model have been used as guess patterns

Multidimensional Spectral Analysis of Monthly Fire Weather Patterns

I-Shang Chow and Keh-Shin Lii
Department of Statistics
University of California, Riverside
Riverside, California 92521
Francis M. Fujioka
Pacific Southwest Research Station
USDA Forest Service
4955 Canyon Crest Drive
Riverside, California 92507

Abstract

We discuss the problem of describing the ensemble characteristics of two-dimensional fire weather fields. In particular, we focus on spectral analysis of data indexed discretely in time, and obtained from an irregular array of observation sites projected in a plane. The data are assumed to be generated from a stochastic process according to :

$$Y(t_i, \mathbf{x}_j) = D(t, \mathbf{x}_j) + \epsilon(t_i, \mathbf{x}_j)$$

where $D(t, \mathbf{x}_j)$ is a deterministic function and $\epsilon(t_i, \mathbf{x}_j)$ is a random variable generated by a stationary stochastic process. The variable t is a cyclical index, e.g. month number or julian date; the subscript i represents repeated measurements for a particular value of t. The variable $\mathbf{x}$ is a vector in the plane. The subscript j on $\mathbf{x}$ indexes the observation sites projected in the plane. The covariance function of Y is given by

$$C(\mathbf{u}) = E[Y(t, \mathbf{x} + \mathbf{u})Y(t, \mathbf{x})],$$

with appropriate regularity conditions on C. The corresponding spatial spectral density is

$$S(\omega) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} e^{-i\omega \cdot \mathbf{u}} C(\mathbf{u}) d\mathbf{u}.$$

An analysis of fire weather data from the contiguous United States is presented.

1 Introduction

A meteorological network characteristically forms an irregular pattern of points on a map. Special interpolation methods have been designed to transform the set of network observations to a regular grid of calculated values that collectively represent the meteorological field. One such method is a surface fitting scheme in which nearby station observations are used to fit a bivariate polynomial function (Akima 1978). Gridpoint values within the neighborhood are then estimated by evaluating the fitted polynomial at the gridpoint location.

Another method of interpolating from network data is statistical, in the sense that it utilizes spatial autocovariance information to estimate field values. Under this characterization of the data, the estimates minimize the mean-squarred error of interpolation (Thiebaux and Pedder 1987). This method is called optimum interpolation in meteorology, and is similar in principle to the method of kriging (Cressie 1991). One criticism of the method, particularly for the analysis of surface fields, is that it tends to oversimplify the spatial autocovariance. In fact, the autocovariance is usually assumed to be homogeneous and isotropic, which renders it to a function of distance only. For surface weather fields, this is a difficult proposition to accept, and even more difficult to demonstrate (McCuttchan and Fujioka 1987).

In this paper, we offer a multidimensional framework for spectral analysis, which yields a two-dimensional spatial autocovariance structure that is far more descriptive than the one-dimensional case.

2 Model Description & Methodology

The data used in this investigation is mean monthly temperature which were obtained at the 1024 stations across the United States for the 78-year period, 1910-1987. There are no missing data throughout this interval for each station.

The data are assumed to be generated from a stochastic process according to:

$$Y(t_i, \mathbf{x}_j) = D(t, \mathbf{x}_j) + \epsilon(t_i, \mathbf{x}_j)$$

where $Y(t_i, \mathbf{x}_j)$ is the mean temperature, $D(t, \mathbf{x}_j)$ is a deterministic function, and $\epsilon(t_i, \mathbf{x}_j)$ is a random variable generated by a stationary stochastic process. The index $t_i, i = 1, \ldots, 78$, are time points, i.e. $t_1 = 1910, \ldots, t_{78} = 1987$, t is a fixed month, and $\mathbf{x}_j = (x_{j1}, x_{j2}), j = 1, \ldots, 1024$, are location coordinates.

In this paper we use July mean temperature as an illustration. For example, let $t = July, i = 78, Y(t_i, \mathbf{x}_j)$ is 1987 July mean temperature, $D(t, \mathbf{x}_j)$ denotes the 78-year July temperature average. i.e.,

$$D(t, \mathbf{x}_j) = \frac{\sum_{k=1}^{78} Y(t_k, \mathbf{x}_j)}{n}, \quad n = 78,$$

$\epsilon(t_i, \mathbf{x}_j)$ is 1987 temperature residual.

Our purpose is to find the spectral density and covariance function of $\epsilon(t_i, \mathbf{x}_j)$. The power spectrum of a stationary random process $\epsilon(t_i, \mathbf{x}_j)$, $S(\omega_1, \omega_2)$,

is defined by

$$S(\omega_1,\omega_2)=\frac{1}{(2\pi)^2}\sum_{u_1,u_2=-\infty}^{\infty}C(u_1,u_2)e^{-i(\omega_1u_1+\omega_2u_2)}.$$

And $C(u_1,u_2)$ is the autocovariance function defined by

$$\begin{aligned}C(u_1,u_2)&=E[\epsilon(t_i,(x_{j1},x_{j2}))\epsilon(t_i,(x_{j1}+u_1,x_{j2}+u_2))]\\&=E[Y(t_i,(x_{j1},x_{j2}))Y(t_i,(x_{j1}+u_1,x_{j2}+u_2))].\end{aligned}$$

We take the inverse Fourier transform of the power spectrum to get the autocovariance function, i.e.,

$$C(u_1,u_2)=\int\int S(\omega_1,\omega_2)e^{i(\omega_1u_1+\omega_2u_2)}d\omega_1d\omega_2.$$

To estimate the power spectrum $S(\omega_1,\omega_2)$, we define the periodogram of $\epsilon(t_i,\mathbf{x}_j)$ as:

$$I(\omega_1,\omega_2)=\frac{1}{(2\pi)^2N}|\sum_{j=1}^{N}\tilde{\epsilon}(t_i,\mathbf{x}_j)e^{-i(\omega_1x_{j1}+\omega_2x_{j2})}|^2,$$

for $\omega=(\omega_1,\omega_2)\neq(0,0)$, where $\tilde{\epsilon}(t_i,\mathbf{x}_j)=\epsilon(t_i,\mathbf{x}_j)-\bar{\epsilon}_i$, $\bar{\epsilon}_i=\sum_j\epsilon(t_i,\mathbf{x}_j)/N$ and N represents the number of temperature stations. We also note that $\sum_j\tilde{\epsilon}(t_i,\mathbf{x}_j)=0$. And let

$$\begin{aligned}I(0,0)&=\frac{1}{(2\pi)^2N}|\sum_j\epsilon(t_i,\mathbf{x}_j)|^2\\&=\frac{1}{(2\pi)^2N}|N\bar{\epsilon}_i|^2\\&=\frac{1}{(2\pi)^2}N\bar{\epsilon}_i^2.\end{aligned}$$

Let $\hat{S}(\omega_1,\omega_2)$ be the smoothed periodogram. Here we use the simple moving average method to smooth the periodogram $I(\omega_1,\omega_2)$. When we take the inverse Fourier transform of the smoothed periodogram, we can obtain the estimated autocovariance function, as

$$\hat{C}(u_1,u_2)=\int\hat{S}(\omega_1,\omega_2)e^{i(\omega_1u_1+\omega_2u_2)}d\omega_1d\omega_2.$$

Statistical properties of these estimates are given in Chow and Lii (1992). General references on the spatial point processes and spectrum estimation with random sampling are given in [2, 6].

Once we have the estimated autocovariance function, $\hat{C}(u_1,u_2)$, we can predict or interpolate the temperature at any location $\mathbf{x}_k$. For example, if $Y(t_i,\mathbf{x}_j)$ is a missing temperature at location $\mathbf{x}_j$, $\hat{Y}(t_i,\mathbf{x}_j)$ is the predictor of $Y(t_i,\mathbf{x}_j)$ based on the observations, $Y(t_i,\mathbf{x}_k),k\neq j,k=1,\ldots,N$, the mean square error (MSE) of the predictor is

$$MSE\{\hat{Y}(t_i,\mathbf{x}_j)\}=E\{[Y(t_i,\mathbf{x}_j)-\hat{Y}(t_i,\mathbf{x}_j)]^2\}.$$

Generally, the problem of determining optimal predictors requires that the class of predictors be restricted. We shall investigate linear predictors. Where $\hat{Y}(t_i,\mathbf{x}_j)=\sum_k c_kY(t_i,\mathbf{x}_k)$, $k\neq j$. Let

$$Q=E[Y(t_i,\mathbf{x}_j)-\hat{Y}(t_i,\mathbf{x}_j)]^2.$$

Letting $\partial Q/\partial c_k=0$, we obtain

$$\begin{aligned}\frac{\partial Q}{\partial c_k}&=2c_k\sigma^2+2\sum_{m\neq k}c_mCov(Y(t_i,\mathbf{x}_m),Y(t_i,\mathbf{x}_k))\\&-2Cov(Y(t_i,\mathbf{x}_j),Y(t_i,\mathbf{x}_k))\\&=0.\end{aligned}$$

Hence,

$$\begin{aligned}\hat{c}_k&=\frac{-\sum_{m\neq k}c_mCov(Y(t_i,\mathbf{x}_m),Y(t_i,\mathbf{x}_k))}{\sigma^2}\\&+\frac{Cov(Y(t_i,\mathbf{x}_j),Y(t_i,\mathbf{x}_k))}{\sigma^2}.\end{aligned}$$

The best linear predictor for a stationary spatial series with known autocovariance function is easily obtained.

3 Data Analysis

The data used in this investigation is the mean monthly temperature from the United States Historical Climatology Network (HCN) (Karl *et al* 1990). The HCN data has been described as probably the best monthly temperature data set available over the contiguous United States for analyzing long-term climate trend on a regional scale (Karl *et al* 1990). The HCN network includes approximately 1200 stations in the contiguous United States, with a rather uniform spatial distribution. Our study period runs from 1910 through 1987. There are missing data exist for some stations throughout this interval. In this analysis we use 1024 stations with no missing data in the study period.

The example illustrates that in predicting or interpolating the missing observation from observations which are selected from the neighborhood of the realization. The selection of the observations is accorded to the contour plot of the autocorrelation function (Figure 4). In this study, we use 106° longitude to seperate the USA into EAST

Table 1: 8 missing data predicted by using Entire USA 1016 observations.

Sta. Id	Obs.	NCAR	Spec.	Nres.	Sres.
31596	82.12	82.527	82.109	-0.407	0.011
95874	81.34	81.124	81.601	0.216	-0.261
115079	78.19	78.046	77.104	0.144	1.086
129557	77.64	79.205	77.211	-1.565	0.429
144559	82.09	80.777	81.209	1.313	0.881
229079	79.31	79.678	79.464	-0.368	-0.155
303319	70.98	71.249	72.033	-0.269	-1.053
444876	74.49	73.415	74.888	1.075	-0.398
ABS.				5.357	4.274
MSE				5.770	3.499

Table 2: 8 missing data predicted by using EAST USA 754 observations.

Sta. Id	Obs.	NCAR	Spec.	Nres.	Sres.
31596	82.12	82.527	81.924	-0.407	0.196
95874	81.34	81.109	80.623	0.231	0.717
115079	78.19	78.046	77.758	0.144	0.432
129557	77.64	79.205	78.128	-1.565	-0.488
144559	82.09	80.777	81.481	1.313	0.609
229079	79.31	79.678	79.120	-0.368	0.190
303319	70.98	71.249	71.574	-0.269	-0.594
444876	74.49	73.415	74.902	1.075	-0.412
ABS.				5.372	3.638
MSE				5.776	1.907

Table 3: Another 8 missing data predicted by using Entire USA 1016 observations.

Sta. Id	Obs.	NCAR	Spec.	Nres.	Sres.
121873	75.70	75.812	75.292	-0.112	0.408
160549	82.40	80.593	81.883	1.807	0.517
252840	77.96	78.380	78.411	-0.420	-0.451
257715	78.27	78.101	77.951	0.169	0.319
303184	72.08	73.622	73.019	-1.542	-0.939
304174	70.35	71.938	71.832	-1.588	-1.482
347012	80.00	80.132	80.970	-0.132	-0.970
385017	81.00	81.515	81.885	-0.515	-0.885
ABS.				6.285	5.971
MSE				8.665	5.541

Table 4: 4 missing data predicted by using WEST USA 248 observations.

Sta. Id	Obs.	NCAR	Spec.	Nres.	Sres.
42294	70.24	71.988	70.851	-1.748	-0.611
49087	68.34	68.714	68.562	-0.374	-0.222
350304	67.21	68.112	68.601	-0.902	-1.391
454154	72.59	72.342	72.563	0.248	0.027
ABS.				3.272	2.251
MSE				4.070	2.358

USA and WEST USA. EAST USA has 762 observations(stations) and WEST USA has 262 observations. We randomly delete 8 observations each time in EAST USA and 4 observations in WEST USA. In Figure 1, × denotes the 8 missing data we want to predict in the Table 1 and Table 2, • denotes another randomly selected missing data set in table 3, and ○ denotes the 4 WEST USA missing observations in table 4. Figure 2 and Figure 3 are the smoothed periodogram and autocorrelation function respectively. From Figure 4, the contour plot of the autocorrelation function, we see that the data are more east-west correlated than north-south. In prediction, we will truncate the autocorrelation function when its values are small. We have selected 0.065 as the truncation point.

In Table 1, Obs. denotes the actual temperatures, NCAR denotes the predictions using a bivariate interpolation, Spec. denotes the predictions using the spectral method, Nres. is the residual of NCAR, and Sres. is the residual of the spectral method. ABS. denotes the sum of the absolute deviation and MSE denotes the mean square error.

Table 2 gives the result of using 754 observations in EAST USA to estimate the autocorrelation function and to predict the same set missing data as in Table 1. We see that the improvement for the spectral method is substantial. Table 3 is another replication. Table 4 is the 4 missing data prediction in WEST USA. Figure 5 is the bar chart of absolute deviations of 8 missing data in Table 2. Black bar is the absolute deviation of NCAR and white bar is the spectral method.

4 Conclusions

We have proposed a method to estimate the spectral density and covariance function of a random field based on data set directly without imposing apriori structure on the covariance function. It is demonstrated with HCN data that the accuracy of prediction can be improved substantially using covariance function thus obtained. It is also shown that the autocovariance function of the monthly mean temperature is not isotropic and not homogeneous between the EAST and WEST USA.

References

[1] H. Akima, *A method of bivariate interpolation and smooth surface fitting for irregularly distributed data*

points, ACM Transactions on Mathematical Software, 4: 148-159, 1978.

[2] D.R. Brillinger, *The spectral analysis of stationary interval functions*, in Proc. Sixth Berkeley Symp. Prob. Statist., Berkeley, CA, pp. 483-513, 1972.

[3] I-Shang Chow and Keh-Shin Lii, *Spectral analysis of random fields with random sampling*, in Probabilistic and Stochastic Methods in Analysis with Applications, Ed. by J. S. Byrnes, Kluwer, 1991.

[4] Noel A. C. Cressie, *Statistics for spatial data*, John Wiley & Sons, Inc., 1991.

[5] Thomas R. Karl, Claude N. Williams, Jr., F. T. Quinlan, and T. A. Boden, *United States Historical Climatology Network (HCN) serial temperature and precipitation data*, Oak Ridge National Laboratory Environmental Sciences Division Publication No. 3404, 1990.

[6] E. Masry, *The estimation of the frequency-wavenumber spectrum using random acoustic arrays - Part I: Performance of beampower pattern estimators and - Part II: A class of consistent estimations*, J. Acoust. Soc. Am. 76 (1984).

[7] M. H. McCutchan and F. M. Fujioka, *Statistical analysis of spatial fire weather fields*, Proc. Ninth Conf. on Fire and Forest Meteor., April 21-24, 1987, San Diego, CA. American Meteorological Society, Boston, 179-184.

[8] H. J. Thiebaux and M. A. Pedder, *Spatial objective analysis*, Academic Press, pp. 299, 1987.

Spatial Distribution of EAST & WEST Missing Data Set

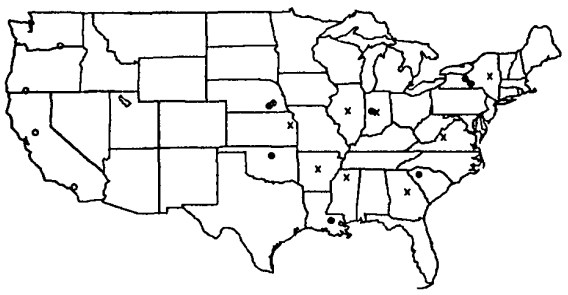


Figure 1

Smoothed Periodogram

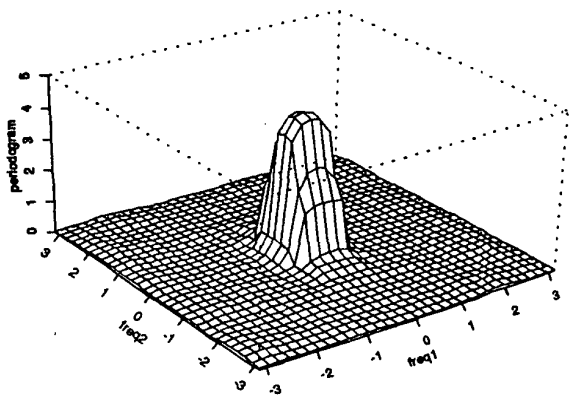


Figure 2

Autocorrelation Function

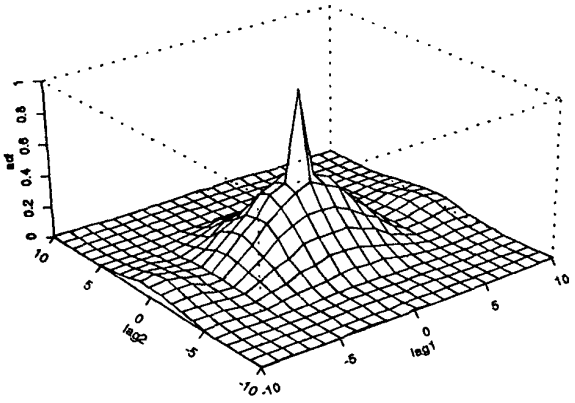


Figure 3

Contour Plot of Autocorrelation Function

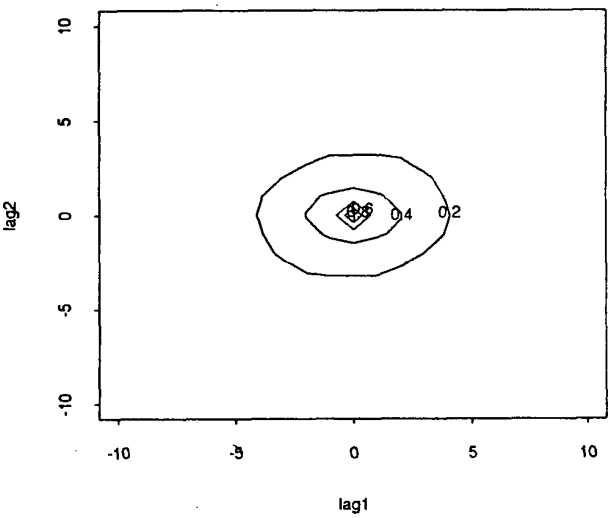


Figure 4

Absolute Deviations of 8 Stations

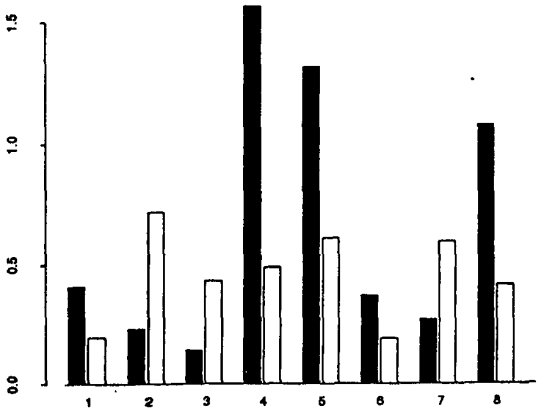


Figure 5

A Comparison Between Dynamical Dimension and Singular Spectrum Analysis for Surface temperature

Compagnucci, R. H. (*) and M. A. Figliola
(*) Dto de CS. de la Atmósfera and (**) Dto de Física
Ciudad Universitaria (1428) Capital Federal, Argentina

1. Introduction

Since the famous article published by Lorenz (1963) , meteorologists keep in mind that the behavior of the atmosphere is a nonlinear dynamical system.
Physicists and mathematicians developed different methods in order to characterized the dynamic of the system.
Packard et al (1981) have suggested an approach to reconstruction a state description from a time series, named the method of delays. It be able study the dynamic of the system since the information of a single variable, that allowed Grassberger-Procaccia (1983) developed the algorithm for calculating correlation dimension.
As an alternative method for reconstruction dynamics, singular system analysis, was developed by Broomhead and King (1986), based on Bertero and Pike (1982). Whose methodology consist of applied Principal Value Decomposition to the covariance matrix. Almost simultaneously, Friederich (1984) by using Empirical Orthogonal Functions analyzed climatic time series, in order to get the time evolution and trying to find climatic attractor's dimension. He compared his results with those obtained using Grassberger-Procaccia algorithm.
Vautard and Ghil (1989) proposed Singular Spectral Analysis (SSA) as an application of Broomhead and King method were the time delay is equal to one. The significantly eigenvalue's number was label as system statistical dimension.
After that, Rasmusson applied this method for analyzing the Southern Oscillation relevant modes.
In this work the correlation dimension was calculated by using the Grassberger -Procaccia algorithm, reconstructing the attractor with the Takens's method. We used this value as a reference of dynamical dimension and to look for the optimal length window for the SSA, in order to obtained the same statistical dimension. To analyzed the waves present in each PCs of the optimal window, we applied Fast Fourier Transform algorithm at the results.

2. The methods

a) The singular spectrum analysis: estimate the number of significat degrees of freedom contained in the data, in order to estimate lower and upper bounds for the number of statistically-significant degrees of freedom, called statistical dimension S. The covariance matrix of the signal for a window $rw = Mrs$ is:

$$Ci,j = (1/(N-(j-i)-1)) * \sum_{i,j>0} x(k)x(i-j)+k) i,j = 0,1 \ldots,M-1$$

were x(t) is the signal, Ci,j is the autocorrelation function, ρs is the sampling interval, M the embedding dimension (or the length of the window). The eigenvalues λk of the matrix Ci,j are real and positive:

$$\lambda k \rho k = C\rho k / M$$

Their squares roots are called singular values and represent the standard deviations. Under good circumstances, λk falls off exponentially with k, until it reaches a floor determined by noise level. The numbers of eigenvalues above a flat noise floor gives the statistical dimension S. The ρk are the corresponding eigenvectors, called also empirical orthogonal functions and they become a natural space for reconstruct the original signal:

$$X(i+j) = \sum_{k} xk(i)\rho k(j)$$

The derivation of the coordinate system is know as the Karhumen-Loeve method.
b) The algorithm of Grassberger-Procaccia: give the correlation dimension.
They calculate the correlation integral Cd(r) give by:

$$Cd(r) = \lim_{N\to\infty} (1/N) * \sum_{i,j} \Theta (r-(xj-xi)d)$$

were Θ is the Heaviside function were $\Theta(s)=1$ if $s>0$ and $\Theta(s)=0$ if $s<1$, and xi, xj the vector position of a point belonging to reconstructing attractor in the d-dimensional phase space.
From Cd(r), the correlation dimension D of the attractor is obtained according to:

$$D = \lim_{r\to 0} \log 2(Cd(r))/\log 2(r)$$

Takens (1981) proof the noise free data, the embedding dimension d is equal to $2D' + 1$, where D' is the earliest greater integer.

c) Power spectrum: The power spectrum (or spectral density) S(w) of a process x(t) is the Fourier transform of its autocorrelation R(t):

$$S(w) = \int \exp(-iwt)\, R(t)dt$$

In this particular case, we calculated the discrete

Fourier Transform of the series,

$$X(k) = (\sum_{n=0}^{N} \exp(-ink)\ x(n))/(N)$$

and

$$S(k) = |X(k)|^2/N$$

3. The Data:

We choose work with the series of Surface Temperature of 'Pampa del Castillo'. It is the only available data series of Argentine sampled in steps of five minutes.
The meteorological station is placed in 45° 47' of latitude and 68° 05' of longitude at 450 mt. high above of level sea, in Chubut (Argentine), approximate in the center of the Patagonia.
The area correspond at the westerlies zone of South hemisphere at the Andes leeward.
These data correspond at an experimental measure since 1-7-1988 at 17 hs. to 1-9-1989 at 16 hs.
We construct an annual series, sampled each hour with 8830 data. and another two season series sampled in steps of five minutes: summer(since 1-7-88 to 3-20-88) and winter (since 7-7-88 to 9-18-88), with 26640 data each one.

4. Results and Discussion

The temperature series for 1988, showed in the figure (1) has a sampling period of an hour. Figure (2) shows power spectral of it.

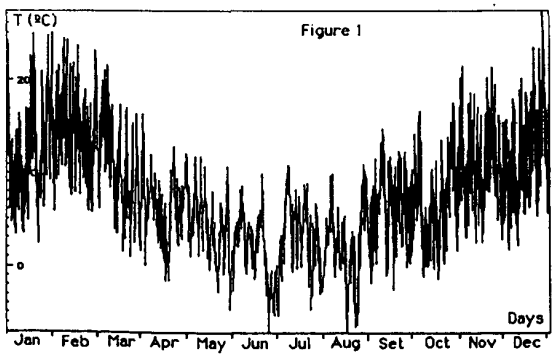


Fig 1: Time Series of the Surface Temperature of Pampa del Castillo (Patagonia, Argentine)

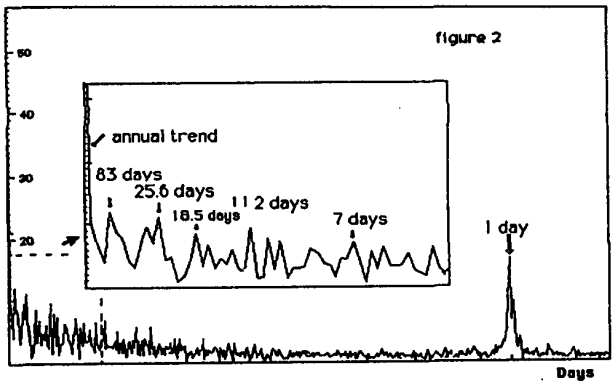


Fig 2 FFT of The Temperature time series

The daily wave can see it with precision furthermore of another significat peak, with periods between 7 and 83 days and the annual trend.
For these series, the characteristic dimension, calculated according to the Grassberger- Procaccia algorithm is 11.182 ± 0.034.
Figure 3 exhibits the natural logarithm of the eigenvalue for the covariance matrix of the data versus the order of its; for different windows lengths (M).

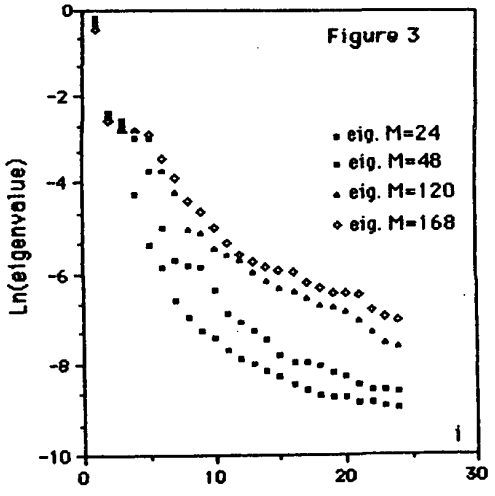
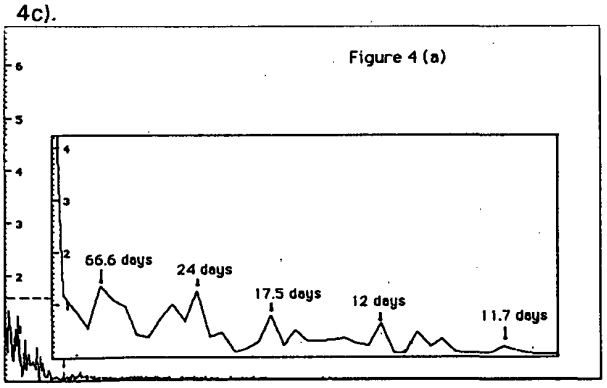
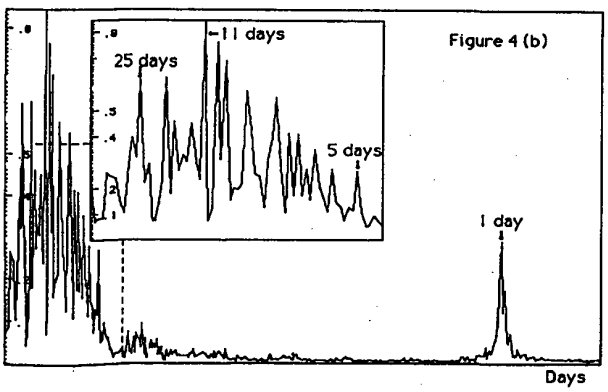


Fig. 3 Ln of Eigenvalue vs order of Eigenvalue for M = 24, M = 48, M = 120 and M = 168

For M = 168 (corresponding to seven days), it is observe that there are 12 relevant PCs in spite of the plateau.
For the cases M = 24 and M = 48 (one and two days), the number of significative eigenvalue is less than 12. For M = 120 (five days) there are 10 or 12

eigenvalue in this situation.
We can this results understand knowing that Rossby's waves have an average period between four and seven days, changing throughout the year. In this case, the winter frequencies are lesser than the summer's; and precisely correspond at the window length which the statistical dimension results greater or equal than the dynamical dimension.
For this reason, we choose M = 168 as the optimal window.
On the another hand, spectral analysis applied at the three first components (PCs), show that it only can discriminate the different frequencies very well for the case M = 168.
For the first PC, the variance explained for the daily wave vanish (see fig. 4a) and the periods greater than ten days make important contributions of the variance. The maximum contribution at the variance correspond for the annual cycle.
The power spectrum for the second PC, shows that greater contribution of the variance. It is explain for the waves with periods in a range of 25 and 5 days, center in 11 days. (Figure 4b). Can it see too, waves of less period and vestiges of the daily wave but with an amplitude lower to the amplitudes of medium scale waves, corresponding at the synoptic scales.
For the third PC, spectral analysis exhibits that the whole variance is centered at the daily wave (Fig. 4c).





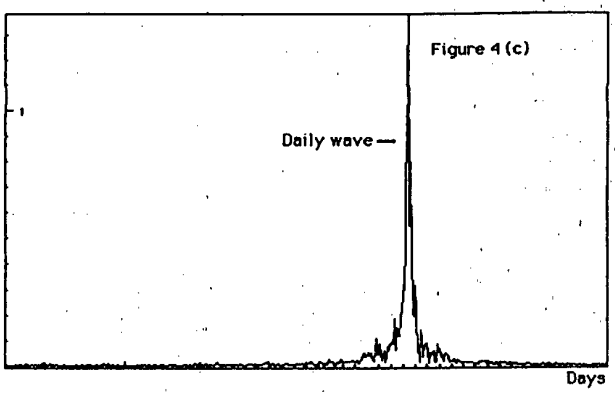


Fig 4: FFT for the three first components

The first Pc is smooth and present the season influence with maximum values in summer and minima in winter. It gives superposed waves of lower periods with minima which correspond the bursting in the area of cold air. Too, it shows the fast transition between summer and winter. It happen since the first days of March when there is a great fall of the values of the temperature, and continue to the middle of april, when appear the first winter's minimum.
The irruptions of cold air are constant and persistent carrying the temperatures to values every time lower during a prolonged period. The winter-summer transition is more slow than the summer-winter's. The cold air irruptions (minima of the data series are following to lengthy periods with average temperature upper than the precedent temperatures. This conclusions are impossible to obtain if we only regard the original data row (Fig. 1)
The second Pc. (fig. 5b) no has the trend and shows the influences of the synoptic scales in Rossby waves.
Waves with periods around ten days or more appear in the winter's months (specially June and July), meantime the summer's waves have a less period (specially in the last months of the studied period). Temporal series of the third PC shows the behavior through the year of the daily wave's amplitude; being less the winter amplitudes than the summer's. This is according with the continental climate where is the meteorologic station.
For the summer series (January-March 1988, 26640 data point, sampled in step of 5 minutes), the dynamical dimension is 7.264 ± 0.0085, while for the winter series (June-September 1988, 26640 data point, sampled in step of 5 minutes) is 6.271 ±

0.036. These differences between winter and summer could be for the presence of the winter waves with period greater than the summer waves. The summer exhibits greater influence of waves with periods between 1 and 7 days. Is notable too, the difference of the annual dynamical dimension and the season dynamical dimension (summer and winter). These can explain as following: Due to the difference in the sampling, season series include micro and macro scale phenomena in the data. Besides, because of the studied series only enveloped a part of the station, waves of phenomena with a scala greater than the synoptic scale, produce no important contribution to the variance and we can observe they as a trend.
For the annual series, (sampled each an hour) micro and macro scale phenomena lost importance, but the synoptic scale is more relevant; specially the annual wave. Furthermore, the non stationary Rosbby waves appears in it (with the winter period greater than the summer period).

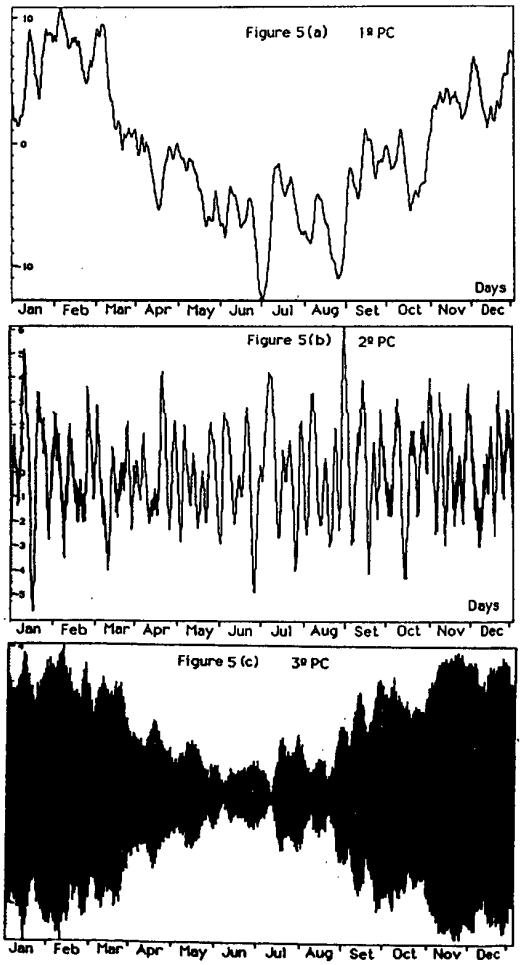


Fig 5: The Three first PCs time series

5. Conclusions

Statistical Dimension calculate with windows length less than seven days is less than the dynamical dimension.
The SSA calculated for windows which can obtain a number of eigenvalues greater than the dynamical dimension is a useful tool for projecting the original series in a convenient base (PCs) and every one of its exhibits clearly only the variance corresponding in a determinate scale of the phenomena; and it allow analyzed different influences of the observed variability.
The joint use of all these techniques, allowed a complete understand of the climatic elements behavior.

References

Lorenz, E. N.; 1963 "Deterministic nonperiodic flow", J. Atmos. Sci., **20**, 130

Packard, N. H., J. P. Crutchfield, J. D. Farmer and R. S. Shaw, 1980: "Geometry from a time series", Phys. Rev. Lett. ,45,712

Grassberger, P and I. Procaccia, 1983: 'Characterization of strange attractors". Phys. Rev. Lett., **35**, 927

Broomhead, D. S. and P. King, 1987: ' Extracting qualitative dynamics for experimental data', PhysicaD, **20**, 217.

Bertero, M. and E. R. Pike, 1982:'Resolution in diffraction-limited imaging, a singular value analysis. I: The case of coherent illumination'. Op. Acta, **29**, 727.

Fraederich, K., 1985;'Estimating the Dimension Weather and Climate Attractors', J. of the Atm. Sci., **43**, 419.

Vautard R. and Ghil, 1989: 'Singular Spectrum Analysis in nonlinear dynamics, with applications to paleoclimatic series', Phys. D, **35**, 395

THE USE OF THE WAVELET TRANSFORM FOR CLIMATIC ESTIMATES

Eduardo SERRANO*, Rosa COMPAGNUCCI**, and Marcela FABIO*

UBA - FCEyN - Dpto de Matematica (*) - Dpto de Ciencias de la Atmósfera (**) - CONICET (**)

Pab I - Ciudad Universitaria - NuNez (cp :1428)
Buenos Aires - República Argentina

Introduction

Since time series of continuous climatic elements may be considered as sampled signals then it is possible to analyzed them with continuous methods both in time and frequency. Usually, a signal is represented in the time domain or, in the frequency domain, by means its Fourier Transform (FT) such as the classic textbooks Shannon (1949) and Pincibono (1966) dealing. Both representations are complementary because each of them have the same information, but the results don't lead to explicit relationship between time and frequency domains.

As it is well known, several different phenomena coexist in the general atmospheric circulation, going from high frequency namely *micro and meso scale perturbations*, to middle frequencies, i.e. *synoptic scales* low frequencies given by the *climatic scale*. This wide range of different and interelated wave lengths appears as a complex behavior in the observed signal.

Time-frequency domain methods are required from extracting physical parameters of interest when these involve join variations in time and frequency.

Wavelet Transform (WT) is a recent mathematical technique witch allow us to make a deep analysis of a signal. It appear to be suitable for the time-frequency analysis in that case. Meyer (1991) provide a comprensive illumination of the WT methods.

They have been applied mostly to signal processing, image coding and numerical analysis. A recently application at the turbulence modeling in order to compute the time evolution of the turbulence flow as given by the Navier-Stokes equations was made by Farge (1992).

With the aim of study the surface pressure field behavior as a primary description to the general circulation we applied the discrete WT to the data from the Compagnucci and Salles (1992) results obtained by the Principal Component Analysis (PCA) of the pressure field. A whole view of the continuous and of the discrete WT are discussed by Hail and Walnut (1989). Nevertheless, here will be given a short outline of the most important features of the WT in the order to its applications to a current signal.

Methodology

A integrable function $\psi(t)$ is called wavelet or analyzing function if it has zero mean and also it is a well localized in both time and frequency domains. From this function ψ, a family of continuously dilated and translated wavelets is generated:

$$\psi_{ab}(t) = |a|^{-1/2}\psi\left(\frac{t-b}{a}\right) \quad a,b \in \mathbb{R} \quad a\neq 0$$

and the associated continuous WT of a square integrable signal s(t) is defined:

$$\Psi\, s(a,b) = \int_{\mathbb{R}} s(t)\psi_{ab}(t)\, dt$$

a is the scale dilation parameter and b is the translation parameter. If an appropiated discrete subset of parameters is selected the discrete WT results.

For some special wavelets and for the parameters:

$$a_j = 2^{-j} \text{ and } b_{jk} = k\, a_j \quad j,k \in \mathbb{Z} \ ,$$

the numerable family:

$$\psi_{jk}(t) = 2^{j/2}\psi(2^j t - k) \quad j,k \in \mathbb{Z}$$

constitutes an orthonormal basis of the space of the finite energy signals, $L^2(\mathbb{R})$.

Orthogonality implies that a square integrable signal s(t) can be represented by:

$$s(t) = \sum_j \sum_k c_{jk}\psi_{jk}(t)$$

with $c_{jk} = \int_{\mathbb{R}} s(t)\psi_{jk}(t)\, dt$ (wavelets coefficient) and the total energy is conserved:

$$\int_{\mathbb{R}} |s(t)|^2 dt = \sum_j \sum_k |c_{jk}|^2$$

Our methods are based on the discrete WT. The cubic spline wavelet $\psi(t)$, is the analizing function.

If $\Delta t=T$ is the sampling rate of the signal s(t), the collection:

$$\psi_{jk}(t) = 2^{j/2}\psi(2^j t/T-k) \quad j,k \in \mathbb{Z}$$

is the current orthonormal basis of $L^2(\mathbb{R})$ These elemental functions are well localized on the intervals $(T/2^j)(k-1/2,k+3/2)$ and its Fourier Transform are band pass filters on the two side bands $(2^j/T)(\pi,2\pi)$

The classical Mallat Algorithm (1989) is used to compute the wavelet coefficients These coefficients give us a time frequency representation of the signal corresponding to a dyadic grid on this domain:

c_{jk} is localized in $((1/2+k)T/2^j;\pi 2^{j-1}/BT)$

In other words, wavelet analysis allows the decomposition of the signal in such way that the band width enlarges as the frequency increases with inversely proportional time resolution. In contrast to the Fourier analysis, the wavelet coefficients keep the locality information present in the signal and the local reconstruction is then possible. Also, abrupt changes in the signal or well localized perturbations can be detected.

However, a better precision in the frequency domain is often necessary to detect and to characterize stationary phenomena, spectral aspects or to extract frequential parameters. Serrano (1991) had proposed a more extensive family of elemental functions. The spline wavelets are included. These functions are generated from the analyzing wavelet $\psi(t)$ and can be considered as Fourier-windows with band width and temporal scale varying in broad range, in inverse proportion. They allow combining in some way Fourier analysis with wavelet analysis. An extensive *library* of orthonormal bases of $L^2(\mathbb{R})$ can be extracted and a suitable representation of the signal in the time frequency domain can be obtained.

Each basis of the library give us a particular structure of the time frequency information contained in the signal. The Shannon entropy criterion, allows one to detect remarkable structures.

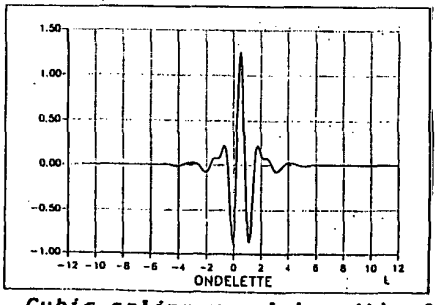


Fig 1. *Cubic spline wavelet $\psi(t)$. Analyzing function.*

Data Analyzed

Daily pressure fields in Southern South America for 1976 and 1977 were previously analyzed by Compagnucci and Salles (1992) using un-rotated PCA with T-mode input matrix (according to Green,1978).

More than the 50% of the variance was account for the first PC. The corresponding pattern has a marked resemblance with the mean sea level pressure field (see fig.2). With zonal westerlies extends south to $40^{o}S$ and the subtropical anticyclons of South Pacific and Atlantic Oceans northern to that latitude. The explained variance for the inverse situation was insignificant. In this work we analyzed the component loadings (see fig.3) corresponding to the first PC which give us the relation between the current field in the sample and the PC pattern feature. Factor loadings near (-1) respond to pressure fields similar to this pattern. When other independing patterns are present in the current field the values approach to (0). The inverse situations have values near to (1). If the pattern persist all along, then surface circulations would be considered like a stable flow. When different patterns are present it is perturbed. In other words, they give us the possible perturbations to the mean general circulation represented by the first pattern.

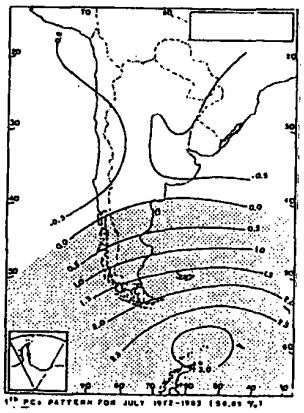


Fig 2. *First PC pattern given by the PCA of daily surface pressure fields.*

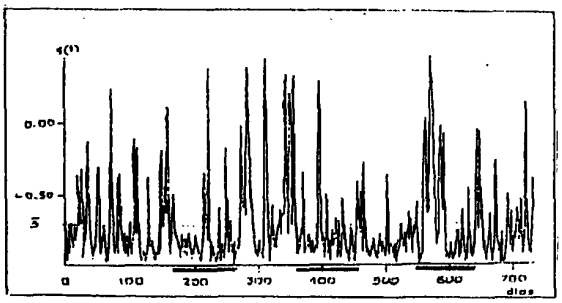


Fig 3. *Factor loadings corresponding to the first PC.*

Conclusions and results

In a first step we compute the wavelet coefficients cjk of the signal, for the levels j=-1 to j=-6. Their wave length and integral energy are showed in table 1. We can see that, the great value of energy occurs in the level j=-4, or for wave length 11 to 22 days, for the whole period analyzed. The fig.4 displays the time distribution of energy for this level. It is remarkable the difference between the energy of the Winter '76 and '77.

In order to extract moré time frequency information, an appropriate basis of the library is selected, according the minimal entropy criterion. The resultant time frequency distribution of the energy is showed in the map of the fig.5 and the topograph feature of the fig.6. Certainly structure can be observed. For instance, the characteristic of the Winter '76 is quite different to the Winter '77. In the first case, the significant energy contribution takes place only at the high frequencies and during the second half time. In the latter case, that contribution occurs at the whole wave lengths.

These observations agree with the results previously obtained by Bischoff and Fernandez (1978), in the analysis of the middle and high troposphere in the same periods. Also it is remarkable the high energy at low frequencies, during the Spring '76 and the Summer '76-77 . These structure lingers over the Autumn and Winter '77. During the last Autumn, the are not contributions of high frequency waves, and we can said that this period is the most stable in the model corresponding to the first PC pattern.

A suddenly and significant phenomenon, given rise a *burst* of energy, is observed at the begin of the Spring '76.

The fig.7 show the general trend of the signal, it resumes the low frequency components or long time oscillations.

level	*wave length (days)*	*energy*
-1	1.4 - 2.8	1.01
-2	2.8 - 5.7	6.13
-3	5.7 - 11.4	12.19
-4	11.4 - 22.8	14.22
-5	22.8 - 45.7	10.33
-6	45.7 - 91.5	10.58

Table 1. *Integral energy for the six levels analyzed.*

The low energy of the most high frequencies, in the level j=-1 tells us that no significant noise is present in the analyzed signal.

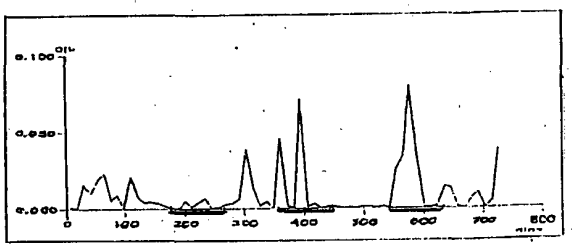


Fig 4. *Wavelet analysis of the FLS for levelj=-4, corresponding to the wave lengths from 11 to 22 days.*

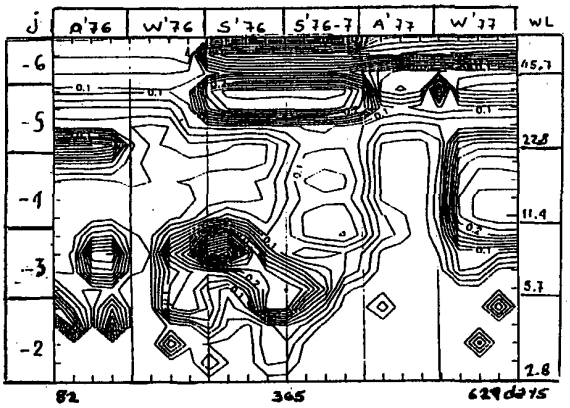


Fig 5. *Map of the time-frequency representation from Autumn '76 to Winter '77, wave lengths from 2 to 90 days*

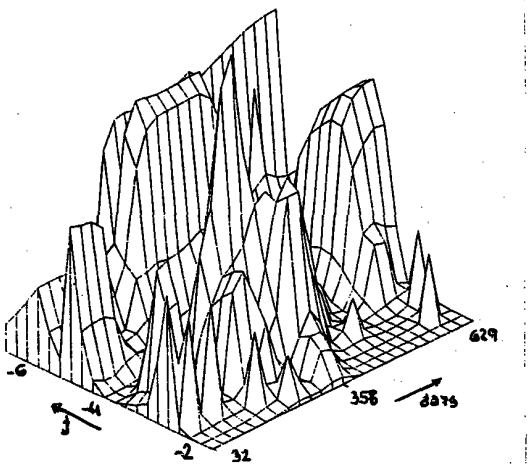


Fig 6. *Topography feature of the time frequency map.*

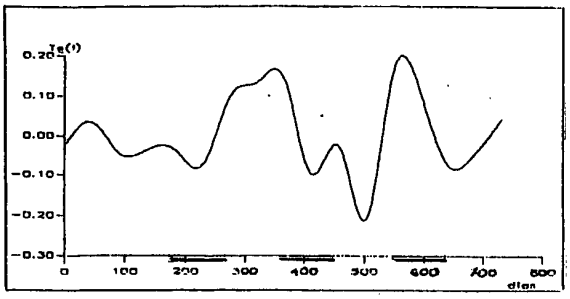


Fig 7. *Trend: residual signal for wave lengths longer than 90 days.*

Prospective

An intersting feature, natural continuation of this work, is to analyze a longer time sequence covering several years in order to determinate some season behavior pattern or to detect eventual chaotic phenomena.

The problem of understanding dynamical system directly of the data is complex, as Mess (1991) and Grassberger (1991) had observed. There have been claims that deterministic chaos has been found in numerous time sequences but these claims have to be taken with very much caution. A deep analysis often is needed. As Arneodo (1991) had proposed, the WT may be effectively applied in that case.

We tried to emphasize the promissory role that the WT could have in Meteorology.

The technique here used is only one in a broad range of available utilities.

References

Arneodo, A., Grasseau, G., Holschneider, M 1988:On the wavelet transform of multifractals. *Phys. Rev. Lett.* 61 (20), 2281-84.

Bischoff, S. and Fernandez, A., 1987: Sobre el comportamiento troposférico en areas argentinas asociadas a la ocurrencia del fenómeno Oscilacion Sur El Niño 1976/1977. *Anales II Congreso Interamericano y V Congremet, C.A.M., Argentina,* 5.4.1 - 5.4.5.

Compagnucci, R. and Salles, M., 1992:Souther South American Anomalies in relation to surface circulation pattern during 1976-1977 ENSO event. *'Paleo ENSO Records' Intern. Symp. (Lima, March 4-7-1992) Ext. Abs., ORSTOM - CONCYTEC,* 44-52.

Farge, M., 1992: Wavelet transform and their applications to turbulence. *Annu. Rev. Fluid Mech.* 24, 395-457.

Grassberger, P.,Schreiber, T. and Schaffrath C., 1991:Nonlinear time sequence anlysis. *Int. J. of Bifurcation and Chaos.* I-3-, 521-547.

Green, P. and Carol, G., 1978:Analyzing multivariate data. *The Dryden Press. Illinois U.S.A.* 519.

Heil, C., Walnut, D., 1990:Continuous and discrete wavelet transform. *SIAM Rev.* 31-4, 628-666.

Mallat, S., 1988:Multiresolution representation and wavelets. *GRSP Lab.* 153.

Meyer, Y., 1990:Ondelettes et operateurs. *Hermann Ed. (Paris)* Tomes I-II-III.

--------- 1987:Wavelets and Operators. *Rapport CEREMADE* 8704. Univ. Paris-Dauphine, Paris-France.

--------- 1989:Ondelettes, filters miroirs en quadrature traitement numeque de l'image. *Gaz. Math.* 40, 31-32.

Pincibono B., 1976:Elements de theory du signal. *Dunon Universite.*

Shannon, C. & Weaver, W., 1949:The mathematical theory of the information. *Univ. Illinois Press.*

Serrano, E., 1991:Construcción de una biblioteca de bases orthonormales de $L^2(\mathbb{R})$ a partir de la transformada wavelet. *Anales de la 4ta Reunion de Trabajo en Procesamiento de la Informacion y control. RPIC'91 (Buenos Aires-Argentina Nov 1991)* 167-174.

COMPARATIVE ANALYSIS OF TIME SERIES OF SPHERICAL HARMONICS BY THE USE OF SSA AND REDUNDANCY CRITERION

Gabriella Csima and Aranka Maller

H-1675 BUDAPEST, P.o.box 32.

Spherical harmonics of meteorological fields contain useful information on the characteristic features of fields and the location and intensity of weather phenomena. By the investigation of the time series of spherical harmonics information can be obtained on climatic fluctuations. This information can be effectively used in long range weather forecasting.

Besides the standard methods of analysing meteorological time series, e.g. harmonic analysis, trend analysis etc., recently some others have been introduced which are in some sence more efficient than the previous ones. One among these is the Singular Spectrum Analysis (SSA), which is based on Karhunen - Loeve theorem and thus it is analogous with the EOF analysis, has been successfully used for analysing the time series of expansion coefficients. This procedure has been introduced by Broomhead and King (1986), who elaborated a method to reconstruct the phase space by the use of experimental time series. The disadvantage of this method is that it is

appropriate only to describe those linear characteristics of the time series which are included within the covariance function. As a following step, it seems to be useful to introduce such methods being able to reproduce nonlinear features of time series. One of these methods is the redundancy criterion analysis (Fraser, 1989), which is based on the utilization of mutual information and seems to be an effective procedure.

Therefore, the aim of this study is to compare the time series of spherical harmonics by using singular spectrum alalysis and redundancy criterion analysis. To this end, 500 hPa hemispheric geopotential fields have been considered. The forty-two-year-old (1949 - 1990) time series of coeficients corresponding to a truncation being optimal on the climatological time scale has been computed.

Redundancy criterion analysis and singular spectrum analysis have been carried out for the time series of daily data as well as for pentade, decade, and monthly averages. These averages have been obtained by averaging spherical harmonics.

The results have been compared by the use of distortion measure (D-field).

The comparison suggest that redundancy criterion analysis is more advantageous than singular spectrum analysis.

CORRECT FORMULA FOR COMPUTATION OF COEFFICIENTS IN HARMONIC ANALYSIS

Xiao-Ping Zhong
Chengdu Research Institute of
Plateau Meteorology
Guang Hua Cun
Chengdu, Sichuan 610072
P. R. C.

ABSTRACT

Harmonic analysis is widely used to analysis of time and space series in many subjects, such as meteorology, hydrology, seismology. There is, however, a severe mistake in the computation formula of harmonic coefficients. The value of the coefficient for the harmonic with a wave length of 2, i. e., the wave length of this harmonic is two times of the interval of the sample, is one time larger than its real value.

KEY WORDS: HARMONIC ANALYSIS, DISCRETIZATION, COEFFICIENT

1. Introduction

Wave-spectrum analysis has been widely applied to various subjects for a many years because of its special advantages. Harmonic analysis, one important branch of wave- spectrum analysis, has also been receiving many attentions due to its clear physical meaning and conveinent computation process. But present formula for computation of coefficient in harmonic analysis is a wrong one for the harmonic with a wave length of 2, although this formula has been using for a long time. The root of this mistake lies in the special property of trigonometric function COS and the method of discretization from continuous integration to discrete summation.

Harmonic analysis is disccused in many books, papers, and teaching materials used in lots of schools and universities. Subjects dealing with this content include statistics, meteorology, hydology, and seismology etc. The same wrong formula is quoted in all materials that the author can find. A correct computation formula of coefficient for the harmonic with a wave length of 2 is derived based on some common principles of mathematics and fumulas. Because the mathematical methodology used here is very simple, and the materials related to the wrong formula is too many, no references are listed at the end of this paper.

2. Basic Principle of Harmonic Analysis

At first, the basic principle of harmonic analysis should be introduced simply in order to locate where is the mistake. It is known from Fourier's thorem that function F(t) under Dirichlet condition can be expanded on an interval (0, 1) into

$$F(t) = A_0 + \sum_{k=1}^{\infty} \left(A_k \cos\left(\frac{2\pi k}{N} t\right) + B_k \sin\left(\frac{2\pi k}{N} t\right)\right). \quad (1)$$

F(t) could be estimated by taking p terms ahead. That is

$$F'(t) = A_0 + \sum_{k=1}^{p} \left(A_k \cos\left(\frac{2\pi k}{N} t\right) + B_k \sin\left(\frac{2\pi k}{N} t\right)\right). \quad (2)$$

The error between F ' (t) and F(t) is

$$\varepsilon(t) = F(t) - F'(t) \quad (3)$$

Appling the method of least-square, the formula for all coefficients could be derived from (3):

$$\begin{cases} A_0 = \frac{1}{N} \int_0^N F(t)\, dt, \\ A_k = \frac{2}{N} \int_0^N F(t) \cos\left(\frac{2\pi k}{N} t\right) dt, \\ B_k = \frac{2}{N} \int_0^N F(t) \sin\left(\frac{2\pi k}{N} t\right) dt, \end{cases} \quad (4)$$

Please note that orthogonality of the trigonometric function, i. e.

$$\int_0^N \sin\left(\frac{2\pi i}{N} t\right) \cos\left(\frac{2\pi j}{N} t\right) dt = 0, \quad (5)$$

$$\int_0^N \cos\left(\frac{2\pi i}{N}t\right)\cos\left(\frac{2\pi j}{N}t\right)dt=0, \quad (i\neq j) \quad (6)$$

$$\int_0^N \sin\left(\frac{2\pi i}{N}t\right)\sin\left(\frac{2\pi j}{N}t\right)dt=0, \quad (i\neq j) \quad (7)$$

and

$$\int_0^N \cos^2\left(\frac{2\pi i}{N}t\right)dt= N/2, \quad (i\neq 0) \quad (8)$$

$$\int_0^N \sin^2\left(\frac{2\pi i}{N}t\right)dt= N/2, \quad (i\neq 0) \quad (9)$$

have been used.

For a discrete random series with constant interval

$$X(t), \quad (t=0, 1, 2, \ldots, N-1) \quad (10)$$

(2) should be written into

$$X'(t)=A_k+\sum_{k=1}^{p}\left(A_k\cos\left(\frac{2\pi k}{N}t\right)+B_k\sin\left(\frac{2\pi k}{N}t\right)\right). \quad (11)$$

If N is an even p= N/2. If N is an odd, p= (N-1) /2. Physical meaningh of (11) is that discret series X(t) can be simulated by using of a linear combination of p harmonics. Based on the results of continuous function, coefficients should be decided by

$$\begin{cases} A_0=\frac{1}{N}\sum_{t=0}^{N-1}X(t), \\ A_k=\frac{2}{N}\sum_{t=0}^{N-1}X(t)\cos\left(\frac{2\pi k}{N}t\right), \\ B_k=\frac{2}{N}\sum_{t=0}^{N-1}X(t)\sin\left(\frac{2\pi k}{N}t\right). \end{cases} \quad (12)$$

(12) appears in all materials that the author has found. In fact, it is not correct for all harmonics. When N/k=2, formula of A_k should have a constant coefficient 1/N, instead of 2/N.

3. Appearance of the Mistake

If F(t) is a continuous function, the process of least-square analysis of (3) is as follows.

$$\int_0^N \varepsilon^2(t)\,dt=\int_0^N (F(t)-F'(t))^2dt$$

$$=\int_0^N \left(F(t)-A_0-\sum_{k=1}^{p}\left(A_k\cos\left(\frac{2\pi k}{N}t\right)-B_k\sin\left(\frac{2\pi k}{N}t\right)\right)\right)2dt$$

$$=\int_0^N F^2(t)\,dt+\sum_{k=1}^{p}A_k^2\int_0^N\cos^2\left(\frac{2\pi k}{N}t\right)dt$$

$$+\sum_{k=1}^{p}B_k^2\int_0^N\sin^2\left(\frac{2\pi k}{N}t\right)dt+\int_0^N A_0^2dt$$

$$-2\int_0^N A_0F(t)\,dt-2\int_0^N F(t)\sum_{k=1}^{p}A_k\cos\left(\frac{2\pi k}{N}t\right)dt$$

$$-2\int_0^N F(t)\sum_{k=1}^{p}B_k\sin\left(\frac{2\pi k}{N}t\right)dt$$

$$+2\int_0^N A_0\sum_{k=1}^{p}A_k\cos\left(\frac{2\pi k}{N}t\right)dt$$

$$+2\int_0^N A_0\sum_{k=1}^{p}B_k\sin\left(\frac{2\pi k}{N}t\right)dt$$

$$+2\int_0^N \left(\sum_{k=1}^{p}A_k\cos\left(\frac{2\pi k}{N}t\right)\right)\left(\sum_{k=1}^{p}B_k\sin\left(\frac{2\pi k}{N}t\right)\right)dt$$

according to orthogonality of the trigonometric function, i. e. (5), (6), (7), (8), and (9), it is obtained that

$$\int_0^N \varepsilon^2(t)\,dt=\int_0^N (F(t)-F'(t))^2dt$$

$$=\int_0^N F^2(t)\,dt+\frac{N}{2}\sum_{k=1}^{p}A_k^2+\frac{N}{2}\sum_{k=1}^{p}B_k^2+N\cdot A_0^2$$

$$-2\int_0^N A_0F(t)\,dt-2\sum_{k=1}^{p}A_k\int_0^N F(t)\cos\left(\frac{2\pi k}{N}t\right)dt$$

$$-2\sum_{k=1}^{p}B_k\int_0^N F(t)\sin\left(\frac{2\pi k}{N}t\right)dt$$

It is demanded that

$$\int_0^N \varepsilon^2(t)\,dt$$

must take the smallest value. So there must be

$$\frac{\partial}{\partial A_0}\int_0^N \varepsilon^2(t)\,dt = 0,$$

$$\frac{\partial}{\partial A_k}\int_0^N \varepsilon^2(t)\,dt = 0, \qquad (k=1, 2, \ldots, p)$$

$$\frac{\partial}{\partial B_k}\int_0^N \varepsilon^2(t)\,dt = 0, \qquad (k=1, 2, \ldots, p)$$

It is easy to find that A0, Ak and Bk should be decided by (4). There is no any doubt about (4) for a continuous function. If (4), however, is discretized into (12) by using of common method, some kind of mistake will be induced because the result of (8) does not hold water in discrete form (8) is correct according to the principle of definite integration. If discretized based on the common method it becomes

$$\int_0^N \cos^2\left(\frac{2\pi i}{N}t\right)dt \approx \sum_{t=0}^{N-1}\cos^2\left(\frac{2\pi i}{N}t\right)$$

$$= 1+\sum_{t=1}^{N-1}\cos^2\left(\frac{2\pi i}{N}t\right)$$

$$= 1+\sum_{t=1}^{N-1}\left(\frac{1}{2}+\frac{1}{2}\cos\left(\frac{4\pi i}{N}t\right)\right)$$

$$= N/2+1/2+\frac{1}{2}\sum_{t=1}^{N-1}\cos\left(\frac{4\pi i}{N}t\right). \qquad (13)$$

According to the summation formula of trigonometric series, the third term at the right end of (13) is equal to

$$\sum_{t=1}^{N-1}\cos\left(\frac{4\pi i}{N}t\right)$$

$$= \frac{\sin((N-1)4\pi i/(2N))\cos(4\pi i/2)}{\sin(4\pi i/(2N))}$$

$$= \frac{\sin(2\pi i-2\pi i/N)\cos(2\pi i)}{\sin(2\pi i/N)}$$

$$= -1. \qquad (14)$$

However, (14) only hold water when $N/i \neq 2$. when $N/i=2$, the result is quite different:

$$\sum_{t=1}^{N-1}\cos\left(\frac{4\pi i}{N}t\right) = N-1,$$

here i is a non- zero integer. Therefore, the correct result of (13) should be

$$\sum_{t=0}^{N-1}\cos^2\left(\frac{2\pi i}{N}t\right) = \begin{cases} N/2. & (N/i\neq 2) \\ N. & (N/i=2) \end{cases}$$

In turn, the correct formula of Ak should be

$$Ak = \begin{cases} \frac{2}{N}\sum_{t=0}^{N-1} X(t)\cos\left(\frac{2\pi k}{N}t\right), & (N/k\neq 2) \\ \frac{1}{N}\sum_{t=1}^{N-1} X(t)\cos\left(\frac{2\pi k}{N}t\right), & (N/k=2) \end{cases} \qquad (15)$$

here k is the k' th harmonic, and N/k is the wave length of it. Therefore, (15) indicates clearly that the computation formula of coefficient for harmonic with a wave length of 2 is different from formula for other harmonics.

4. Root of the Mistake

At the bottom, this mistake comes from the special property of trigonometric function COS, and common method of discretizing. In general, discretization is relized by using of rectangular summation. For a definite integration

$$\int_0^N \cos^2\left(\frac{2\pi k}{N}t\right)dt,$$

N values of the integrand are taken on the interval (0, N) usually. Distances between each values is identical. The discrete form is

$$\int_0^N \cos^2\left(\frac{2\pi k}{N}t\right)dt \approx \sum_{t=0}^{N-1}\cos^2\left(\frac{2\pi k}{N}t\right). \qquad (16)$$

Please note that N is a integer here. The right end of (16) is equal to the left end only if N/k is not equal to 2. When N/k=2, the former is two times of the later. It has been demonstrated by means of the summation method of the

trigonometric function. Here it is going to be explained further from the principle of mathematics by using of a example with N=4.

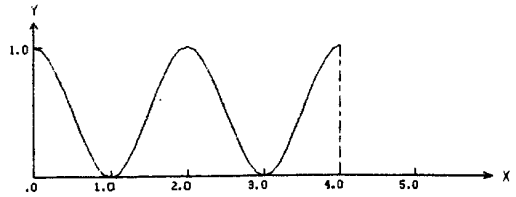


Fig. 1 Continuous integral curve of $COS^2(2\pi*1*t/4)$

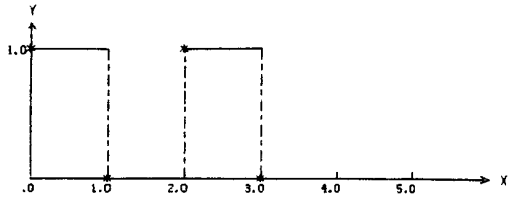


Fig. 2 Discrete values and summation curve of $COS^2(2\pi*1*t/4)$

When N=4, k=1, i.e., $N/k\neq 2$, the curve of the integrand at the left end of (16) is shown in Fig. 1, and the discrete values and its rectangular summation curve is shown in Fig. 2. It is clear that the area surrounded by the continuous integral curve is almost the same as the area surrounded by the discrete summation curve. So (16) does hold water in this case. However, when N=4, k=2, i.e., N/k=2, the curves of the functions at the two ends of (16) draw out two areas with different size (Fig. 3 and Fig. 4). It is easy to find that the area formed by the discrete curve is about one time larger than that formed by the integral curve. In the former situation, the continuous integral curve and the discrete summation curve both have the change of increasing and decreasing. But in the later situation, only the continuous curve increases in one interval and decreases in another interval, and the discrete curve becomes a straight line because of

$$COS^2\left(\frac{2\pi k}{N}t\right)\equiv 1. \qquad (t=0,1,2,\ldots,N-1)$$

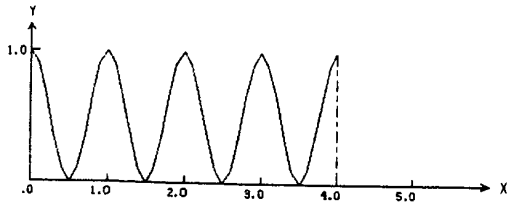


Fig. 3 Continuous integral curve of $COS^2(2\pi*2*t/4)$

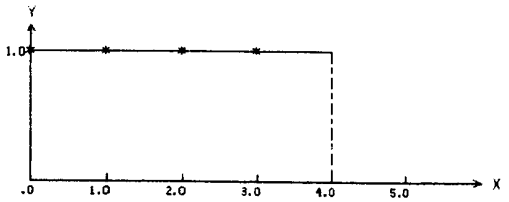


Fig. 4 Discrete values and summation curve of $COS^2(2\pi*2*t/4)$

Therefore, the area surrounded by the discrete summation curve is much larger than that by the continuous integral curve. This is the key reason why the mistake comes into being.

5. Conclusion

A wrong formula for coefficient computation of the harmonic with a wave length of 2 in harmonic analysis has been used for a long time because of ignoring the special property of trigonometric function COS. The value of the wrong coefficient is one time larger than its accurate value. The correct formla for the harmonic with a wave length of 2 should be

$$\begin{cases} A_0=\dfrac{1}{N}\sum_{t=0}^{N-1}X(t)\,dt, \\ A_k=\dfrac{\delta}{N}\sum_{t=0}^{N-1}X(t)\,COS\left(\dfrac{2\pi k}{N}t\right)dt, \quad \delta=\begin{cases}1, & (N/k=2)\\ 2, & (N/k\neq 2)\end{cases} \\ B_k=\dfrac{2}{N}\sum_{t=0}^{N-1}X(t)\,SIN\left(\dfrac{2\pi k}{N}t\right)dt, \end{cases} \qquad (17)$$

instead of (12).

There is no doubt that the use of this wrong formula has been making some wrong conclusions and results in meteorology, hydrology, and seismology, etc. It is possiable for scientists to make the harmonic with a wave length of 2, which may not be obvious, or even not exist in fact, become more remarkable by this formula with a mistake. This kind of harmonic will confuses physical explanations and applications.

Discovery of the mistake does not have any influence on Fourier expansion, or Fourier series of a continuous function. When it is, however, nessceary to discretize a Fourier series and compute their coefficients for a discete function, or a random series, for example in discete mathematics, statistics and other subjects concerned, this discovery bears a special meanings. Because of the special property of the trigonometric functions, it is possiable for similar problems to occure under some kind of conditions. Cares must be taken.

Association Between Time Series Of Various Meteorological Elements In The Pacific Ocean During ENSO Events

Jose A. Maliekal
Dept. of the Earth Sciences
SUNY College at Brockport
Brockport, NY 14420

Introduction

That "El Niño also appears in varying intensity and like any other large-scale event in the ocean-atmosphere system, El Niño certainly does not have only a single cause" was noted by Wyrtki (1975). On the other hand, the evolutionary cycle of the canonical El Niño-Southern Oscillation (ENSO) in Rasmusson and Carpenter (1982) is based on a composite of individual events. Subsequent studies have shown that some of the associations implied in the Rasmusson-Carpenter canonical model are rather loose. The present study delves into similarities and differences between some recent ENSO events.

Methodology

Most of the observed monthly meteorological time series can be treated as the sum of three components, they are the low frequency component, the seasonal component, and the irregular component. Cleveland *et al.*, (1982) proposed an iterative procedure, involving a series of linear and non-linear resistant filters, called SABL (Seasonal Adjustment Bell Laboratories) to decompose time series into individual components. In this study, time series of sea level pressure (SLP) and sea surface temperature (SST) are first decomposed using SABL. The low frequency components of these are then graphically analyzed following the superpose epoch analysis procedure.

To follow up on the results obtained from the above analysis, time-longitude plots of the Outgoing Longwave Radiation (OLR) from the equatorial belt (5°S-5°N) are prepared. As the first step, the component due to the annual march of the sun is removed from the daily OLR data by subtracting the 12 year mean from daily values. Seasonal averages of OLR values are then formed. These averages are centered (by removing the zonal mean values) and standardized (by dividing with the zonal standard deviation). Further more, they are smoothed by a median filter. Time-longitude plots are then prepared.

To validate initial results, multivariate spectrum analysis is employed. Area averaged daily OLR time series from three regions along the equatorial belt are used for this purpose; regions are eastern Indian Ocean, (70-100° E; IO), maritime continent, (102.5-132.5° E; MC), and western Pacific, (135-165° E; WP). Power spectra are estimated by fitting autoregressive models to data; model orders are determined by Akaike's Information Criterion. An estimate of the number of degrees of freedom of an autoregressive model is N/p (Gersch and Goddard, 1970), where N is the length of the time series and p is the model order. In the present study N is 365 and p is less than 4 and hence the number of degrees of freedom is about 90. A conservative estimate of the 95% confidence level for zero spectral coherence and partial spectral coherence with 45 degrees of freedom is 0.07. A similar procedure was used by Gersch and Goddard (1970) to locate the site of epileptic focus.

Results

Fig. 1. shows that during the 1976-77 and 1986-87 events, warm water spread westward from the tropical South American coastline. On the other hand, during the 1982-83 event, the appearance of the warm water was near simultaneous along the coastline as well as along the eastern and central Pacific Ocean. A similar plot of the SLP (Fig. 2) show that peak warming of the coastal surface waters coincided with the peak in the rise in the SLP at Darwin during the 1976-77 event. The decrease in the SLP at Tahiti and Easter Island bottomed only months later. Extremes in the sea level pressure at these three sites occurred months before the coastal waters registered their warming climax.

Time-longitude plots (Fig. 3) show that Prior to the 1976-77 ENSO, the western Pacific remained cloud free, but the maritime continent was cloudy. This pattern flip-flopped in the spring of 1976. Dry conditions continued to persist over the maritime continent throughout 1977 with central Pacific

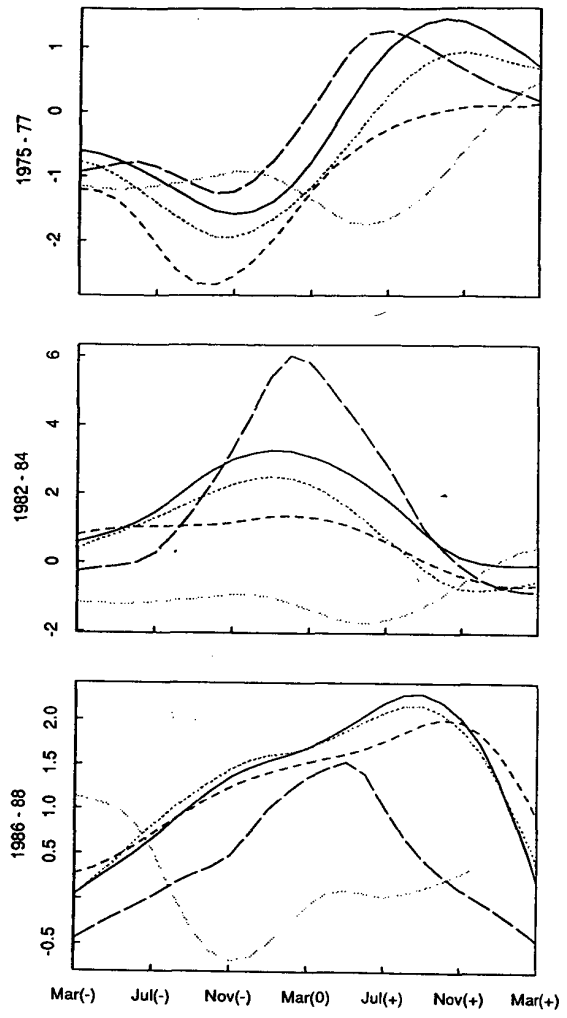


Fig. 1. Standardized low frequency component of SST time series at Puerto Chicama (- -), Niño 3 (—), Central Pacific (····), and Niño 4 (- - -). (-), (0) and (+) represent months before, during and after ENSO events.

experiencing the opposite polarity. No systematic eastward propagation of anomalous convection is evident either prior to or after this ENSO event. The 1982-83 event had a distinctly different behavior. The origin of the anomalous convection associated with this event can be track back to the maritime continent, beginning in the summer of 1981. Once started, the convection propagated systematically eastward.

Previous studies, especially Gutzler and Harrison (1986), have indicated that anomalous low level wind convergence associated with ENSO first appears in the vicinity of Indonesia. Judging from OLR anomalies, this appears to be the case during the 1982-83 event. Recent study by Ramanathan and Collins(1991) have shown that the thermodynamic effect of ocean warming is to enhance deep convection. Cirrus anvil, the byproduct of deep convection subsequently shields the ocean from solar radiation. This inhibits continued warming of the ocean surface, there by reducing convection. Next logical step is to assess the role of the maritime continent, situated between warm waters of the eastern Indian Ocean and western Pacific Ocean, in regulating convection over these warm waters.

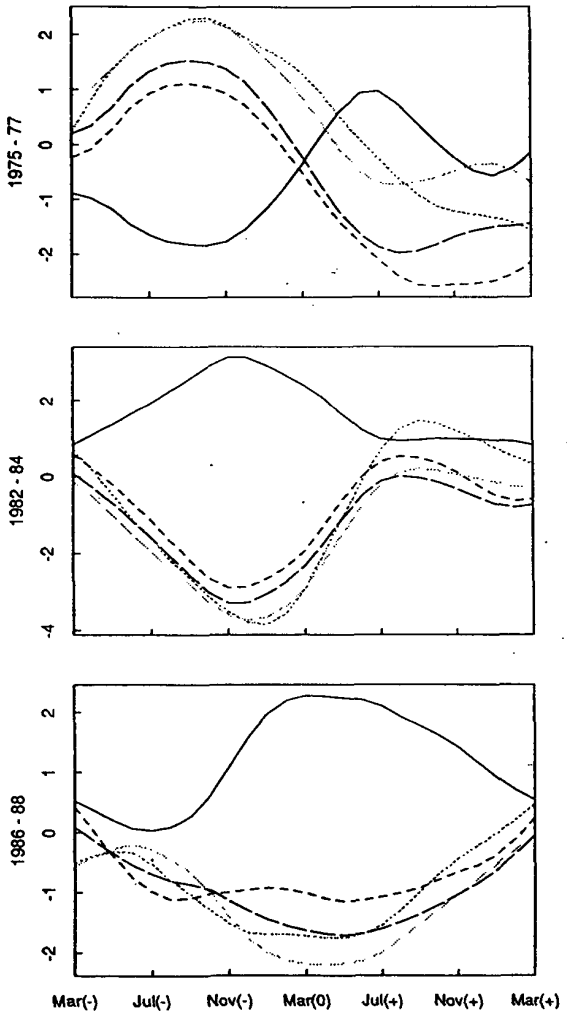


Fig. 2. Same as in Fig. 1, but for SLP at Darwin (—), Tahiti (····), Easter (- -), Tahiti - Darwin (—) and Easter - Darwin (- -).

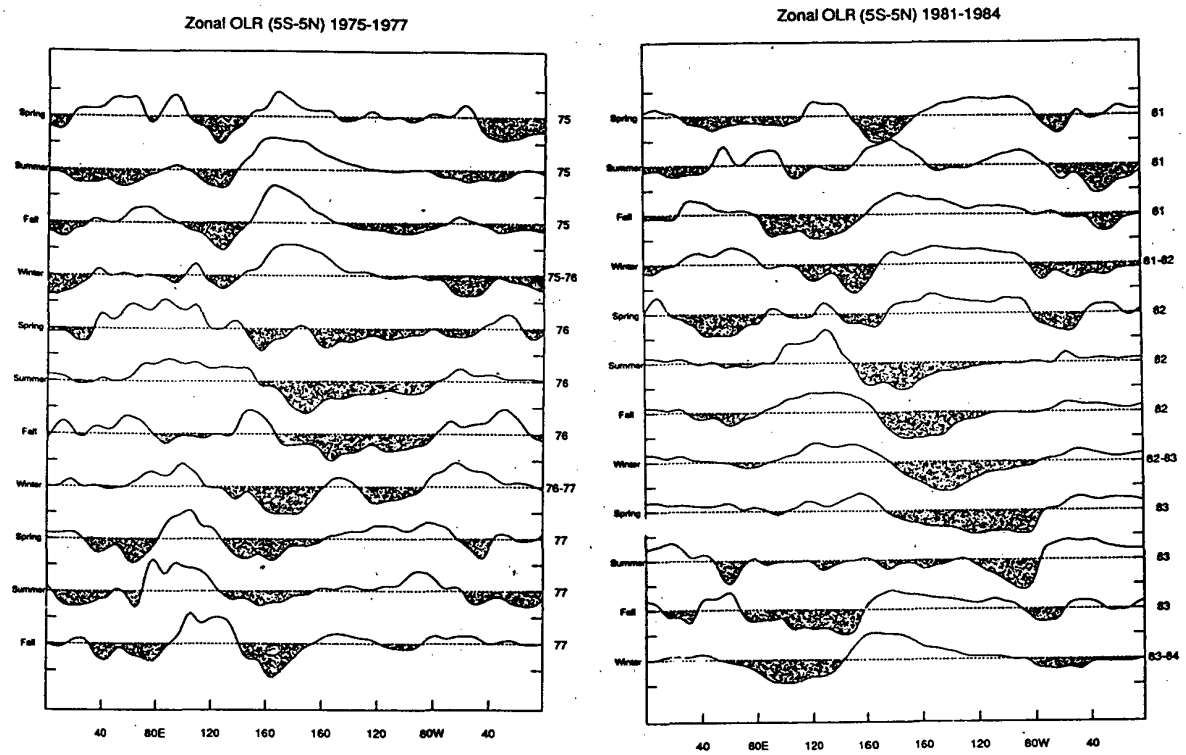


Fig. 3. Time longitude plots of standardized and smoothed OLR from the equatorial belt.

To accomplish the above objective spectral coherence and partial coherence are estimated. Partial coherence is the coherence remaining between two OLR time series after the influence of third OLR time series is removed from each of the first two, between time series. The basic premise of this approach is that convection in one region is causal to two other neighboring regions, if the region in question accounts for the linear association between the other two regions over a frequency interval. Notice that all OLR time series have greatest energy concentrations on the low frequency side (Fig. 4). Partial coherence is not very much different from spectral coherence in most cases; the years 1981 and 1983 are exceptions. During these years there is significant coherence between the OLR time series of the eastern Indian Ocean and western Pacific (Fig. 5). However the partial coherence between these two regions, after the influence of the OLR time series from the maritime continent is removed, is not significantly different from zero. Therefore, the coherent fluctuations in convection over eastern Indian Ocean and western Pacific Ocean during 1981 and 1983 appears to be the result of the influences exerted by the maritime continent. Thus low frequency convective activity over the maritime continent appears to have played a critical role in the initiation as well as the cessation of the 1982-83 ENSO event.

In what follows the principle behind this observed association between maritime continent and neighboring oceans is described. Heating of the maritime continent in conjunction with the convergence of the prevailing surface winds produce deep local convection. The predominant easterly flow of the upper troposphere shears the cirrus anvil to the west. Here the anvil region extends 100 to 1000 km away from the main convective region (Houze and Hobbs, 1982). This anvil temporarily shades the eastern equatorial Indian Ocean from solar radiation. To the east, equatorial western Pacific Ocean is free of dominating cirrus clouds. Continued exposure to solar radiation heats the western Pacific and as a result

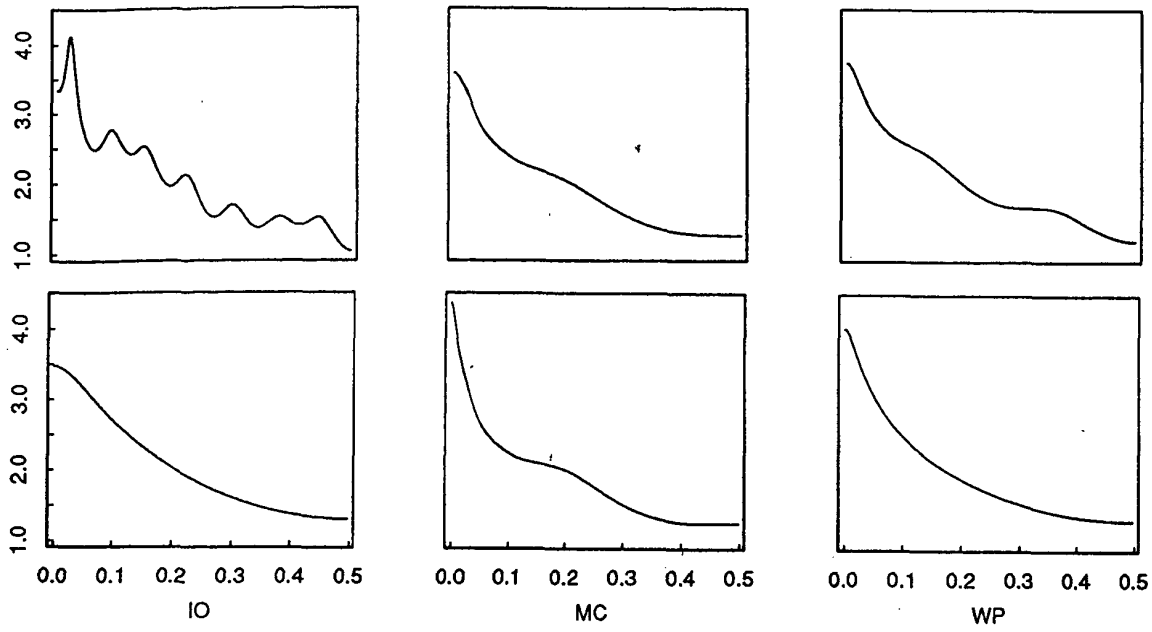


Fig. 4. Power spectra of OLR in the IO, MC, and WP. In this diagram natural logarithm of power is plotted against frequency.

convection increases over the western Pacific. Tops of these clouds are also sheared preferentially to the west. They now shade the maritime continent from solar heating. Consequently, the cirrus shield over the Indian Ocean starts to disappear. Soon this region is subject to more solar heating. Subsequently the convection shifts to Indian Ocean. Cirrus shielding of the maritime continent is lifted and heating of the maritime continent resumes.

References

Gersch W., and G.V. Goddard, 1970: Epileptic focus location: Spectral analysis method. *Science*, **169**, 701-702.

Gutzler, D.S., and D.E. Harrison, 1987: The structure and evolution of seasonal wind anomalies over the equatorial Indian and western Pacific Oceans. *Mn Wea.Rev.*, **115**, 169-192.

Houze, R.A., and P.V. Hobbs, 1982: Organization and structure of precipitating cloud systems. In *Advances in Geophysics*, **24**, 225-315. Academic Press, New York.

Ramanathan, V. and W. Collins, 1991: Thermodynamic regulation of ocean warming by cirrus clouds deduced from observations of the 1987 El Niño. *Nature*, **351**, 27-32.

Rasmusson, E. H. and T. H. Carpenter, 1982: Variations in the tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Mon. Wea. Rev.*, **110**, 354-384.

Wyrtki, K. 1975: El Niño: The dynamical response of the equatorial Pacific to atmospheric forcing. *J. Phys. Oceanogr.*, **5**, 572-584.

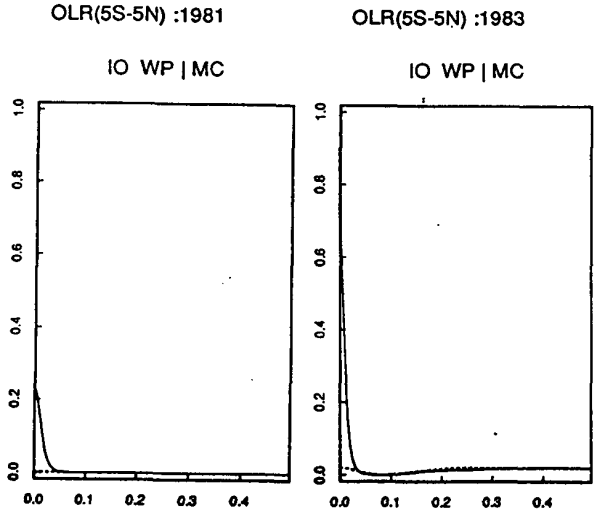


Fig. 5. Coherence (full line) and partial coherence (dashed line) between OLR time series. IO, WPI MC indicates that the plotted spectral coherence is between IO and WP and the plotted partial coherence is coherence after the influence of the MC is removed.

A Test for Differences in Autocorrelation Between Climate Time Series Based on the Jackknife

J.J. Beersma and T.A. Buishand

Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands

1 Introduction

In this paper a procedure is presented for testing for differences in autocorrelation coefficients. The procedure requires the estimated autocorrelation coefficients with their standard errors. It is not necessary to specify the form of the underlying distribution nor to model the autocorrelation structure.

First the estimation of autocorrelation coefficients for seasonally varying data is considered. The jackknife is introduced as a technique to reduce bias and to obtain standard errors of the estimated autocorrelation coefficients. This leads to a simple statistic for testing for differences in autocorrelation coefficients of two independent time series. The method is then extended to test for a regional change in autocorrelation with data from several locations. The use of the test is illustrated with an observed record of daily temperatures in the Netherlands and with data generated by the Canadian Climate Centre (CCC) second generation GCM for the present CO_2 level and a doubling of the CO_2 level.

2 Estimation of autocorrelation coefficients

For ease of presentation it is assumed that time series of daily values are available. The lag k autocorrelation coefficient, ρ_k, then represents the correlation between two values separated by an interval of $k-1$ days. The estimation of the autocorrelation coefficients should be done with some care, because the systematic annual cycle may introduce serious bias. A popular method is to estimate the autocorrelation coefficients for each calendar month separately. Let $x_{i,j}$ denote the value on day i ($i = 1, \ldots, n$) of a certain month in year j ($j = 1, \ldots J$). Then the lag k autocovariance is estimated as

$$c_k = \frac{1}{N_k} \sum_{j=1}^{J} \sum_{i=1}^{n-k} (x_{i,j} - \overline{x})(x_{i+k,j} - \overline{x}), \quad k = 0, 1, \ldots \tag{1}$$

where $\overline{x}$ is the mean over all daily values

$$\overline{x} = \frac{1}{nJ} \sum_{j=1}^{J} \sum_{i=1}^{n} x_{i,j} \ . \tag{2}$$

Quite often N_k is set equal to the total number of days, $N_k = nJ$ for all k (Katz 1982; Trenberth 1984). In this paper, however, $N_k = (n-k)J$, which results in less biased estimates. For $k = 0$, c_k represents the sample variance. The estimated autocorrelation coefficients are obtained as

$$r_k = c_k / c_0 \ , \quad k = 0, 1, \ldots \ . \tag{3}$$

In the jackknife method the required statistic is recomputed a large number of times after successive deletion of a group of observations from the entire data set (Efron 1982). To obtain a simple test it is necessary that the data in different groups are almost independent. For the climatic data considered here, it is therefore most appropriate to delete each time a complete year. The jackknife estimate of ρ_k then reads

$$r_{k,JACK} = J r_k - (J-1) r_{k(\cdot)} \ , \quad k = 0, 1, \ldots \tag{4}$$

where

$$r_{k(\cdot)} = \frac{1}{J} \sum_{j=1}^{J} r_{k(j)} \ , \tag{5}$$

with $r_{k(j)}$ the estimate of ρ_k after omitting the data for year j. The use of the jackknife results in almost unbiased estimates of ρ_k. Table 1 demonstrates this for 10–year samples of 30 daily values generated by a normal first-order autoregressive (AR1) process with $\rho_1 = 0.8$. For the AR1-process $\rho_k = \rho_1^k$, $k = 0, 1, \ldots$.

Table 1: Empirical means of the estimates $r_{k,JACK}$ and r_k of the autocorrelation coefficient ρ_k for an AR1-process with $\rho_1 = 0.8$ ($J = 10$, $n = 30$, 5000 simulations).

lag	ρ_k	$E(r_{k,JACK})$	$E(r_k)$
1	0.800	0.800	0.789
2	0.640	0.641	0.621
5	0.328	0.327	0.297
10	0.107	0.107	0.076

To derive the jackknife estimate, it is convenient to rewrite Eq. 1 as

$$c_k = \frac{1}{N_k}\{P_k - S[S_{L,k} + S_{R,k} - (n-k)S/n]/(nJ)\}, \qquad k = 0, 1, \ldots \qquad (6)$$

where

$$P_k = \sum_{j=1}^{J}\sum_{i=1}^{n-k}(x_{i,j}x_{i+k,j}), \quad S = \sum_{j=1}^{J}\sum_{i=1}^{n} x_{i,j},$$

$$S_{L,k} = \sum_{j=1}^{J}\sum_{i=1}^{n-k} x_{i,j}, \quad S_{R,k} = \sum_{j=1}^{J}\sum_{i=1}^{n-k} x_{i+k,j}.$$

The estimate $r_{k(j)}$ in the right-hand side of Eq. 5 is then easily obtained by subtracting the contribution of year j from the quantities P_k, S, $S_{L,k}$ and $S_{R,k}$.

3 Standard errors of estimated autocorrelation coefficients

The jackknife is not only a method of bias reduction. The recomputed statistics $r_{k(j)}$ can also provide a distribution-free estimate of the standard error. The jackknife estimate of the variance of r_k is given by (Efron 1982)

$$\hat{\sigma}^2_{k,JACK} = \frac{J-1}{J}\sum_{j=1}^{J}[r_{k(j)} - r_{k(\cdot)}]^2. \qquad (7)$$

For the generated 10-year samples of the AR1-process in the previous section, Table 2 compares the mean of $\hat{\sigma}_{k,JACK}$ with the standard deviations $\sigma(r_{k,JACK})$ and $\sigma(r_k)$ of $r_{k,JACK}$ and r_k. The table also presents these results for an AR2-process with the same value of ρ_1 as the AR1-process (0.8) but with a much smaller value of ρ_2 (0.45 instead of 0.64). The statistic $\hat{\sigma}_{k,JACK}$ yields an almost unbiased estimate of $\sigma(r_k)$. It slightly underestimates $\sigma(r_{k,JACK})$, in particular for large lags. This is consistent with Efron (1982) where Eq. 7 is considered to provide an estimate of $\mathrm{var}(r_k)$ rather than of $\mathrm{var}(r_{k,JACK})$.

4 Testing for differences at a single site

Table 2 shows that the estimated lag 1 autocorrelation coefficient has a much smaller standard error than the higher order autocorrelation coefficients. A systematic difference in the values of ρ_1 for two climates (e.g. a $1\times CO_2$ run and a $2\times CO_2$ run) is therefore better detectable than differences in the other autocorrelation coefficients. Let $\rho_1(\mathrm{I})$ and $\rho_1(\mathrm{II})$ be the theoretical lag 1 autocorrelation coefficients for Climate I and Climate II respectively. A test for differences in autocorrelation can then be based on their jackknife estimates $r_{1,JACK}(\mathrm{I})$, $r_{1,JACK}(\mathrm{II})$ and the jackknife estimates $\hat{\sigma}_{1,JACK}(\mathrm{I})$, $\hat{\sigma}_{1,JACK}(\mathrm{II})$ of the standard errors of the autocorrelation estimator from Eq. 7. A possible test-statistic is

$$T_1 = \frac{r_{1,JACK}(\mathrm{I}) - r_{1,JACK}(\mathrm{II})}{\left[\hat{\sigma}^2_{1,JACK}(\mathrm{I}) + \hat{\sigma}^2_{1,JACK}(\mathrm{II})\right]^{1/2}}. \qquad (8)$$

A similar statistic was used by Miller (1968) to test for differences between the variances of two independent samples. Under the null hypothesis, $\rho_1(\mathrm{I}) = \rho_1(\mathrm{II})$, the statistic T_1 is approximately a Student-t variable with d degrees of freedom, where d is given by

$$d = \frac{(J+K)^2}{K^2/(J-1) + J^2/(K-1)} \qquad (9)$$

with J and K the number of years for Climate I and Climate II respectively. For details and alternative test-statistics, see Buishand and Beersma (1992). For non-integer d the t-distribution is still defined through its relation with the beta distribution. The critical values of the test-statistic can be obtained numerically (Gardiner and Bombay 1965).

The quality of the approximation to the null distribution of T_1 depends on J, K, the group size n and the probability distribution of the climate variable involved. A Monte Carlo experiment with the normal AR-processes in Table 2 showed that for $n = 30$ the proposed t-distribution performs quite well, even for J and K as small as 5 (Buishand and Beersma 1992). The test is rather robust against departures from normality. For samples from a very skewed exponential AR1-process the use of the t-distribution turned out to be dubious.

Table 2: Empirical mean of the estimate $\hat{\sigma}_{k,JACK}$ compared with the empirical standard deviations of $r_{k,JACK}$ and r_k for an AR1-process with $\rho_1 = 0.8$ ($\rho_2 = 0.64$) and an AR2-process with $\rho_1 = 0.8$ and $\rho_2 = 0.45$ ($J = 10$, $n = 30$, 5000 simulations).

	AR1-process			AR2-process		
lag	$E(\hat{\sigma}_{k,JACK})$	$\sigma(r_{k,JACK})$	$\sigma(r_k)$	$E(\hat{\sigma}_{k,JACK})$	$\sigma(r_{k,JACK})$	$\sigma(r_k)$
1	0.038	0.039	0.038	0.022	0.023	0.023
2	0.065	0.067	0.065	0.053	0.055	0.054
5	0.108	0.115	0.107	0.097	0.102	0.098
10	0.134	0.147	0.135	0.112	0.117	0.113

5 Increasing the power of the test by combining several samples

For GCM runs of limited length the statistic T_1 can only detect quite substantial differences in autocorrelation. For two normal AR1-processes with $\rho_1(\mathrm{I}) = 0.8$ and $\rho_1(\mathrm{II}) = 0.7$ the probability of a significant result at the 5%-level is about 0.35 if $J = K = 10$ (group size $n = 30$). A more powerful test is possible by combining the samples at several grid points within a region. Regional analogs of the test-statistic T_1 can be derived as follows. For each grid point the estimates $r_1, r_{1(1)}, \ldots, r_{1(J)}$ are calculated. Taking averages over the grid points results in $\bar{r}_1, \bar{r}_{1(1)}, \ldots, \bar{r}_{1(J)}$. The jackknife estimates in the test-statistic T_1 are then obtained by applying Eqs. 4, 5 and 7 to these average autocorrelation estimates. By a similar averaging procedure one can get a single test-statistic for the December, January, February (DJF) period and other seasons, as well as for the whole year. Averaging over grid points and/or months leads to an increase in power because of a reduction of the standard error. A simultaneous test on the results for individual calendar months is also possible using the binomial distribution or by means of Fisher's test (Sneyers 1990). These methods, however, assume independence among the values of T_1 in successive months.

6 Results for daily mean surface air temperatures

In this section the proposed method is illustrated with the 1951–1980 temperature record at De Bilt (52°06′N, 5°11′E) and generated time series over Europe of the CCC GCM. Ten-year samples for both the $1\times CO_2$ and $2\times CO_2$ simulated climates were made available over a grid of 3.75° × 3.75°.

Table 3 presents autocorrelation estimates for the De Bilt record and the GCM $1\times CO_2$ run at two grid points. For the GCM data the autocorrelation coefficients are systematically lower than those for the observed temperatures. This leads to very significant differences in lag 1 autocorrelation coefficients averaged over the year. For individual seasons the underestimation is significant for both grid points in winter (DJF) and autumn (SON) and for the grid point west of De Bilt in spring (MAM).

Table 3: Tests for differences in the lag 1 autocorrelation coefficients ρ_1 of daily temperature observations at De Bilt (1951-1980) and the $1\times CO_2$ run of the CCC GCM (10 years). B refers to De Bilt, W and E refer to the nearest grid point west and east of De Bilt respectively.

statistic	DJF	MAM	JJA	SON	year
$r_{1,JACK}$(B)	0.83	0.82	0.78	0.78	0.80
$r_{1,JACK}$(W)	0.73	0.73	0.73	0.69	0.72
T_1(B,W)	-2.88*	-2.18*	-1.82	-3.42†	-4.98†
$r_{1,JACK}$(E)	0.75	0.77	0.73	0.71	0.71
T_1(B,E)	-2.58*	-1.58	-1.69	-2.54*	-4.47†

* (†) differences significant at the 5% (1%)-level.

To test for differences in autocorrelation between the generated $1\times CO_2$ and $2\times CO_2$ temperature data an area of 25 grid points around the Netherlands (extending from 40.8° to 59.4° N latitude and 5.6° W to 13.1° E longitude) was considered. The area includes Great Britain, Denmark, France, Germany and parts of Spain and Italy. Sixteen grid points have been designated as land points, the other nine are sea points. Land and sea points were treated seperately. The autocorrelation estimates are summarized in Table 4. Both for the land and the sea points the standard error of the autocorrelation estimate is reduced by a factor 1.7 as a result of spatial averaging. This reduction is comparable to that obtained by averaging over

three consecutive calender months. Table 4 shows that the statistics based on the regional seasonal averages are only significant for the land points in winter (DJF) and for the sea points in summer (JJA). According to the binomial distribution the joint result for the four seasons is not significant at the 5 %-level. There is also no statistical evidence of differences in the annual mean of the lag 1 autocorrelation coefficient.

Table 4: Tests for differences in the lag 1 autocorrelation coefficients ρ_1 of daily temperature between the $1\times CO_2$ (C) and $2\times CO_2$ (D) runs of the CCC GCM (both 10 years) for an area of 25 grid points around the Netherlands.

statistic	DJF	MAM	JJA	SON	year
	Land grid points (16)				
$r_{1,JACK}$(C)	0.70	0.72	0.74	0.68	0.71
$r_{1,JACK}$(D)	0.63	0.70	0.76	0.70	0.70
T_1	-2.63*	-0.37	0.77	0.92	-0.85
	Sea grid points (9)				
$r_{1,JACK}$(C)	0.81	0.84	0.87	0.86	0.85
$r_{1,JACK}$(D)	0.85	0.84	0.91	0.88	0.87
T_1	1.00	-0.19	2.26*	1.23	1.49

* differences significant at the 5%–level.

7 Discussion

In the present paper it is demonstrated that the jackknife can be used to reduce the bias in autocorrelation estimates of climate time series and to calculate the standard errors of these estimates. The examples in Tables 3 and 4 show that the test based on the jackknife performs quite well. From Table 3 it is seen that this test is able to detect differences in the range of 0.05 to 0.10 between the seasonal autocorrelation estimates of the 30–year observed temperature record and a 10–year GCM simulation. Table 4 indicates that changes in autocorrelation larger than 0.05 would generally be significant at the 5 %–level when 10–year samples over a region of $\sim 2.5 \times 10^6$ km^2 are combined.

The method is not suitable for very short GCM simulations. For small sample sizes, Student's t-distribution does not always provide a good approximation to the critical values of the test-statistic T_1. But, even if this distribution can be used, the critical values will be large as a result of lack of degrees of freedom. Unfortunately, averaging of monthly values does not lead to more degrees of freedom. Because of this and the relatively large standard errors the test is then unable to detect meaningful differences in autocorrelation.

The jackknife requires much less computer time than other resampling techniques, especially when use is made of Eq. 6 to obtain the various autocorrelation estimates. Other statistical properties of climatic time series can also be investigated with this technique. In particular, a jackknife estimate of the variance is easily obtained from the estimation procedure of the autocorrelation coefficients. Tests for differences in variability can then be conducted in the same way. It is sometimes possible, however, to obtain more powerful tests by removing the autocorrelation from the data first ('prewhitening'). A paper on this topic is in preparation.

Acknowledgment. We would like to thank the Canadian Climate Centre for providing the GCM output.

References

[1] Buishand, T.A. and J.J. Beersma, 1992: Tests for differences in autocorrelation between climate time series based on the jackknife. Submitted to *Journal of Climate.*

[2] Efron, B., 1982: *The Jackknife, the Bootstrap and other Resampling Plans.* Society for Industrial and Applied Mathematics, 92 pp.

[3] Gardiner, D.A., and B.F. Bombay, 1965: An approximation to Student's t. *Technometrics,* **7**, 71–72.

[4] 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.

[5] Miller, R.G., 1968: Jackknifing variances. *Ann. Math. Statist.*, **39**, 567–582.

[6] Sneyers, R., 1990: *On the Statistical Analysis of Series of Observations.* World Meteorological Organization, Tech. Note No. 143, WMO No. 415, 192 pp. (original French version published in 1975).

[7] Trenberth, K.E., 1984: Some effects of finite sample size and persistence on meteorological statistics Part I: Autocorrelations. *Mon. Wea. Rev.*, **112**, 2359–2368.

A Vector Time Series Modeling Approach with Application to Seasonal Precipitation in Florida

Pao-Shin Chu[1], Ping Ding[1], and Richard W. Katz[2]
[1]Dept. Meteorology, Univ. Hawaii; [2]ESIG/NCAR, Boulder, Colorado

Introduction

The large scale atmospheric and oceanic circulation anomalies over the Pacific Ocean (ENSO) have been known to affect climate in Florida (Horel and Wallace, 1981; Douglas and Englehart, 1981; Ropelewski and Halpert, 1986; Brenner, 1991). If the ENSO-precipitation teleconnections are valid, the natural questions would be how to identify the nature of their statistical relation and how to use this relation to improve long-range forecasts?

Motivated by the recent development of vector time series models (Tiao and Box, 1981), it is tempting to apply these methods to address the aforementioned questions. In this paper, the regional, seasonal precipitation over southern and central Florida (i.e., spatial average of standardized precipitation anomalies at 9 locations) and the Southern Oscillation Index (SOI) are modeled jointly in the general framework of multivariate autoregressive (AR) models. As one example of how these models could be used, seasonal precipitation variations have been forecast.

Vector Time Series Modeling

Just like the univariate case (Box and Jenkins, 1976), there are three basic stages for multivariate time series modeling; namely, model identification, parameter estimation and diagnostic checking. This strategy is an iterative procedure of model building to ensure satisfactory model fitting and utilization. The following are brief reviews of the techniques used in these three stages. Details are given in Tiao and Box (1981).

The vector $\mathbf{X}_t$ is a K-dimensional AR(p) process if

$$\boldsymbol{\phi}_p(B)\mathbf{X}_t = \mathbf{a}_t \qquad (1)$$

where

$$\boldsymbol{\phi}_p(B) = \mathbf{I} - \boldsymbol{\phi}_1 B - \ldots - \boldsymbol{\phi}_p B^p$$

are matrix polynomials in B, and B is the backward shift operator. The $\boldsymbol{\phi}$'s are K x K unknown parameter matrices, and $\mathbf{a}_t$ is a series of white noise vectors, independently and identically distributed as $\mathbf{N}(\mathbf{0},\boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is the covariance matrix. In (1), p is the order of the vector AR process.

Given data in the form of a time series, the first step is to identify the structural form of the model. In this regard, two fundamental statistics are employed to help specify the model structure. The first one is the cross correlation matrix $\rho(k)$, where k is the lag. The second statistic is the partial cross correlation matrix $\mathbf{P}(k)$. It is defined in terms of the cross covariance matrix $\boldsymbol{\Gamma}(k)$ as

$$\mathbf{P}(k)=\begin{cases} \boldsymbol{\Gamma}(1)\boldsymbol{\Gamma}^{-1}(0) & k=1 \\ [\boldsymbol{\Gamma}(k)-\mathbf{c}_k'\mathbf{A}_k^{-1}\mathbf{b}_k] \\ (\boldsymbol{\Gamma}(0)-\mathbf{b}_k'\mathbf{A}_k^{-1}\mathbf{b}_k)^{-1} & k>1 \end{cases}$$

where the superscript prime and -1 denote the transpose and the inverse of a matrix, respectively, and

$$\mathbf{A}_k = \begin{bmatrix} \Gamma(0) & \cdots & \Gamma(k-2) \\ \vdots & \ddots & \vdots \\ \Gamma'(k-2) & \cdots & \Gamma(0) \end{bmatrix}$$

$$\mathbf{b}_k = \begin{bmatrix} \Gamma(k-1) \\ \vdots \\ \Gamma(1) \end{bmatrix} \quad \text{and}$$

$$\mathbf{c}_k = \begin{bmatrix} \Gamma'(1) \\ \vdots \\ \Gamma'(k-1) \end{bmatrix}$$

If $\mathbf{X}_t$ follows a vector AR(p) model, then all the elements of $\mathbf{P}(k)$ will be zero for $k>p$.

When the sample cross correlation matrix (CCM) is displayed in sequence as a function of lag k, the slow decrease in the matrix elements would imply that the data are consistent with a vector AR process. Meanwhile, the rapid decrease of the sample partial cross correlation in magnitude implies an AR process is an appropriate model for the data.

We use the Akaike Information Criteria (Akaike, 1974) to determine the order p of the vector AR process. In general, a model with a minimum AIC value is preferred to others.

After the order of the vector model is tentatively identified, estimates of the associated parameter matrices can be determined (Brockwell and Davis, 1990). Specifically, the ϕ's and Σ for an order p can be calculated recursively using the multivariate version of the Durbin-Levinson algorithm.

Following the identification and parameter estimation procedures is the model diagnostic checking. If the selected model fits the joint series adequately, the residuals should be serially uncorrelated in both autocorrelation and cross correlation terms. A check on this property is made through cross correlation matrices of the residuals.

Forecasting can be made, say, at time origin t for lead time ℓ as

$$\hat{\mathbf{X}}_t(\ell) = \sum_{i=1}^{p} \phi_i \mathbf{X}_{t+\ell-i} + \mathbf{a}_{t+\ell}$$

The forecasted values are substituted for $\mathbf{X}$ values that have not yet been observed, and error terms that have not yet been observed are set equal to zero.

Applications

As indicated in Table 1, the sample autocorrelation function of the SOI decays slowly and becomes negative at lags five and six, whereas the autocorrelation function of the precipitation series decreases more rapidly with the increasing lags. Significant, negative cross correlations are noted at lags one (-0.25) and two (-0.22) seasons, with the SOI leading precipitation. Hence, when SOI is negative and large (ENSO-like condition), more precipitation in Florida is expected in the subsequent two seasons. On the other hand, negative cross correlations are found at lags one, two, and three seasons, with precipitation leading SOI. These values range from -0.27 at lag one to -0.15 at lag three. This result suggests a possible

feedback relationship between the two variables, although it is not physically clear how changes in Florida precipitation may affect the SO.

Table 1. Cross Correlation Matrices (CCM) and Partial Cross Correlation Matrices (PCCM) for the seasonal SOI and precipitation time series: Winter 1936 through Fall 1983. A plus sign indicates a value greater than $2T^{-1/2}$, a minus sign a value less than $-2T^{-1/2}$ and a dot a value in-between. T is the data length. For the PCCM, element estimates are divided by their corresponding standard errors to obtain standardized coefficients (SC). A plus sign denotes a SC greater than 2, a minus sign a SC is less than -2, and a dot for values in-between.

Lag k	CCM	SC
1	+ - - +	+ - - +
2	+ - - .	. . . .
3	+ . - .	- . . .
4	. . . .	. . . +
5	- . . .	. . . .
6	- + . .	. . . .
7	. + . .	. . . .
8	. + . .	. . . .
9	. + . .	. . . .
10	. . . .	. . . .
11	. . . -	. . . .
12	. . . .	. . . .
13	. . + .	. . . .
14	. . . .	. . . .
15	. . + .	. . . .

Of major interest here in Table 1 is that the elements in the CCM decrease slowly with increasing lags and significant values can be found at relatively large lags, for example, at lag 15. This behavior suggests that the joint series might be described by a purely AR process. For the partial cross correlation matrix, all the elements after lag four are small compared to their estimated standard errors. This cut-off behavior suggests the possibility of a low-order vector AR process.

Table 2 shows the results of the AIC. A vector AR(3) process exhibits the lowest AIC value. Based on both AIC and parsimonious considerations, a two-dimensional AR(3) process appears to be more appropriate than the others.

Table 2. Summary of vector model order determination.

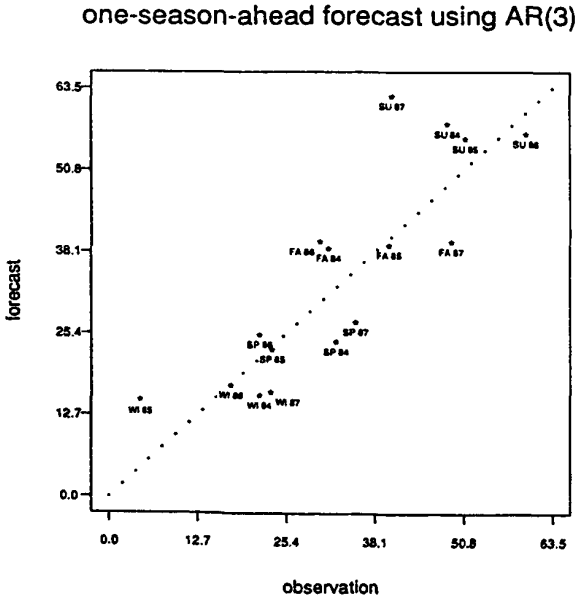
LAG	AIC
1	0.4263
2	0.4632
3	0.3780
4	0.3895
5	0.4036
6	0.4232

Cross correlation matrices of residuals after fitting vector AR(1), AR(2), and AR(3) processes are examined. After AR(1) and AR(2) fit, significant values still appear at various lags. However, after AR(3), the residual pattern is close to a random process, suggesting there is no need to

go beyond AR(3).

To illustrate the utility of the models, one-season-ahead forecasts are made using the bivariate AR(3) model. In order to compare forecasts with actual precipitation amounts (in cm), the normalized value is multiplied by the seasonal standard deviation and then added to the corresponding seasonal mean. Fig. 1 shows that the forecasts match the observations reasonably well, as most points are not far away from the perfect forecast line. Furthermore, the forecasts do not tend to systematically overestimate or underestimate the observations as the number of points on both sides of the perfect line is nearly equal. The correlation between forecasts and observations is 0.86, significant at the 99% level.

Fig. 1. Scatter diagram between observed and predicted seasonal precipitation from a bivariate AR(3) model for the period between winter 1984 to fall 1987. Units are in cm. Broken line denotes 45° angle.



Acknowledgments

This research has been supported by the USDA-Forest Service under contracts PSW-90-0001CA and PSW-91-0002CA. NCAR is sponsored by the National Science Foundation.

References

Akaike, H., 1974: A new look at the statistical model identification, IEEE Trans. Automatic Control AC-19: 716-722.

Brockwell, P.J., and R.A. Davis, 1990: Time series: Theory and Methods, 2 ed., Springer-Verlag, 577pp.

Box, G.E.P., and G.M. Jenkins, 1976: Time series analysis: Forecasting and control (rev.), Holden-Day, 575 pp.

Brenner, J., 1991: Southern Oscillation anomalies and their relationship to wildfire activity in Florida. Int. J. Wildfire, 1, 73-78.

Douglas, A.V., and P.J. Englehart, 1981: On a statistical relationship between autumn rainfall in the central equatorial Pacific and subsequent winter precipitation in Florida. Mon. Wea. Rev., 109, 2377-2382.

Horel, J.D., and J.M. Wallace, 1981: Planetary scale atmospheric phenomena associated with the Southern Oscillation. Mon. Wea. Rev., 109, 813-829.

Ropelewski, C.F., and M.S. Halpert, 1986: North American precipitation and temperature patterns associated with the El Nino Southern Oscillation (ENSO). Mon. Wea. Rev., 114, 2352-2362.

Tiao, G.C., and G.E.P. Box, 1981: Modeling multiple time series with applications. J. Amer. Statis. Asso., 76, 802-816.

Global Warming: A Nonparametric Statistical Approach

S. JAIN
Department of Industrial Engineering
University of Toronto
Toronto, Ontario, Canada M5S 1A4

R. K. JAIN
Department of Mathematics and Statistics
Memorial University of Newfoundland
St. John's, Newfoundland, Canada A1C 5S7

Abstract

The problem of global warming, greenhouse and urban heat-island effect, etc. are some of the most debated subjects by the environmental scientists, public and politicians. In this paper, we review nonparametric statistical procedures which could be applied to detect the changing trend of these characteristics. A numerical example will be used to illustrate its application.

1 Introduction

The objective of this paper is to review nonparametric statistical methods which can be applied to detect the global warming trend, harmful effects of pollutants on humans, and the "greenhouse effect." Meadows, et al. (1974) in their report, "The Limits to Growth," investigated in a world model five major trends of the global problem-accelerated industrialization, rapid population growth, widespread malnutrition, depletion of nonrenewable resources, and deterioration of the environment. The degenerating environment was linked to pollution. They discussed the following basic points in terms of a global perspective:

(a) The pollution (such as mercury in ocean fish, lead particles in the city air, mountains of urban trash and oil slicks on beaches) increased exponentially over a period of time.

(b) Upper limits of these pollutants.

(c) Harmful effects of pollutants.

For example, carbon dioxide (CO_2) pervaded the atmosphere by burning fossil fuels such as coal, oil, and natural gas. Meadow's et al. (1974) further stated in their report that the atmospheric waste and heat around cities caused the formation of "Urban Heat-Islands." A recent article published in the *Globe and Mail* (February 15, 1992), used the heat-islands effect to explain the warming trend of metro Toronto.

It is well-known that an increase in Thermal pollution will have serious climatic effects. Further, the increase in organic wastes results in a catastrophic effect on fish life. The increasing amount of industrial and municipal wastes enter the water system and decrease the dissolved oxygen content in water. The question that is raised by many environmental scientists is what will be the upper limit of these pollutants? This is a difficult question to answer! However, we can apply our current knowledge to extrapolate the future upper limits and then somehow we shall try to control these pollutants.

It is now known that air pollution is linked to major diseases such as emphysema, bronchitis, asthma, and lung cancer. Jain (1986) used the Markov Chain model to demonstrate that the severity of asthma is dependent to weather.

This paper studies the application of nonparametric statistical methods to determine how the trend of global warming can be detected. The changes in climate conditions over a specified period of time may be considered as a random variable. The specific parametric model required several assumptions. For example, to apply the first-order autoregressive process, we

must assume that the errors are normally distributed. However, the use of nonparametric tests will eliminate such problems. Ali and Giaccotto (1982) discussed several nonparametric procedures and found that they were locally most powerful. Moreover, these tests have high asymptotic efficiencies compared to the most favourable parametric tests under some specific alternatives.

Section 2 discusses the developments of linear rank-test statistics. In this section, Wilcoxon, Van Der Waerden (or Normal Score), Savage, and Maximum Efficiency Robust tests are obtained in a natural fashion from the linear rank tests. Finally, in section 3, the temperature-changes data of metro Toronto is used to illustrate the application of these tests. The temperature-changes in metro Toronto are employed to explain the heat-island effect which results in warming trend.

2 Rank-Tests

Page (1954) remarked that the sequence of observations $X_1, \ldots, X_n$ could be divided into subsets, each of which could be regarded as a random sample from a common distribution. These subsamples may be considered as obtained under different conditions. For example, the warming trend can be due to the condition of the excess release of CO_2 in atmosphere and industrial pollution, etc. Hâjek (1969) stated that linear rank statistic could be used to decide whether the variation in condition caused variation in distribution of individual observation or not.

Let $X_1, X_2, \ldots, X_N$ be independent random variables with continuous distribution functions $F_1(x), F_2(x), \ldots, F_N(x)$ respectively. A linear rank statistic S is defined by

$$S = \sum_{i=1}^{N} c_i a(R_i), \tag{1}$$

where c_i's are known regression constants, R_i is the rank of X_i among $X_1, X_2, \ldots, X_N$ and $a(R_i)$ is the value of the "score function" $a(\cdot)$ at R_i.

The Two-Sample Case

Let $x_1, x_2, \ldots, x_m$ be m independent observations on a random variable X with CDF $F(x)$ and $x_{m+1}, x_{m+2}, \ldots, x_{m+n}$ be n independent observations on a random variable Y with CDF $G(y)$. Consider testing the hypothesis

$$H_0 : p(x_1, x_2, \ldots, x_N) = \Pi_{i=1}^{N} f(x_i)$$

against an alternative

$$H_1 : q(x_1, x_2, \ldots, x_N) = \Pi_{i=1}^{m} f(x_i - d_i) \times \Pi_{j=1}^{n} f(x_{m+j} - d_{m+j}) \tag{2}$$

where $d = (d_1, d_2, \ldots, d_N)$ is an arbitrary vector and $N = m + n$.

Hâjek and Sidak (1967) defined the following statistic S_c for testing the hypothesis (2):

$$S_c = \sum_{i=1}^{N} (c_i - \bar{c}) a(R_i), \tag{3}$$

where $\bar{c} = \frac{1}{N} \sum_{i=1}^{N} c_i$.

Wilcoxon Test

If we put $a(i) = i$ and $c_i = 1$ for $1 \leq i \leq m$ and $c_i = 0$ for $m < i \leq N$ in linear rank statistic (1), we obtain the two-sample Wilcoxon statistic

$$S = \sum_{i=1}^{m} R_i. \tag{4}$$

Van Der Waerden (or Normal Score) Test

The Van der Waerden statistic (or Normal Score statistic) is determined by

$$S = \sum_{i=1}^{m} a(R_i), \tag{5}$$

where

$$a(R_i) = E(S_{(i)})$$

and $S_{(1)} < \ldots < S_{(m+n)}$ are the $m + n$ order statistic from the $N(0, 1)$ distribution.

An approximation (Hettmansperger, 1984, page 145) of expected value is given by

$$E(S_{(i)}) = \Phi^{-1}(\frac{i}{n+m+1}), \quad (6)$$

where

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} \exp(-x^2/2)dx.$$

Thus, we may express the Normal Score statistic as

$$S = \sum_{i=1}^{m} \Phi^{-1}(\frac{R_i}{N+1}). \quad (7)$$

Savage Test

The Savage statistic is defined by

$$S = \sum_{i=1}^{N} a(R_i), \quad (8)$$

where

$$a(i) = \sum_{j=N-i+1}^{N} (\frac{1}{j}).$$

It can be shown that

$$\sum_{j=N-i+1}^{N} \frac{1}{j} = E[-\ln(1-U^{(i)})], \quad (9)$$

where $U^{(i)}$ is the i^{th} order statistic in a sample of size N from the uniform distribution over (0, 1).

Maximum Efficiency Robust Test

Gastwirth and Mahmoud (1986) proposed a Maximum Efficiency Robust Test (MERT) based on a rank test as follows:

$$R(u) = \frac{[-\ln(1-u) - 1 + \Phi^{-1}(u)]}{(2(1+\rho))^{\frac{1}{2}}}, \quad (10)$$

where

$$\rho = \int_0^1 (-\ln(1-u) - 1)\Phi^{-1}(u)du.$$

Define Z_1 and Z_∞ by

$$Z_1 = -\ln(1-u) - 1 \quad (\textit{Savage component})$$

and

$$Z_\infty = \Phi^{-1}(u) \quad (\textit{Normal Score component}).$$

Thus, the standard normal form of MERT (10) is given by

$$R(u) = \frac{Z_1 + Z_\infty}{(2(1+\rho))^{\frac{1}{2}}}. \quad (11)$$

The value of ρ was obtained by numerical integration, and is approximately equal to 0.90228. Hence, the MERT test is given by

$$R(u) = \frac{Z_1 + Z_\infty}{1.9505}. \quad (12)$$

3 Application

The average monthly and yearly temperatures recorded at Toronto Pearson International Airport for the period 1975-1991 are obtained from Environment Canada annual report. Based on the article published in the *Globe and Mail* (February 15, 1992), we have decided to divide the data set into two periods: January 1975-December 1981 and January 1982-December 1991. The period 1982-1991 was identified as the warming trend. The test statistic values for testing the location shifts of monthly, yearly, average temperatures and yearly snowfalls are presented in Table 5.1.

Table 5.1

Average Temperature	Tests			
	MERT	Wilcoxon	Normal Score	Savage
January	2.2902*	2.1170*	2.1378*	2.3293*
February	1.4236	1.2509	1.3377	1.4391
March	0.7023	0.7698	0.5482	0.8216
April	2.1203*	1.5877	1.8524*	2.2833*
May	0.2574	0.2405	0.2172	0.7192
June	1.6213	1.4433	1.4678	1.6946*
July	1.5958	1.0584	1.1961	1.9116*
August	1.8661*	1.6358	1.7115*	1.9285*
September	2.5167*	2.2131*	2.1643*	2.7446*
October	2.2330*	1.8282*	1.9830*	2.3726*
November	0.2385	0.5773	0.4097	0.5732
December	0.7524	0.7216	0.5621	0.9056
Yearly	3.0782*	2.8868*	2.9165*	3.0876*
Yearly Snow	1.2517	1.6350	1.3156	1.0856

* indicates values are significant at 5 percent.

4 Concluding Remarks

In testing for the location shifts, all the test statistics discussed in section 2 are highly significant for the average yearly temperature data. This implies that the yearly temperature averages for the period 1982-1991 are stochastically larger than the period 1975-1981. When the changes are examined, corresponding to monthly averages for the temperature data, we find that all statistics are significant for the months of January, April, August, September, and October except Wilcoxon test for the months of April and August. These significant values seem to point that there is an upward shift in location for the period 1982-1991, identified as the warming trend.

Further, the standardized statistic values for the location shifts of yearly snowfall for the period under consideration do not indicate any trend. However, the Wilcoxon test is significant at 6 percent. Hâjek (1969) remarked that if the tails are diffused, the Wilcoxon test is more powerful than the normal score test. The choice of test statistic does not necessarily imply that data should follow some particular distribution. However, if the true distributions are closer than the loss of efficiency will be small.

Acknowledgement

The authors are grateful to Professor I. B. MacNeill for suggesting this area of research. The research of the first author is supported by NSERC grant OGP131722.

References

1. Ali, M. M. and Giaccottio (1982). The identical distribution hypothesis for stock market prices, location and scale-shift alternatives. *JASA*, V. 77, 19-28.

2. Gastwirth, J. L. and Mahmoud, H. (1986). An efficiency robust nonparametric test for scale change for data from a Gamma distribution. *Technometrics*, Vol. 28, 1, 81-84.

4. Hâjek, J. and Sidak, Z. (1967). Theory of rank-tests. Academic Press.

5. Hâjek, J. (1969). A course in nonparametric statistics. Holden-Day.

6. Hettmansperger, T. P. (1984). Statistical inference based on ranks. Wiley.

7. Jain, S. (1986). Markov chain model and its application. *Computers and Biomedical Research*, Vol. 19, 374-378.

8. Meadows, D. H.; Meadows, D. L.; Randers, Jorgen and Behrens III, W. W. (1974). The limits of Growth. Universe Books, New York.

9. Page, E. S. (1954). Continuous Inspection Schemes. *Biometrika*, 41, 100-114.

10. Savage, I. R. (1956). Contributions to the theory of rank order statistics – the two sample case. *Annals of Math. Stat*, Vol. 27, 590-615.

11. Whiteside, B. S.; Duran, B. S., and Boullion, T. L. (1975). A comparison of some nonparametric tests for scale. *J. Statist. Comput. Simul.*, Vol. 4, 121-132.

12. Woinsky, M. N. (1972a). Nonparametric detection using spectral data. IEEE Trans. on Infor. Theory. 18, 120-126.

13. Woinksy, M. N. (1971b). A composite nonparametric test for a scale slippage alternative. *Annals of Math. Stat.*, 43, 65-73.

14. *The Globe and Mail* (February 15, 1992). Chill is gone from Toronto winters: Decade-long warming trend resulting in less snow, rinks closing.

Modeling the fluctuations of Caspian sea level

M.Reza Meshkani
Department of Statistics,Shahid Beheshti University
Evin,Tehran 19832 IRAN

Abstract

Caspian sea is a closed body of water shared by Iran and USSR. This sea is of great importance to both countries;its oil and world famous caviar are among its precious resources.Fluctuations of Caspian sea level has always concerned both nations. History has shown there have been great raises and falls since its creation about 7 millenia ago,general trend being downward. Presently its average level is about 27 meters below open sea levels.During past decade the level has raised up,flooding the fertile shore land accomodating farms,orchards,and industries.

To ascertain the short term behavior of sea level fluctuations,we have studied it as a stochastic system.Taking the sea level as the output and temperature and rainfall as the inputs, we have modeled a transfer function for this system.Specifically let z(t) denote the monthly average of sea level.Let X(t) and Y(t),respectively,denote the monthly average of temperature in centigrades and monthly total rainfall in milimeter at the gauging station in Anzali,Iran.The follwing seasonal models have been found which describe the system very well:

for $x(t)=X(t)-X(t-12)$ and $B^{j}X(t)=X(t-j)$

$$x(t)=(1-0.56B)(1-.95B^{12})(1-0.20B-0.11B^{2})a(t),$$

where a(t) is a white noise process.For $y(t)=Y(t)-Y(t-12)$ and b(t) another white noise process independent of a(t)

$$y(t)=(1-0.14B)(1-0.91B^{12})b(t)$$

Finally,the transfer function model relating output to inputs is

$$z(t)=(-0.75+0.53B)(1-0.60B-.25B^{2})^{-1}x(t)+0.01y(t)+$$

$$(1-0.77B^{12})(1-.13B)(1-0.97B)^{-1}n(t)$$

Where n(t) is the noise process independent of other noises.

Based on these models,forecasts of future values of the variables and their 95% probability limits have been provided.

Multi-step Prediction Model in Time Series Analysis with Mean Generating Function

Hong-Xing Cao Feng-Ying Wei
(Chinese Academy of Meteorological Sciences,
46 Baishiqiaolu, Beijing, 100081, China)

1. INTRODUCTION

A method of time series analysis with mean generating function (MGF) has been suggested and applied to meteorology, hydrology, environmental science, market prediction, etc. (Cao and Wei, 1989; Cao and Wei, 1991). This method not only can be used to study the statistical properties of time series but also is suitable for multiple steps forward prediction, for example, making a climatic projection for the future decade. From this point, the MGF time series analysis is superior to the AR and ARMA models, which are limited to a-few-step forward prediction.

2. ESTABLISHMENT OF MEAN GENERATING FUNCTION

Calculate first-order and second-order differences of a time series x(t), i.e.

$$\Delta x(t)=x(t+1)-x(t) \qquad t=1,2,\ldots,N-1$$

$$\Delta^2 x(t)=\Delta x(t+1)-\Delta x(t) \qquad t=1,2,\ldots,N-2$$

and

$$x^1(t)=\{\Delta x(1),\Delta x(2),\ldots,\Delta x(N-1)\}$$

$$x^2(t)=\{\Delta^2 x(1),\Delta^2 x(2),\ldots,\Delta^2 x(N-2)\} \qquad (1)$$

where N is the sample size.

Constitute MGFs of x(t) with

$$\bar{x}_l(i)=\frac{1}{n_l}\sum_{j=1}^{n_l-1} x(i+jl) \qquad i=1,2,\ldots,l \quad 1<l<M \qquad (2)$$

where n_l is the largest integer required by $n_l < \mathrm{INT}(\frac{N}{l})$, $M=\mathrm{INT}(\frac{N}{2})$, INT represents taking the integer part of $\frac{N}{l}$ or $\frac{N}{2}$. Do the same for $x^1(t)$ and $x^2(t)$ with (2).

Make periodic extensions for all MGFs, three MGF's series $F^{(0)}(t)$, $F^{(1)}(t)$ and $F^{(2)}(t)$ coming from x(t), $x^1(t)$ and $x^2(t)$ whose element is denoted as f(t), t=1,2,...,N, namely

$$f_l(t)=\bar{x}_l(t-l\mathrm{INT}(\frac{t-1}{l})) \quad t=1,2,\ldots,N; \; l=1,2,\ldots,M$$

Meanwhile, make integrating extensions for MGFs of the first-order difference, namely

$$f_l^{(3)}(t)=\sum_{i=1}^{t} \bar{x}_l^{\,1}(i) \qquad t=1,2,\ldots,N; \; l=1,2,\ldots,M$$

thus we obtain the fourth set of MGF, denoted as $F^{(3)}(t)$.

3.SCREEN PROCEDURE BY COUPLE SCORE CRITERION

Denote S_1 as a fine score which verifies the quantitative error of a forecast and S_2 as a rough score which verifies the qualitative error of a forecast, for example, a trend category. A couple score criterion is defined as (Cao et al.)

$$CSC= S_1 + S_2 \tag{3}$$

A residual square sum is

$$Q_k = \frac{1}{n}\sum_{t=1}^{n} (y_t-\hat{y}_t)^2 \tag{4}$$

where k is the number of independent parameters in a statistical model, y is a dependent variable and $\hat{y}_t$ is its estimation. Obviously, Q_k is equivalent to root mean square error (RMSE)

$$\epsilon=\left[\frac{1}{n}\sum_{t=1}^{n} (y_t-\hat{y}_t)^2\right]^{1/2} \tag{5}$$

total square sum (TSS) is

$$Q_y=\frac{1}{n}\sum_{t=1}^{n} (y_t-\bar{y})^2 \tag{6}$$

where $\bar{y}$ is the mean of dependent variable y. From a view of forecast score, TSS means one makes forecast every time only by taking its mean $\bar{y}$, so we say this is a forecast made by a fool, TSS is the score of the fool forecast.

Let

$$S_1=\frac{Q_k}{Q_Y} \tag{7}$$

or equivalently

$$S_1=1-\frac{Q_k}{Q_y} \tag{8}$$

S_1 implies a good model should be better than a fool forecast, i.e. $Q_k<Q_y$.

As to trend measure, we take minimum discrimination information statistics 2I (Kullback)

$$S_2=2I=2\left[\sum_{i=1}^{G}\sum_{j=1}^{G} n_{ij}\ln n_{ij} +n\ln n - \left(\sum_{i=1}^{G} n_{i.}\ln n_{i.} +\sum_{j=1}^{G} n_{.j}\ln n_{.j}\right)\right] \tag{9}$$

where G is the number of trend categoris in both observations and forecasts, n_{ij} is the number of occurrence in the i category observed and the j category forecasted in G*G contingency table,

$n_{i.}=\sum_{j=1}^{G} n_{ij}$, $n_{.j}=\sum_{i=1}^{G} n_{ij}$.

2I is asymptotically distributed as χ^2 with $\nu_2=(G-1)(G-1)$ degree of freedom (Kullback).

Incorporate (8) and (9)

$$CSC=\lambda(1-\frac{Q_k}{Q_y})+2I \tag{10}$$

where λ is a weight factor which make a order-of-magnitude of S_1 be corresponding to one of S_2. λ is determined by forecast requirement.

We take

$$S_1=(n-k)(1-\frac{Q_K}{Q_y}) \tag{11}$$

in the linear model. If k is fixed, when $n\to\infty$, $(n-k)\to\infty$ and $(n-k)(1-\frac{Q_k}{Q_y})\approx nR^2$ is asymptotically distributed as χ^2 with $\nu_1=k$ degree of freedom (Kendall). From the theorem of χ^2 addition

$$\chi^2_{\nu}=\chi^2_{\nu_1}+\chi^2_{\nu_2}$$

$\nu=\nu_1+\nu_2$, the freedom degree of CSC is $\nu=k+(G-1)^2$ and CSC is asmptotically distributed as χ^2.

Some studies on comparison of CSC with AIC, BIC have been made (Cao et al, Wei and Cao), it shows CSC is an available criterion for determining dimension of statistical model.

As a primary screen, calculate the CSC for every MGF, setting up a univariate regression between the series x(t) and one of MGF extension series. For the MGF if

$$CSC<\chi^2_{\nu,\alpha}$$

where α is a confindence level, e.g. $\alpha=0.05$, ν is a freedom degree of statistic χ^2, the MGF is deleted, In this way, K MGFs are selected and determined. Then we make a fine screen for the K MGFs. Input the MGF one by one according to the order from big CSC to small one, when CSC reachs maximum, for example, it is K_0 MGFs, $K_0<K$, the MGFs which have already input into the regression are just what we want. Thus, the regression equation is obtained, i.e.

$$\hat{x}(t)=\Phi_0+\sum_{i=1}^{K_0}\Phi_i f_i(t) \tag{12}$$

Having made the q-step periodic and integrating extensions for MGF f(t), the q-step forecast with

$$\hat{x}(N+q)=\phi_0+\sum_{i=1}^{K_0}\phi_i f_i(N+q) \qquad q=1,2,... \qquad (13)$$

can be made, where ϕ_i, $i=0,1,...,K_0$, is the regressive coefficient.

4.TEMPERATURE AND PRECIPITATION FORECAST

We employ the above model to the global temperature series during 1951 to 1988 (Folland,1990). Through screen procedure, the regression equation calculated from both the periodic and integrating extensions of MGFs is acquired.

$$\hat{x}(t)=-0.1399-0.7978f_{10}^{(3)}(t)+0.5957f_{12}^{(3)}(t)+0.7450f_{7}^{(3)}(t)-$$

$$0.2832f_{6}^{(3)}(t)+1.6303f_{10}^{(0)}(t)+0.4652f_{8}^{(3)}(t)-0.4777f_{4}^{(3)}(t)$$

$$+0.3083f_{11}^{(3)}(t)-0.1060f_{5}^{(3)}(t)$$

The RMSE of $\hat{x}(t)$ is 0.07 C. For comparsion reason, the regression calculated from periodic-extension MGFs only, denoted as $\hat{x}'(t)$, is also obtained, the RMSE of $\hat{x}'(t)$ is 0.09C. The precipitation over the middle and lower reaches of Yangtse River is also modeled. The model fitting of the precipitation during 1951 to 1988 is reasonable compared with observations and the RMSE for the fitting is 89.13mm. we made the independent forecasts of three years (1989 to 1991), two years of them are in good agreement with observations.

5.REFERENCES

Cao H-X. and F-Y.Wei, 1989: Time series modelling based on mean generating function and its application to long-term climate prediction, preprints of 4 International Meeting on Statistical Climatology, 27-31 March, 1989, Rotorua, New Zealand, 227-231.

Cao H-X. and F-Y.Wei, 1991: A method of time series analysis based on the mean generating function, Chinese J. Num. Math.& Appl. 13:4, 23-30.

Kendall, S.M. and A.Stuart, 1979: The Advanced Theory of Statistics, Vol.2, Charles Griffin & Company Ltd., London,358-366.

Kullback,S., 1968: Information Theory and Statistics, Dover Publications, Inc., New York, 113-119.

Wei F-Y. and H-X. Cao, 1990: Mathematical Model of Long-term Prediction and Its Application, China Meteorological Press, 29-36 (in Chinese).

Folland, C.K. et al., 1990: Observed Climate Variation and Change, Climate Change, The IPCC Scientific Assessment, Houghton J.T. et.al (editors), Cambridge University Press, 195-238.

THE CASE OF LARGEST SAMPLE VALUES TO ESTIMATE A DISTRIBUTION'S RIGHT TAIL

by

J. Tiago de Oliveira
Academy of Sciences of Lisbon, Portugal

Abstract

The "classical" yearly method uses only the sample formed by each maximum observation, per (natural) year to estimate the right tail of a distribution. Here, after a brief description of some approaches to estimate the distribution's right tail using largest observations, we will concentrate in one of them, based on the use of a fixed number of successive largest observations and describe its chief features.

1. Introdution; the three approaches

Until now there are only three approaches to the question of estimating the right tail of a distribution using only the largest values of an i.i.d. sample with a distribution function F(x) attracted for maxima to the one of the asymptotic distributions of maxima (Weibull, Gumbel or Fréchet).

The first one, due to Pickands in 1975, opened the way by showing that the asymptotic distribution of excesses over thresholds (thresholds that must converge deterministically or randomly — like order statistics — to the upper bound or right-end point of the distribution) is a generalized Pareto distribution if and only the maxima are attracted to one of the asymptotic distributions of maxima, the shape coefficient being the same. This method is, in one way, a generalization of the P.O.T. (peaks over threshold) method, largely used by engineers. In practical applications, data are *supposed* to follow a generalized Pareto distribution, although some correction for bias is made for the fact that the i.i.d. hypothesis plus the assumed generalized Pareto distribution are only approximations. The indices of order statistics used increase to infinity with the sample size but the relative rate of increase converges to zero in some way. As, in the applications the method splits in accordance with the form of the asymptotic distribution, a

procedure for statistical choice was developed by Tiago de Oliveira (1991).

In the second approach, due to Weissman in 1978, although the sample size is supposed to increase to infinity, the number of order statistics used is fixed and in practice not only the sample is assumed to be i.i.d. but also the asymptotic distribution is assumed to be the *exact* distribution of the fixed finite number of successive large order statistics. We will follow this approach owing to the observed data available, using some unpublished complementary results by Tiago de Oliveira (1992').

Both approaches assume that the asymptotic results is valid without discussing the *real* applications conditions.

The third approach assumes *only* the existence of an asymptotic distribution for maxima and the (dependent) sample is "cleaned" from the large values clustering in such way that the "cleaned" sample can be considered, approximately, as an independent one. In fact, the "independentization" procedure reduces strongly (to about 10% to 20%) the sample size of the correlated (periodically made) observations used. This method, due to de Haan and Dijk in 1991, seems the best one to be used, although it imposes a large amount of computation.

The details about the procedures can be found in Tiago de Oliveira (1992).

2. Gumbel asymptotics; the one sample case

Consider a sample of n i.i.d. random variables with distribution function F(x) and let $x''_1 \geq x''_2 \geq \ldots \geq x''_n$ be the descending order statistics. Let us suppose that F(x) is attracted, for maxima, to the Gumbel distribution $\Lambda(x)$, i.e., that exist coefficients $a_n > 0$ and b_n, not uniquely defined, such that $\text{Prob}\{(x''_1 - b_n)/a_n \leq x\} = F^n(b_n + a_n x) \to \Lambda(x)$ as $x \to \infty$. Then Weissman (1978) shows that denoting by $z_i = (x''_i - b_n)/a_n$, for i = 1, 2, ... the corresponding reduced (descending) order statistics, the limiting upper ones have the joint asymptotic density

$$\exp(e^{-z_k} - \sum_1^k z_i)(-\infty \geq z_1 \geq z_2 \geq \ldots \geq z_n > \infty)$$

each one having the density

$$\frac{1}{(i-1)!} \exp(-e^{-z} - i\,z) \quad (-\infty < z < +\infty).$$

Assuming that the asymptotic distribution of the x''_i are the exact ones and denoting by $\bar{x}''_k$ the average $\sum_1^k x''_i/k$, the maximum likelihood estimators of a (= a_n) and b (= b_n) are

$$\hat{a} = \bar{x}''_k - x''_k \quad \text{and} \quad \hat{b} = x''_k + \log k \,.\, \hat{a},$$

whose mean values are $E(\hat{a}) = \frac{k-1}{k}\, a$ and $E(\hat{b}) = b + (\gamma - \frac{k-1}{k} \log k - S_k)\, a$, the variances being $V(\hat{a}) = \frac{(k-1)a^2}{k^2}$ and $V(\hat{b}) = (\sigma_k^2 + \frac{k-1}{k^2} / (\log k)^2)\, a^2$ and the minimum variance unbiased estimators are

$$a^* = \frac{k}{k-1}\,\hat{a} \quad \text{and} \quad b^* = x''_k + (S_k - \gamma)\, a^*,$$

the variances of a^* and b^* being respectively $V(a^*) = a^2/(k-1)$ and $V(b^*) = (\sigma_k^2 + (S_k - \gamma)^2/(k-1))\, a^2$. x''_k and a^* are independent and $2(k-1)a^*/a = 2k\,\hat{a}/a$ is a **chi-square** with $2(k-1)$ degrees of freedom or $(k-1)\, a^*/a = k\,\hat{a}/a$ a standard gamma distribution with shape parameter $k-1$, as it will be explained below. Here, S_k denot $\sum_1^{k-1} i^{-1}$ the mean value of z''_i is $\gamma - S_i$ and the variance of z''_i is $\sum_k^{+\infty} i^{-2} = \frac{\pi^2}{6} - \sum_1^{k-1} i^{-2}$. As an immediate consequence we get, for the covariance $C(\hat{a}, \hat{b}) = C(\hat{a}, x''_k + \log k\,.\,\hat{a}) = \log k\,.\,(k-1)\, a^2/k^2$ and, thus, the correlation coefficient is

$$\rho(\hat{a}, \hat{b}) = \log k \,/\, \sqrt{(\log k)^2 + \frac{k^2 \sigma_k^2}{(k-1)}}\,.$$

We will always deal with the maximum likelihood estimators having in mind the property of invariance of the estimators which is essential because we have the transformations in the section 3. The reasoning could be made, with minor changes, for the minimum variance unbiased estimators.

The maximum likelihood estimator of a quantile of probability p is given by.

$$\hat{\chi}_p = \hat{b} + \chi_p\, \hat{a} = \hat{b} - \log(-\log p)\, \hat{a} =$$

$$x''_k + (\log k - \log(-\log p))\, \hat{a}\,.$$

We have, evidently,

$$E(\hat{\chi}_p) = b + (\gamma + \frac{k-1}{k} \log k - \log(-\log p))/a$$

and

$$V(\hat{\chi}_p) = \frac{\pi^2}{6} - \sum_1^{k-1} i^{-2} + (\log k + \log(-\log p)^2)\, a^2.$$

The bias is $B(k)a$, where the coefficient $B(k) = \gamma + \frac{k-1}{k} \log k - S_k$. If we are using return periods T, as $T = 1/(1-p)$, we get the estimator for the quantile

$$\hat{\chi}(T) = \hat{b} - \log(-\log(1 - 1/T)\, \hat{a}\,.$$

Thus, to obtain one-sided confidence regions, we have to compute Prob$\{\hat{b} - \log(-\log p)\, \hat{a} \le x\}$, or as $\hat{b} = x''_k + \log k\,.\,\hat{a}$, we have, more generally, to compute Prob$\{x''_k + t\,.\,\hat{a} \le x\}$, where t is $\log k - \log(-\log p)$ in the specific case we are considering. It imposes a good deal of computation although we are dealing with the convolution of two independent distributions: the (asymptotic)

distribution of z''_k (or x''_k) and the gamma distribution for $\hat{a}$, conveniently reduced.

Until now we have been dealing with the estimation of the parameters and of the quantiles (i.e., forecasting) *assuming* that the asymptotic model is applicable to the observed data. Let us give a test for the validity of the assumed model, at least as an approximate one. It has been shown that the $k-1$ random variables $E_j = x''_j - x''_{j+1}$ are independent and exponentially distributed with scale parameter a and, thus, a test for exponentiality, for instance the estimated Kolmogoroff-Smirnov one, can be used to validate the assumption that the asymptotic model is applicable. The statement made above that $(k-1)\,a/a = k\,\hat{a}/a$ has the standard gamma distribution with shape parameter $(k-1)$, is a consequence of the exponentiality of the E_j because $\hat{a} = \sum_1^{k-1} E_j$.

As the clustering problem may appear. Bartels number can be used to steer the splitting of the sample in subsamples for which the model may be applicable. This will not be needed in our case.

3. Gumbel asymptotics; the many samples case

Suppose that we have now N samples from which we recorded, or want to use only, the k largest values. The first step to the gobal analysis is to test the observed data for homogeneity, i.e., that they belong to the same asymptotics to Gumbel model for largest values. We can not use the likelihood ratio statistics because the number of conditioning equations for the equality of the parameters would be $2(N-1)$, increasing with N! Naturally we are going to build one distance test (or W-test), like the ones described in Silvey (1975). As the logarithm of the denominator of the likelihood ratio statistics has the expression

$$\log \hat{L} = -k \sum_1^N \log \hat{a}_j - N\,k + N\,k \log k$$

where the $\hat{a}_j$ denote a the maximum lilkelihood estimators of a_j obtained in the previous section, it seems natural to use a statistic analogous to the likelihood ratio one but obtained putting for the log-numerator $\bar{a} = \sum_1^N \hat{a}_j / N$ instead of the $\hat{a}_j$, i.e., use the statistic $2\,k\,N(\log \bar{a} - \sum_1^N \log \hat{a}_j / N)$ or yet the statistic obtained developing in series around 1 the values of $\log(\hat{a}_j / \bar{a})$ which is $k/2 \cdot \sum_1^N (\frac{\hat{a}_j}{\bar{a}} - 1)^2$, equal to $kNCV^2/2$ where CV denotes the well known coefficient of variation, or the equivalent and simpler statistic $t_n = \sum_1^N \hat{a}_j / N\,\bar{a}^2$.

By the Law of Large Numbers we know that this statistic converges stochastically to $\frac{E(\hat{a}^2)}{E(\hat{a})^2} = \frac{k}{k-1}$ and, also that $\sqrt{N}(t_n - \frac{k}{k-1})$ is asymptotically normal with zero mean value and variance $2k/(k-1)^3$, by the δ-method, see Tiago de Oliveira (1982). Once homogeneity is accepted, although we could use the maximum likelihood estimators of a and b, it is simpler and more convenient, having the maximum likelihood estimators per sample, to use their averages as estimators, like it was done above, because they are asymptotically normal. They are not only good linear estimators but by their variability we can also evaluate the approximation used. With this approximation, as $\bar{a} = \sum_1^N \hat{a}_j / N$ and $\bar{b} = \sum_1^N \hat{b}_j / N$, we have, as before, for the estimator of the p-quantile is

$$\hat{\chi}_p = \hat{b} - \log(-\log p)\,\bar{a}$$

whose mean value and variance are $E(\hat{\chi}_p)$ and $V(\hat{\chi}_p)/N$, respectively.

4. The case for Fréchet and Weibull asymptotics

As it is well known the Fréchet distribution is $\Phi_\alpha(x) = \Lambda(\alpha \log x)$, for $x > 0$, and so if the sample maxima have an asymptotic Fréchet distribution the distribution of $Y = \log X$ is asymptotic to Λ and the maximum likelihood estimator of a, in the case of one sample, is given by $\hat{\alpha}^{-1} = \bar{Y}'_k - Y''_k$. For large values of α ($\alpha > 5$), Gumbel and Fréchet distributions, with convenient parameters, are practically identical in applications and so if the underlying distribution is unknown, the estimation of α will steer the use of Gumbel distribution or of the Fréchet distribution by taking then logarithms and resorting to the Gumbel one and after reversing the procedure.

If the asymptotic distribution is the Weibull one $\Psi_\alpha(x) = \Lambda(-\alpha \log(-x))$, with $x < 0$, the distribution must have an right-end, or upper bound, $x_0 = \inf\{x : F(x) < 1\}$ and then $Y = -\log(x_0 - X)$ has an asymptotic Gumbel distribution. Notice that to use this transformation, x_0 must be known, at least approximately. In that case α is estimated by the same procedure as for Fréchet distribution. As for large values of $\alpha(> 5)$ Gumbel and Weibull distribution are practically identical for the choice of procedure we can follow the lines above.

5. A case study

In Sneyers and Vandiepenbeeck (1983) we can find the quantiles relative to the maximum wind speeds at Uccle (Belgium) for the years of 1961 to 1989 as well as the 10 maxima wind speeds for the years 1976/80. We are

going to use those data for analysis. The table, with the 10 values of the wind speeds in m/s, ordered, is

Year	x''_{10}	x''_9	x''_8	x''_7	x''_6	x''_5	x''_4	x''_3	x''_2	x''_1	$\tilde{x}''_{10}$
76	20.4	20.7	21.2	22.0	23.4	23.9	24.2	25.7	30.4	32.0	24.4
77	23.9	24.2	25.2	25.5	26.2	27.2	27.7	27.9	26.7	32.0	26.9
78	21.7	22.0	22.0	22.2	22.2	23.0	23.0	23.2	24.4	26.5	23.0
79	22.7	23.4	25.2	25.2	25.2	26.2	27.2	29.2	29.2	30.0	26.4
80	21.3	21.7	22.1	22.4	23.2	23.5	24.0	24.0	24.1	25.3	23.2

As a consequence we obtain the following table of values

Year	$\hat{a}$	$\hat{b}$	KS	%
76	4.0	41.2	.259	.777
77	3.0	38.9	.282	.846
78	1.3	29.5	.333	.999
79	3.7	38.5	.333	.999
80	1.9	29.7	.246	.738

and the average estimators of a and b are $\bar{a} = 2.78$ and $\bar{b} = 35.56$. As all the values of $\sqrt{k-1}$ KS are smaller than 1.36 (the 5% point for the KS test statistic) and thus smaller than 1.63 (the 1% point for the KS statistic) we can accept the hypothesis of exponentiality the weighted spacings E_j, equivalent to the acceptance of the Gumbel asymptotics to the largest sample values (upper order statistics).

The table of the observed quantiles in 1949/60 and of the (lightly smaller) estimated quantiles (e.q.) $\hat{\chi}(T)$ and their estimated standard deviations (e.s.d.) for the return period $T(= 1 - 1/p)$ is the one below

T	2	6	10	25	30	50	100
data(49/60)	29.7	34.2	36.1	39.4	40.0	41.8	44.2
e.q. (76/80)	29.4	33.0	34.5	37.1	37.7	39.1	41.0
e.s.d.	3.21	4.86	5.56	6.72	6.95	7.59	8.46

It shows a good agreement between previous observations and estimations based on later ones. Note that the value of the estimated bias is $B(10) \cdot \hat{a} / = -.179426 \times 2.78 = -.4988$, as data suggests.

References

Sneyers, R. and Vandiepenbeek, M,1983. *On the use of large values for the determination of the distribution of maximum values.* Arch, Met. Geoph. Biocl., B,32,279-286.

Silvey,S.D., 1975. *Statistical Inference.* Chapman and Hall.

Tiago de Oliveira,J. 1982.*The δ-method for obtention of the asymptotic distributions; applications.* Publ. Inst. Statist. Univ. Paris, XXVIII, 49-70.

Tiago de Oliveira, J. 1991. *Extremes and tail estimation; statistical choice.* To be publ. in Memor. Acad. Ciências Lisboa.

Tiago de Oliveira, J. 1992. *Tail estimation and extremes.* To be publ. in Order Statistics and Non-parametrics; Theory and Applications (Homage to Prof. Ahmed E. Sarhan).

Tiago de Oliveira, J. 1992'. *On the estimation of the large quantiles from the largest observation.* In preparation.

Weissman. I.,1978. *Estimation of parameters and large quantiles based on the k largest observations.* J. Amer. Satist. Assoc., 73,364,812-815.

ESTIMATING THE SENSITIVITY OF EXTREME EVENTS TO CLIMATE CHANGE: THE EFFECTS OF AUTOCORRELATION AND CHOICE OF EXTREME VALUE DISTRIBUTION

Barbara G. Brown and Richard W. Katz

National Center for Atmospheric Research
Boulder CO 80307 USA

1. INTRODUCTION

Because extreme climate events have large impacts on society and the environment, it is of interest to anticipate how the frequency of such events might change with a change in climate. For example, it might be desirable to know how the frequency of July daily maximum temperatures exceeding a high-temperature threshold would be expected to change with global warming (Mearns et al., 1984). The statistical theory of extreme values can be used to study the sensitivity of extreme events to changes in the location or scale of the underlying distribution. Analytical results obtained so far on this topic (e.g., Katz and Brown, 1989, 1992) are based on certain large-sample approximations. Specifically, it has been assumed that temperature extremes can be modeled by a Type I extreme value distribution (EVD). The accuracy of this approximation may be affected by the fact that daily temperature observations are strongly correlated over time. In addition, examination of observed temperature extremes suggests that the Type I EVD may not always provide the best model for these data (particularly in the case of minimum temperature). Example curves demonstrating the sensitivity of extreme events to climate change, with the extremes modeled by the Type I, Type II, and Type III extreme value distributions, are illustrated in Section 2. The effects of autocorrelation on relative sensitivity are considered, using computer simulations, in Section 3.

2. SENSITIVITY OF EXTREME EVENTS

2.1 Theoretical framework

A climate variable X is assumed to have a distribution [say with distribution function (d.f.) F], possessing a location parameter μ and scale parameter σ. Climate change is assumed to involve changes in either μ or σ, with the shape of the distribution remaining the same. Now consider a sequence of observations of this climate variable $\{X_t: t = 1, 2, \ldots\}$, assumed to be stationary with common d.f. F. The extreme event E of interest is

$$E = \{M_T > c\}, \tag{1}$$

where $M_T = \max\{X_1, X_2, \ldots, X_T\}$ denotes the maximum of a sequence of length T, and c denotes a threshold.

We are interested in how the probability of the extreme event, $P(E) = \Pr\{M_T > c\}$, changes as the location parameter, μ, changes. The *sensitivity* of the extreme event E to μ is defined as the partial derivative, $\partial P(E)/\partial\mu$. Because extreme events vary in their likelihood, it is more reasonable to work with the *relative sensitivity*

$$\frac{\left[\frac{\partial P(E)}{\partial \mu}\right]}{P(E)}, \tag{2}$$

comparing the sensitivity of an event to its probability.

Under a wide range of dependence conditions (e.g., satisfied by a first-order autoregressive [AR(1)] process), the limiting distribution of the maximum M_t is one of the three so-called extreme value distributions. That is,

$$\Pr\left\{a_T\left(\left[\frac{M_T-\mu}{\sigma}\right]-b_T\right) \le x\right\} \to G(x), \text{ as } T\to\infty, \tag{3}$$

where $a_T > 0$ and b_T are "normalizing constants" (Leadbetter *et al.*, 1983; Chapter 3). The three types of EVD, $G(x)$, are:

$$\text{Type I: } G(x) = \exp(-e^{-x}); \tag{4}$$

$$\text{Type II: } G(x) = \exp(-x^{1/k}), \quad x>0, \quad k<0; \tag{5}$$

Type III: $G(x) = \exp[-(-x)^{1/k}]$, $x < 0$, $k > 0$. (6)

The limiting form of EVD, G, depends on the shape of the right-hand tails of the underlying parent d.f. F . For example, if F is the normal d.f., then the Type I EVD is the limiting form.

If use is made of this large-sample approximation (3), then the relative sensitivity of the extreme event E to μ is approximately given by the hazard function of the d.f. G,

$$H_G(x) = \frac{G'(x)}{[1 - G(x)]} \quad (7)$$

(Katz and Brown, 1992). For G the Type I EVD, this hazard function (7) has the property that $H_G(x) \approx 1 - e^{-x}$, for large x. Thus, the relative sensitivity of E to μ increases to a constant as the threshold c increases (i.e., as the event becomes more extreme). For G the Type II EVD, the hazard rate increases for intermediate values of x, and then decreases for large x. For G the Type III EVD, the hazard rate increases toward infinity as x increases to zero. Thus, for the Type II EVD, relative sensitivity to μ decreases for large values of c. For the Type III EVD, relative sensitivity increases toward infinity as c approaches the location parameter for the EVD.

2.2 Example

The theoretical results regarding relative sensitivity are illustrated in Figure 1 for an application to July daily maximum temperature at Berne, Indiana. These curves were computed using estimated parameter values based on analyses of the long-term data at Berne (Brown and Katz, 1992). The d.f. F is assumed to be the normal distribution, so that μ is the mean and σ is the standard deviation. For the Type II and Type III distributions, the shape parameter k was chosen as $k = -0.20$ and $k = 0.20$, respectively. These values, based on the so-called "penultimate approximation," represent approximately the largest values of k (in absolute value) that might be observed for a series of extremes that actually arise from the Type I EVD (Reiss, 1989; p. 172). Figure 1a shows the relative sensitivity to μ for G the Type I and Type II extreme value distributions; the relative sensitivity to μ for the Type III EVD is shown separately in Figure 1b because of its much larger range of values (i.e., on the y-axis). The three curves have the shapes anticipated according to the analytical results discussed in Section 2.1.

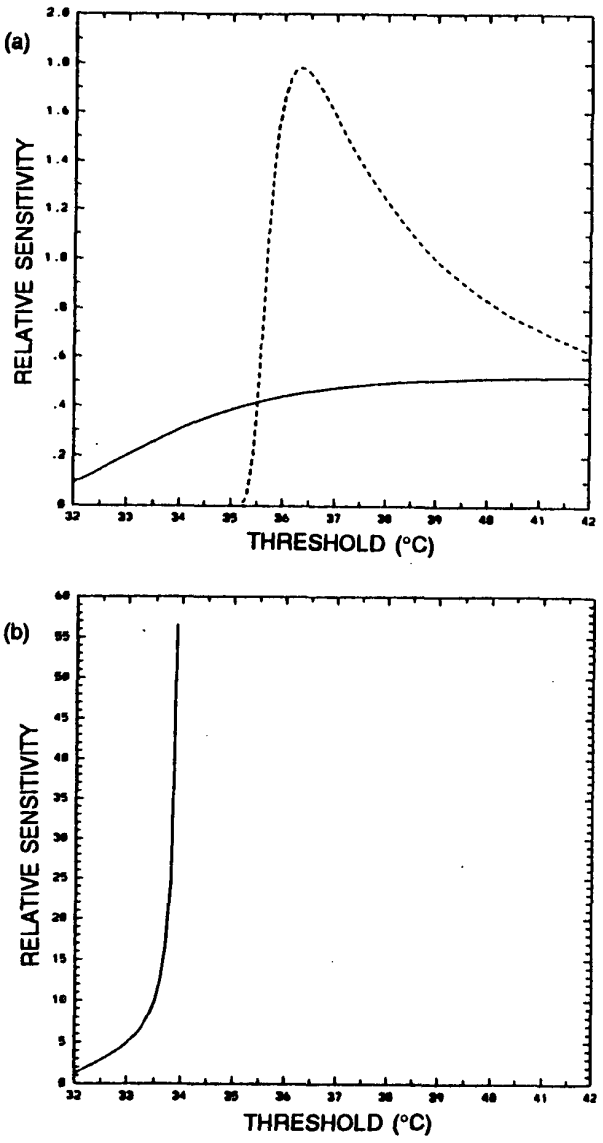


FIGURE 1. Relative sensitivity to μ as a function of c, for (a) G the Type I (solid line) and Type II (dashed line) EVD; and (b) for G the Type III EVD, with distribution parameters appropriate for July at Berne, Indiana.

3. EXTREME VALUE THEORY UNDER DEPENDENCE

Katz and Brown (1992) made use of the Type I extreme value approximation in drawing conclusions about the relative sensitivity of extreme high temperature events, assuming that the underlying observations are independent. Extreme value theory for dependent sequences has been extensively studied. In fact, many

forms of dependence have little effect on the accuracy of approximations for the distribution of the maximum of a sequence. What is evidently not so well understood is the effect of dependence in terms of hazard functions. Here we consider how the actual shape of the hazard function for the exact distribution of the maximum of a finite number T of autocorrelated variables differs from that of the Type I EVD.

3.1 Theoretical framework

We assume here that the d.f. F of the climate variable X is the normal distribution [written $N(\mu,\sigma^2)$]. We further assume that the serial dependence can be modeled as an AR(1) process. That is,

$$X_t - \mu = \phi(X_{t-1} - \mu) + \varepsilon_t,\ t = 2, 3, \ldots . \qquad (8)$$

Here the parameter ϕ, $-1 < \phi < 1$, is the first-order autocorrelation coefficient, and the ε_t are independent $N(0,\sigma_\varepsilon^2)$ with $\sigma_\varepsilon^2 = (1-\phi^2)\sigma^2$. For climate variables, a tendency to persist is typical, with $0 < \phi < 1$. Under these conditions, (3) holds, where G is the Type I EVD (4) (Leadbetter *et al.*, 1983; Chapter 4).

An "associated independent sequence" is defined as having the same marginal d.f. F as the X_T sequence, but being independent [i.e., $\phi = 0$ in (8)]. For this sequence, the exact d.f. of the maximum is

$$\Pr\{M_T \le x\} = [F(x)]^T. \qquad (9)$$

This equation (9) should provide a good approximation for the distribution of the maximum, even when autocorrelation is present, at least provided that F is the normal d.f. (Leadbetter *et al.*, 1983; Chapter 4).

3.2 Simulation results

To study the effects of autocorrelation on relative sensitivity, a large number of random sequences of $N(\mu,\sigma^2)$ variables, with autocorrelation structure as specified in (8), were generated via computer simulations. Each sequence had length $T = 30$ (i.e., roughly corresponding to the number of days in a month). The maximum value in each simulated autocorrelated sequence was determined and the hazard rate, H_A say, was estimated for these values. Two other hazard rates were also computed for purposes of comparison: (i) the hazard rate, H_G say, of the Type I EVD (4); and (ii) the hazard rate, H_I say, for the exact distribution of the maximum under independence (9). Without loss of generality, the parameters of the AR(1) process were selected so that the X_t's have zero mean and unit variance (i.e., F is the standard normal, Φ); that is, we think of the threshold c as being in standardized units.

For convenience, the simulation results are presented in terms of the shape of the *hazard rate*. However, it is important to keep in mind that the hazard rate curve is equivalent to the curve showing the *relative sensitivity* of the extreme event E to the mean μ (2) as the threshold c defining the event is increased (i.e., as the event becomes more extreme). Like the more familiar problems of spectral density function or probability density function estimation, it is necessary to smooth the raw estimates of the hazard rate in order to produce a consistent estimation procedure. Here, with 10,000 replications of the simulation procedure, a moving average of order 3,000 was employed for intermediate values of the standardized threshold; a lower order of moving average (with a minimum order of 1,000) was used to smooth the "tails" of the hazard function.

Figure 2 shows the actual hazard rates H_A for simulations based on two values of non-zero autocorrelation (dotted lines), along with the two theoretical hazard rate curves, H_I and H_G. For $\phi = 0.25$ (Figure 2a), the exact rate for independence, H_I (dashed curve) is a good approximation for the actual hazard rate (in particular, much better than the limiting Type I EVD, consistent with the claim of Leadbetter *et al.*, 1983, Chapter 4). With a high degree of autocorrelation [i.e., $\phi = (0.5)^{1/2}$; Figure 2b], substantial discrepancies arise, with the differences being greatest for relatively small threshold values c. The curve for H_A (dotted) in Figure 2b is straighter in appearance than the curves for either of the two alternatives that, in effect, ignore autocorrelation.

Recalling that the normal distribution itself has an approximately linear hazard rate for large values, the effect of autocorrelation might be somewhat akin to taking the maximum of a smaller number of independent random variables. Nevertheless, the behavior is not at all analogous to the concept of "effective sample size," derived from the effect of autocorrelation on the variance of time averages. The nature of these discrepancies is such that dependence produces at least as steep a curve as the exact hazard rate under independence (i.e., H_I), which is in turn steeper than the theoretical asymptotic rate (i.e., H_G).

4. CONCLUSIONS

Utilizing the Type I EVD, characteristics of the sensitivity of extreme events to changes in parameters of the

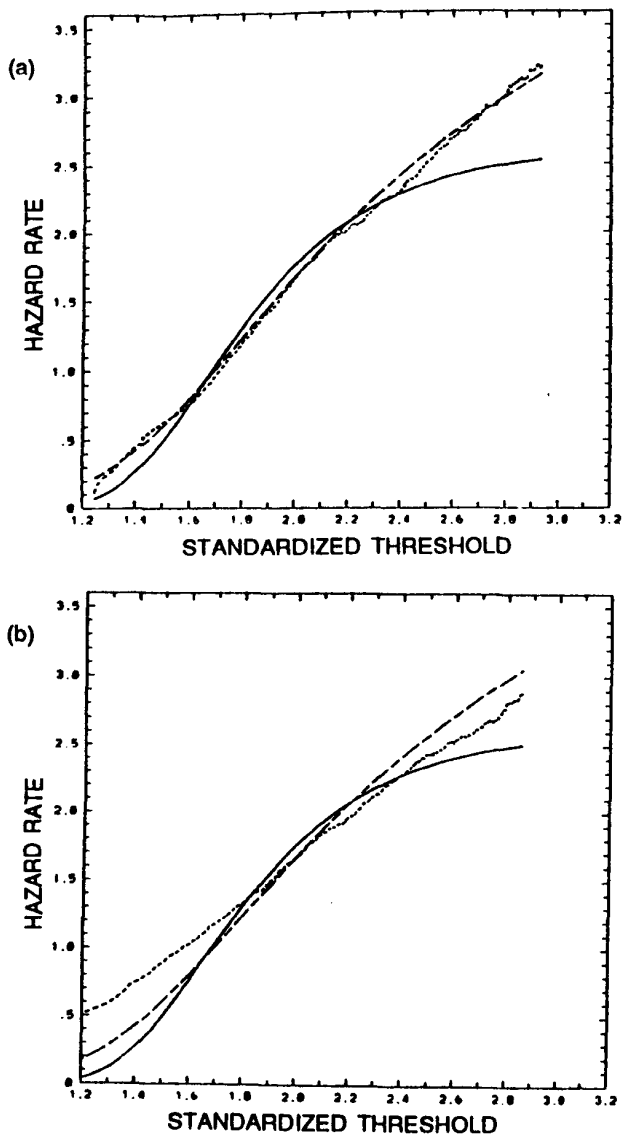


FIGURE 2. Hazard rates for simulated series with $T = 30$ and (a) $\phi = 0.25$ and (b) $\phi = (0.5)^{1/2}$ (dotted lines). For comparison, the hazard rate for the exact distribution of the maximum under independence, H_I, (dashed line) and the hazard rate for the Type I EVD, H_G, (solid line) are also shown.

underlying climate distribution were previously considered in some detail (Katz and Brown, 1989, 1992). Here we have shown that relative sensitivity relationships are quite dependent on the choice of EVD. Previous studies also ignored the day-to-day autocorrelation that is a feature inherent in most climate variables. However, simulations described here indicate that our characterization of the relative sensitivity is quite robust to the possible presence of autocorrelation. Further, what discrepancies exist are in the direction of even more sensitivity than either the asymptotic theory or the exact theory under independence would predict.

ACKNOWLEDGMENTS

This research was supported in part by the Environmental Protection Agency under Cooperative Agreement CR-815732-01-0. The views expressed are those of the authors and do not necessarily reflect those of the U.S. government. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

REFERENCES

Brown, B.G., and R.W. Katz, 1992: Regional analysis of temperature extremes, with implications for climate change. In preparation.

Katz, R.W., and B.G. Brown, 1989: Climate change for extreme events: An application of the theory of extreme values. *Preprints, Eleventh Conference on Probability and Statistics in the Atmospheric Sciences.* 2-5 October, Monterey, CA. American Meteorological Society, Boston, 10-15.

Katz, R.W., and B.G. Brown, 1992: Extreme events in a changing climate: Variability is more important than averages. *Climatic Change,* in press.

Leadbetter, M.R., G. Lindgren, and H. Rootzen, 1983: *Extremes and Related Properties of Random Sequences and Processes,* Springer-Verlag, New York.

Mearns, L.O., R.W. Katz, and S.H. Schneider, 1984: Extreme high-temperature events: Changes in their probabilities with changes in mean temperature. *Journal of Climate and Applied Meteorology,* **23**, 1601-1613.

Reiss, R.-D., 1989: *Approximate Distributions of Order Statistics, with Applications to Nonparametric Statistics.* Springer-Verlag, New York.

Deriving the Characteristics of Cold Spells from Crossing Theory.

Peter R. Waylen
Department of Geography
University of Southampton
Southampton, England, SO9 5NH[1]

1. Introduction

Knowledge of the statistical properties of excursions of random processes above or below some particular level of interest is important in many aspects of applied climatological studies. Practical examples include critical wind speeds, precipitation levels, heat waves and cold spells. The theoretical behaviour of such excursions has been studied extensively since the work of Rice (1945) and is summarized by Leadbetter et al. (1983). Generally applications involve the use of simple fixed truncation level and a stationary Gaussian process (for example, Nordin and Rosbjerg, 1970), although attempts have been made to remove seasonal non-stationarity by subtracting long-term means from the data series (Rodriguez-Iturbe, 1968) and by applying seasonally varying truncation levels (Griffiths, 1990). Recently the crossing properties of log-normal and Chi-squared processes have been detailed (Desmond and Guy, 1991) in addition to Gaussian processes.

The particular variable of interest in this study in that of daily minimum temperatures, crossing below freezing. The occurrence of sub-freezing temperatures is of considerable practical importance to the citrus industry in Florida, one of the study areas, while the dates of the first and last excursions below freezing (freezes) determine the length of the available growing season at the other study site, Regina Airport, on the Prairies of Canada. First and last freezes have long been studied as random variables (Reed and Tolley, 1916; Thom and Shaw, 1958). Dates of first and last freeze are generally considered to follow a Gaussian distribution (see for example, Schmidlin and Dithier, 1986) and have been considered as indicators of climate change (Skaggs and Baker, 1985). This paper investigates the expected number and timing of excursions of daily minimum temperatures below freezing by means of crossing theory and develops distributions of the dates of first and last freezes of the year as a result.

2. Theoretical considerations

The average number of down-crossings $E[D_q(0,1)]$ per unit time, by a discrete stationary Gaussian process (zero mean, unit variance), U_t, of interval di, of a truncation level q, is

$$E[D_q(0,1)] = P[U_1<q] - P[U_1<q, U_2<q] \quad (1)$$

This value is determined by the bivariate normal distribution and is dependent upon the first order autocorrelation phi. In this study the value E[Dq(0,1)] is interpreted as the daily risk of a sub-freezing period commencing on that day.

Leadbetter et al. (1983) point out that the number of crossings of the threshold in some period greater than the basic interval, becomes Poisson distributed as the threshold gets further from the mean of the process. A time dependent, non-homogeneous Poisson distribution may be used to model the number of such events in a particular time period, t.

$$P[m(t)] = [\exp(-G(t))\, G(t)\, m(t)] / m(t)! \quad (2)$$

where

[1] Currently on research leave from Department of Geography, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1.

$$G(t) = \sum_{i=1}^{365} g(i) \qquad (3)$$

and g(i) is the intensity function of the Poisson process derived from equation 1 as the daily rate of freeze occurrence. The cumulative density function of the date of first freezes, F(t), becomes

$$F(t) = 1 - \exp[-G(t)] \qquad (4)$$

and the cumulative density function of last freezes, H(t) becomes

$$H(t) = \exp[-G(365) - G(t)] \qquad (5)$$

Assuming that the dates of first and last freeze are independent then the discrete probability distribution, Y(l), of the length, l, of the growing season is

$$Y(l) = \sum_{i=1}^{\infty} F(i) \cdot H(i+l) \qquad (6)$$

Equation (1) assumes a stationary Gaussian process, but daily minimum temperatures are non-stationary in, at least, mean and variance. The use of Fourier analysis to represent the seasonality of climatological variables is well established. Other functional forms could be used to model the seasonal trends in the parameters but the Fourier technique was selected on the basis of the ease of parameter estimation, physical interpretability and parsimony in the number of coefficients, and the subsequent goodness of fit. Twelve monthly estimates of first order daily serial correlation are made empirically because of the sensitivity of estimates of serial correlation to small sample sizes.

3. Application

Daily minimum temperatures are analyzed from Regina Airport, Saskatchewan (1885-1982) and Lake City, Florida (1954-1985). Both sets of data are collected by the respective national climatological institutes. Although the record length of the latter is not the longest in the state, it does represent one of the most freeze prone regions and has examples of years in which one or fewer freezes occur, which have caused analytic problems (Vestal, 1971).

a. Daily temperatures

A simple sinusoid is insufficient to model the complexity of the empirical functions of daily minimum temperatures (figure 1). There is a flattening of the variance functions at both sites during summer, and a flattening of the mean function during summer at Lake City, which can be modeled by the addition of a second harmonic.

b. Daily freeze rates

Historic freezes are extracted from the data and expressed as relative frequencies, or daily freeze rates by date of occurrence (figure 1). Lake City experiences a clear winter maximum of freeze occurrence, no freezes ever having been recorded outside of the October-April period. Regina Airport by contrast depicts a distinctly bimodal pattern of occurrence. Mean minimum daily temperatures are so low in winter that, despite the high variance, above-freezing temperatures, and therefore down-crossings, are exceedingly rare. The highest rates of down crossing occur in the spring and fall. Expected freeze rates based on equation (1), the Fourier representation of the mean and variance functions and empirical autocorrelation functions, reproduce the historic qualities of modality, timing and appropriate values of daily freeze rates.

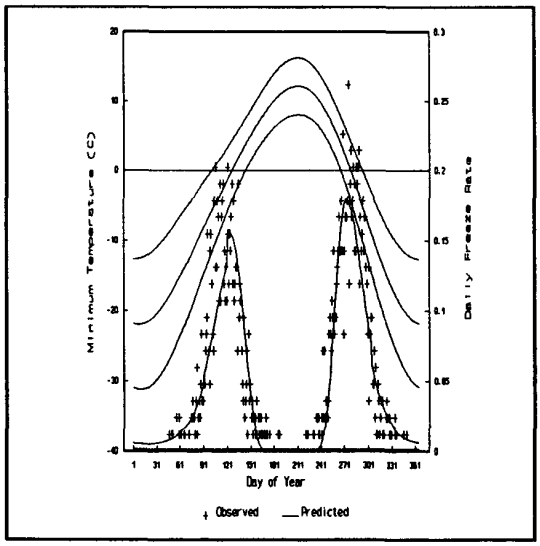


Figure 1. Mean minimum daily temperatures plus and minus one standard deviation and observed and predicted freeze rates at Regina Airport.

Some interesting statistical properties of freezes at each location are also highlighted. At Regina the peak intensities of freeze occurrences vary between

the two modal times and during each season the freeze rates are asymmetrically distributed about the modal value. Both periods show longer tails going into the winter season and very abrupt changes going into and leaving the summer season, resulting from higher variances in winter than in summer. The predicted freeze rates at Lake City display a lesser degree of asymmetry and preserve the leptokurtic behaviour observed in a previous empirical study of freeze rates in north and central Florida (Waylen et al., 1985).

c. Numbers of freezes.

A temperature threshold of 0 C is always at least one standard deviation below the mean on any day of the year at Lake City and therefore the application of the Poisson distribution appears appropriate to model the number of freezes per year. The relative frequencies of historic annual freeze counts are extracted and are not significantly different from a Poisson distribution using a parameter G calculated as the summation of predicted daily freeze rates. However, at Regina, the required threshold temperature corresponds to the annual mean minimum daily temperature (0.006 C), and the applicability of the Poisson distribution seems less likely. Because of the high average number of freezes per year the Regina data are sub-divided into two groups, "Spring" (Jan. 1 to Jul. 10) and "Fall" (remainder of year). This does not affect the statistical characteristics of the annual count because of the additive properties of the Poisson distribution. Both subgroups can be represented by a Poisson distribution. The number of freezes in each season during a year appears to be independent of one another (correlation coefficient $3.5*10^{-6}$, coefficient of linear regression not significantly different from zero at 0.01 level).

d. First and last occurrences

Dates of the first (Fall) and last (Spring) freezes are extracted from the historic records. As records are kept on the basis of a calendar year this may lead to differing numbers of first and last freezes. At Lake City the situation may become more complex as there may be one or fewer freezes per year, and the first freeze may actually occur in January or February.

Traditionally the dates of first and last freeze have been modelled by a Gaussian distribution (Thom and Shaw, 1958) and this practice continues. Historic data (figure 2) illustrate the difficulty of this assumption, particularly in the tails of the distribution. Distributions estimated from crossing theory more faithfully reproduce this unusual tail behaviour. Estimates of higher order moments are notoriously sensitive to small sample sizes (even the 84 years at Regina), however estimates of the mean and standard deviation from the moments of the predicted distribution (table 1) closely match the historic data. At Regina both the first and last freezes are predicted to be negatively skewed, yet observed dates of last freeze are positively skewed. At Lake City the signs of skewness are correctly estimated, but the magnitude is underestimated for the last freeze. The kurtosis at Regina is estimated as highly leptokurtic for both first and last freeze, but the observed data are only so to a lesser degree. Predictions of platykurtic behaviour for last freezes at Lake City are reasonable, but the predicted leptokurtic behaviour of first freezes is in error.

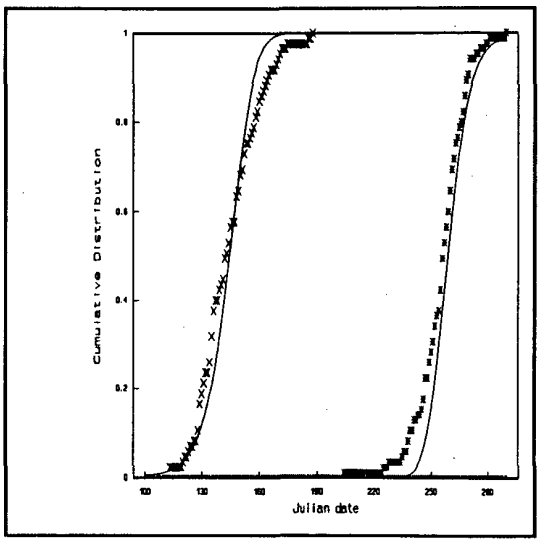


Figure 2. Observed and predicted cumulative distributions of the dates of first and last freezes at Regina Airport.

e. Growing season

The length of the growing season (time between last freeze of spring and first freeze of fall) is important at Regina on the Prairies of Canada, while the length of the freeze-prone period (first freeze to last freeze) is importance in north Florida. An estimate of the distribution of both variables can be made from equation (6) by transposing the distributions of first and last freezes. An assumption of independence is essential to the operation of equation (6) and is supported by a plot of first and

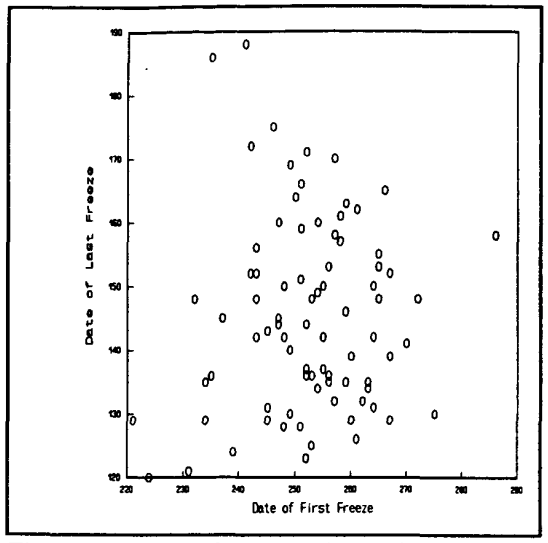


Figure 3. Scatter plot of the dates of first and last freeze at Regina Airport.

last dates in figure 3 (correlation and regression coefficients not significantly different from zero at 0.05 level).

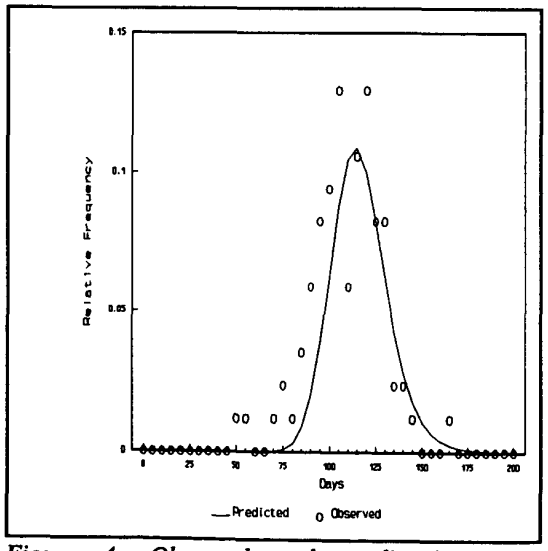


Figure 4. Observed and predicted relative frequencies of the length of the growing season at Regina Airport.

Observed and predicted functions of season lengths are plotted in figure 4. Again the tails appear to diverge from Gaussian behaviour. The predicted descriptive statistics of the season lengths do not match the observed data so well as the first and last freezes because their estimates compound the errors in the latter two variables. The predicted mean length of growing season at Regina is 12 days longer than the observed value and the skewness is slightly positive compared to small negative observed value. Prediction of the mean length at Lake City is within 2 days of the true value while the standard deviation is in error by 8 days (44% of true value).

f. Changing the truncation level

The use of crossing theory permits the extension of the results to truncation levels other than the one initial employed. This may be of particular interest when the truncation level is seldom, if ever experienced in the historic record. At -2.2 C (27 F) new buds on citrus trees will be destroyed, the occurrence of such temperatures is also recorded by the National Climate Data Center (U.S.) as one of their levels of "freeze data".

The preceding analyses are repeated at the new truncation level. The predicted annual numbers of "freezes", dates of first and last freeze and the length of the freeze-prone season cannot be considered significantly different from the observed data. Data are not plotted on normal paper because of the difficulties in handling years in which no "freezes" occurred. These estimates are make for crossings of a temperature which is far less common in north central Florida. Empirical estimates of the distribution of the dates of first and last events would have to be based on sample sizes of 20 and 18 respectively.

5. Discussion and Conclusion

Although lacking parsimony in its parameters (5 parameters for each of the daily mean and variance functions, and 12 monthly observations of daily autocorrelation) the model provides practical information on the characteristics of the frequency and timing of cold spells, which exceeds that available from the standard approach. Inclusion of the behaviour of the mean, variance and autocorrelation functions permits the prediction of the dates of all freezes (not just first and last). It also leads to a great deal of flexibility in the derived form of the distribution of first and last freezes, as opposed to the generally assumption of normality. Based as it is, on the distribution of daily minimum

temperatures and crossing theory, rather than empirical observations of sub-freezing events, it also offers a great deal of flexibility in handling differing truncation levels between sites (as in the comparison of Canadian and Floridian data) and differing practical truncation levels at specific sites. These differing levels may be crop specific or time dependent as crops may vary in their susceptibility to differing temperatures in each season.

The theoretical link between above (below) threshold crossings and annual maxima (minima) is well established (Leadbetter et al., 1983). In this practical exercise the application of the theory becomes difficult because of the time variant nature of the Poisson crossings and non-stationarity of the Gaussian process. In previous studies of similar types of problems, certain simplifying assumptions have been made which are not completely applicable in this case. In flood studies (Rasmusson and Rosbjerg, 1991) the rate of crossings has been permitted to vary seasonally, but the peak above threshold crossing in each event has been considered to follow a single probability distribution, such as the gamma, log-normal and generalized Pareto. Others have divided the year according to some dominance of flood generating processes like hurricanes and fronts (Waylen, 1991). Each process has associated with it a separate time varying rate of occurrence (not necessarily mutually exclusive in time) and a fixed probability distribution of peak magnitudes within an event. The distribution of annual extremes generated by each process is estimated separately then compounded to produce a distribution of overall maxima.

Such parallels cannot be easily extended to the study of temperatures where the means and variances of daily minimum temperatures change continuously throughout the year (figure 1). When the threshold of interest is sufficiently far from the general level of the process under study, then the assumption of a single simple distribution of within-event maxima is reasonable. Waylen (1988) was able to assume exponentially distributed event peaks when dealing with temperature time series in central and southern Florida and thereby able to produce the distribution of annual minimum temperatures at each site and a time varying function for the distribution of minimum temperatures within any times period of interest. However the choice of, relatively, so low a level leads to the loss of potentially pertinent information and adds to the uncertainty of the resultant distribution. Secondly, the threshold level of interest is, as in this case, frequently not selected upon statistical grounds, but upon very meaningful practical requirements.

Acknowledgements. This research was supported in part by a grant from the Division of Sponsored Research, University of Florida and a fellowship from the Committee for Advanced Studies, University of Southampton.

REFERENCES CITED

Desmond, A.F. and B.T. Guy, 1991. Crossing theory for non-Gaussian stochastic processes with application to hydrology. *Water Resour. Res.*, **27**, 2791-2797.

Griffiths, G.A., 1990. Rainfall deficits: Distribution of monthly runs. *J. Hydrol.*, **115**, 219-229.

Leadbetter, M.R., Lindgren, G. and H. Rootzen, 1983. *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag, New York.

Nordin, C.F. and D.M. Rosbjerg, 1970. Applications of crossing theory in hydrology. *Bull. Int. Ass. Sci. Hydrol.*, **15**, 27-43.

Rasmusson, P.F. and D.M. Rosbjerg, 1991. Prediction uncertainty in seasonal partial duration series. *Water Resour. Res.*, **27**, 2875-2883.

Reed, W.G. and H.R. Tolley, 1916. The probable growing season. *Mon. Wea. Rev.*, 44, 509-512.

Rice, S.O., 1945. Mathematical analysis of random noise. *Bell Sys. Tech. J.*, **24**, 24-156.

Rodriguez-Iturbe, I., 1969. Applications of run theory to hydrology. *Water Resour. Res.*, **5**, 1422-1426.

Schmidlin, T.W. and B.E. Dethier, 1986. A statistical analysis of the freeze hazard in New York State. *Phys. Geogr.*, 7, 246-257.

Skaggs, R.H. and D.G. Baker, 1985. Fluctuations of the length of the growing season in Minnesota. *Climate Change*, 7, p.403-14.

Thom, H.C.S. and R.H. Shaw, 1958. Climatological analysis of freeze data for Iowa. *Mon. Wea. Rev.*, 86, p.251-7.

Vestal, C.K., 1971. First and last occurrence of low temperatures during cold season. *Mon. Wea. Rev.*, 99, 650-652.

Waylen, P.R., 1988. A statistical analysis of freezing temperatures in central and southern Florida. *J. Climatol.*, **8**, 607-628.

Waylen, P.R., 1991. Modelling the effects of tropical cyclones on flooding in the Santa Fe river basin, Florida. *Geojournal*, **23**, 361-373.

Waylen, P.R., Chen, E.Y. and J.Gerber, 1985. A method of estimating the probability of cold spells in Florida. *Proceedings of the Florida State Horticultural Society*, 99, p.1-8.

Table 1. Estimates of the descriptive statistics of the predicted and observed (in parentheses) values of freeze variables.

Variable	Mean	Standard Deviation	Skewness	Kurtosis
Regina				
First freeze fall	254.8 (252.0)	16.3 (12.8)	-0.495 (-0.781)	10.986 (5.846)
last freeze spring	139.4 (144.6)	12.1 (15.3)	-1.139 (+0.476)	8.985 (3.054)
Growing season	121.0 (108.2)	15.6 (19.2)	+0.653 (-0.333)	4.037 (3.886)
Lake City				
First freeze fall	335.7 (331.7)	17.9 (15.4)	+0.417 (+0.341)	3.790 (2.587)
Last freeze spring	55.2 (58.3)	12.1 (18.3)	-1.728 (-0.599)	2.565 (2.763)
Freeze-prone season	91.0 (92.8)	26.8 (18.7)	-0.204 (+0.301)	3.035 (1.557)

TEMPERATURE THRESHOLDS FOR CLIMATE SCENARIOS

Peter J. Robinson
Department of Geography
University of North Carolina
Chapel Hill, NC 27599-3220, USA

The current concern with climate change and its impacts requires the development and use of scenarios of future climate containing sufficient detail for impact assessment. The current physically based general circulation models do not provide the required spatial and temporal detail, while historical analogues, which contain sufficient detail, do not encompass the range of likely future conditions. Approaches combining model output and historical data are needed (Lamb, 1987; Robinson and Finkelstein, 1991). One approach, designed to estimate the number of days per month exceeding a specified temperature threshold after a doubling of atmospheric CO_2 has been demonstrated to provide a simple method of scenario development for impact assessment purposes (Robinson, 1991, 1992). This paper briefly reviews that approach and explores the nature of the temperature variability assumptions which support it.

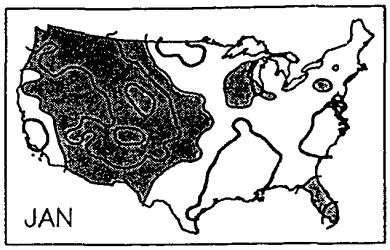
The scenario model

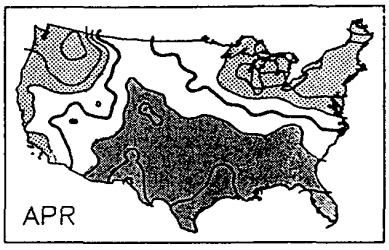
In this scenario development method the historical temperature record for an individual station is used to develop a regression relationship between monthly average temperature and the number of days in the month exceeding a user-selected temperature threshold. The resulting equation is then extrapolated from the observed data to produce an estimate of the number of days for the monthly average temperatures postulated by general circulation models for a doubled CO_2 atmosphere. Results from several models are used, and thus a series of scenarios are produced, giving a range of possible future conditions for consideration by a decision-maker.

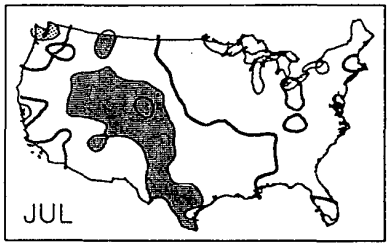
The method was designed to allow impact assessors to develop their own scenarios. Simplicity and ease of interpretation was paramount. The validity of the method depends on the suitability of a linear approach, the strength of the regression relationship, the amount of extrapolation used, and the assumption that daily temperature variability is a linear function of monthly average temperature. The first three of these are mainly statistical in nature, and preliminary tests have provided encouraging results (Robinson, 1991, 1992). The assumption concerning daily temperature variability, however, has both statistical and physical aspects and has not been adequately tested.

Temperature variability as a function of mean temperature

In one of the few explorations of the relationship between the variability of daily temperatures and monthly mean temperature, Mearns et al. (1984) found that the variance of the daily maximum temperature decreased as the monthly mean temperature increased. Their analysis, however, was confined to the Summer months for four stations in the







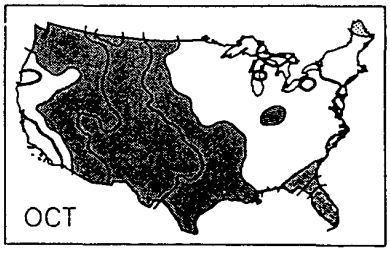


Fig. 1. Slope of regression relationship between monthly mean temperature and the standard deviation of daily average temperature. The zero contour is bold, slopes >0.1 are lightly stippled, those <-0.1 heavily stippled.

United States Corn Belt, but here is expanded using the same dataset as that used for the scenario development. This dataset consists of 126 stations with virtually complete daily temperature records for the 1950-1986 period, in an approximately regular grid across the conterminous 48 states.

When the standard deviation of the daily mean temperature was assessed as a function of monthly mean temperature a negative slope was most common, indicating a decrease in standard deviation with increasing mean temperature (Fig. 1). Slopes greater than approximately ±0.1 (the stippled areas) are significant at the 95% level. For most of the year the west has negative values, expanding into the southeast in the Spring. Areas with positive values are restricted to the northeast in Spring and Autumn and along the west coast in Spring.

A similar analysis was performed for the relationship between monthly mean temperature and the monthly mean maximum temperature. Since in the United States mean temperatures are calculated as

Mean = (Maximum + Minimum)/2,

analyses using the maximum also provide information about the relationships between monthly means and both minimum temperatures and diurnal ranges. For the relationship using maximum temperatures, regression slopes can vary from 0 to 2. A zero slope indicates no change in the mean maximum as the monthly mean increases, all increase being in the minimum, which leads to a decreasing diurnal range. The opposite, an increasing diurnal range, little change in the minimum, and rapid increase in the

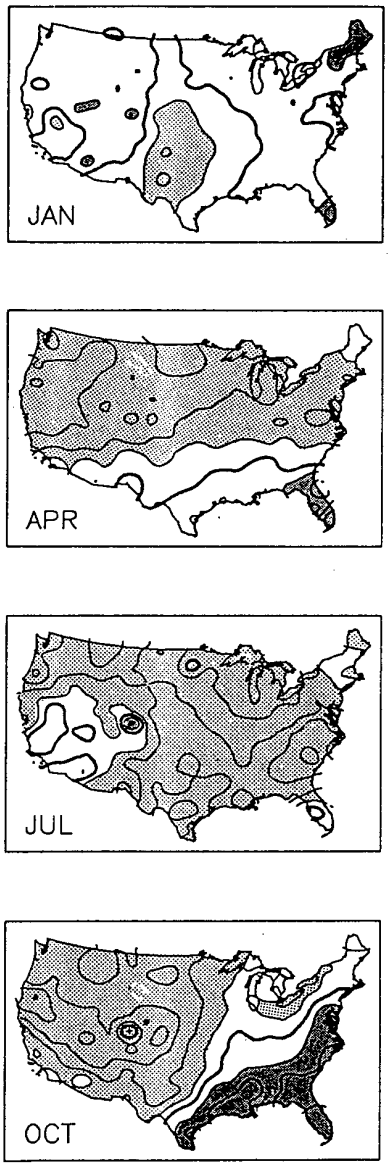


Fig. 2.Slope of regression relationship between monthly mean temperature and mean monthly maximum temperature. The 1.0 contour is bold, slopes >1.1 are lightly stippled, those <-1.1 heavily stippled.

maximum, occurs when the slope approaches 2. At unity the maximum, mean, and minimum increases are identical. A significant positive slope occurs through much of the nation in Spring and Summer, indicating an increase in diurnal range as mean temperature increases (Fig. 2). In January only the south-central area has a significant positive slope, while the south in Autumn is the only region with a significant negative slope.

Secular temperature trends

It is possible that the above trends could be influenced by secular changes in temperature during the period of the observational record. One possible cause of an increasing temperature is the greenhouse effect itself, where minimum temperatures may increase faster than the maximum. If this effect is already occurring, it would influence the regression used to develop the scenarios of temperature exceedence frequency. No attempt was made in this analysis to ascertain whether or not the greenhouse effect was present, only to identify a possible source of bias within the scenario development method. Consequently regressions of both maximum and minimum temperatures against time were undertaken and monthly and seasonal analyses performed. Only the seasonal results for maximum temperatures are shown here (Fig. 3). Results using the minimum were virtually identical, there being no significant trends for almost all of the area. The exceptions were, for the change in maximum, a small area in the north-central region with a significant increase in Spring and portions of the central area with a Winter decrease.

For the north-central United States a comparison is possible with the long-term trends presented by Karl et al. (1991) (Table 1). The present short-period results show more rapid changes. This is expected given the fluctuations in temperature trends within the period of instrumental records. In addition, the comparison area is largely that having the greatest temporal trends (Fig. 2). Thus it emphasizes that area where secular changes would have the most influence on the regression-based scenario results. In Spring the increases in both maximum and minimum temperatures indicate little change in the diurnal range, while the different directions of the Winter change suggest an increase in diurnal range. These differences are not reflected in the relationship between mean temperature

TABLE 1: Temporal temperature trends for the central United States. Values are in °F/100 years.

Period of Record	Karl et al. 1901-1978	Present 1950-1986
Maximum temperature		
Spring	0.68	7.89
Summer	0.81	0.94
Autumn	-0.99	-6.76
Winter	0.36	-6.55
Minimum temperature		
Spring	1.69	8.16
Summer	1.76	0.59
Autumn	1.63	0.27
Winter	1.06	-2.51

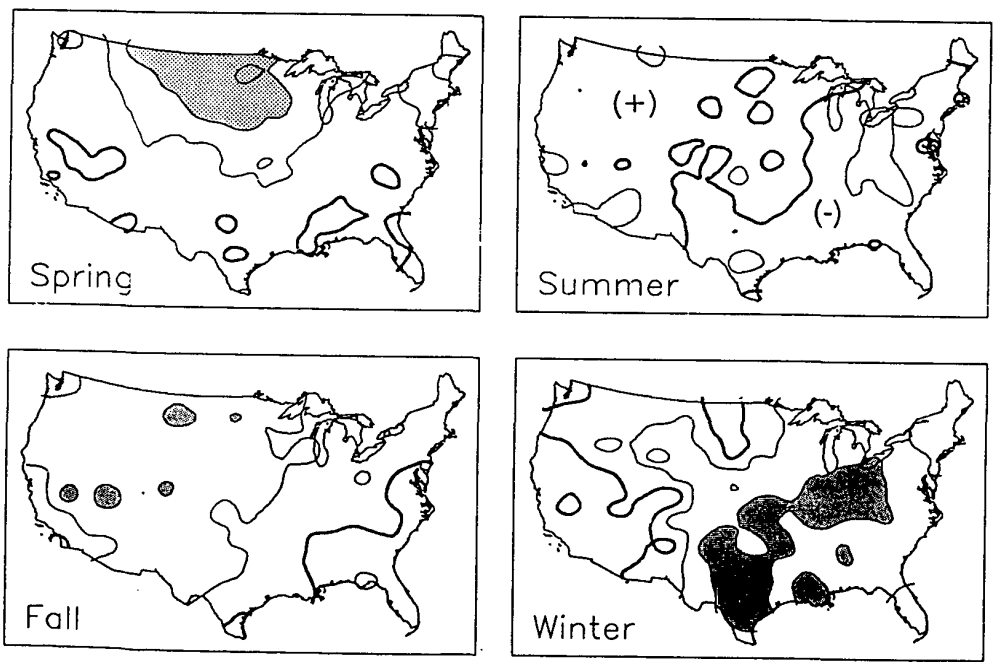


Fig. 3.Trend in seasonal mean maximum temperature as a function of time (°F/yr) as derived from linear regression analyses for the 1950-1986 period. The zero contour is bold, slopes >0.1 are lightly stippled, those <-0.1 heavily stippled.

and mean maximum temperature (Fig. 2), suggesting that the scenario development method is relatively free from bias which might be introduced as a result of any already established greenhouse effect.

Discussion

The standard deviation of mean daily temperature decreased with increasing monthly mean temperature for most of the nation. However, there were some areas and times with the opposite trend, indicating that any scenario of the frequency of threshold exceedences developed under the assumption of a constant daily temperature variability as mean temperatures increases will overestimate the frequency in many areas, but underestimate in others. There is similar temporal and spatial variability of the relationship between the diurnal temperature range and mean temperatures. For most of the United States and in most seasons the diurnal temperature range increases as the monthly mean temperature increases. The only regions where there has been a significant trend in the diurnal range as a function of time over the last 40 years were the north-central portion of the nation in Spring, with an increase, and in the south-central region, which had a Winter decrease.

For scenario development, therefore, there is uncertainty in the assumption that the variability of daily temperatures, whether mean or maximum, responds to changes in monthly mean temperature is a temporally and spatially uniform way. For much of the nation, and for most seasons, the assumption appears reasonable, given the uncertainties in other aspects of the method and in the general circulation model outputs themselves. Nevertheless, in general for the United States the method is likely to underestimate the increase in frequency of temperature exceedences resulting from climatic change.

References

Lamb, P.J., 1987: On the Development of Regional Climatic Scenarios for Policy-Oriented Climatic-Impact Assessment. *Bull. Amer. Meteor. Soc.*, **68**, 1116-1123.

Karl, T.R., R.R.J. Heim, and R.G. Quayle, 1991: The greenhouse effect in central North America: If not now, when?. *Science*, **251**, 1058-1061.

Mearns, L.O., R.W. Katz, and S.H. Schneider, 1984: Extreme high-temperature events: Changes in their probabilities with changes in mean temperature. *J. Climate Appl. Meteor.*, **23**, 1601-1613.

Robinson, P.J., 1991: Scenarios of future climate for impact assessment. *Environ. Prof.*, **22**, 8-15.

Robinson, P.J., 1992: Developing scenarios of temperature thresholds, *Phys. Geog.*, **13**, 1, 14-30.

Robinson, P.J. and P.L. Finkelstein, 1991: The development of impact-oriented climate scenarios, *Bull. Amer. Meteor. Soc.*, **72**, 481-490

TRENDS IN EXTREME SEA LEVELS

J. A. Tawn and M. J. Dixon
Department of Probability and Statistics
University of Sheffield, Sheffield, S3 7RH. U.K.

Summary

Knowledge about current and future extreme sea level trends is vital for the design of coastal flood prevention schemes and the understanding of the possible impact arising from climate change. Since the traditional method of estimating extreme sea level trends uses data from only the site of interest, estimates are often too inaccurate to be of any value to coastal planners and provide only limited knowledge about the spatial variation in trends. In this paper these trends are estimated within a spatial model for extreme sea levels which accounts for inter-site dependence and coherent changes in the underlying generating mechanism. The resulting trend estimates are much improved exhibiting smooth, but significant, spatial variation and are similar in value to mean sea level trends. The methodology is illustrated with U.K. and Australian applications.

1 Introduction

For coastal regions rising sea levels is one of the most serious consequences of possible climate change. The majority of studies of sea level change have focussed on current and future trends in mean sea levels (Barnett, 1984, and Woodworth, 1987). However, for the design of coastal flood prevention schemes coastal engineers require knowledge about current and future trends in extreme sea levels. Trends in extreme sea levels principally can arise due to trends in (*i*) mean sea level; (*ii*) the number of extreme storms; and (*iii*) the variability of extreme storms: thus, trends in mean and extreme sea levels may be significantly different. The problem of trend estimation for sea level data is complicated by the fact that the data are recorded using tide gauges which measure the height of the ocean surface relative to the adjacent land. Consequently, observed sea level trends result from a combination of vertical land movements and actual sea level variations. Here we consider two case studies, U.K. and Australia, where data on both spatial and temporal scales are used to estimate trends in extreme sea levels. These are discussed in Sections 2 and 3 respectively.

In the U.K. study the data are observations of sea level annual maxima from 62 sites. For estimation of the extreme sea level trend at a site, a simple approach is to use standard regression techniques and fit a linear model to the annual maxima data. This approach leads to highly variable trend estimates. However, by exploiting information about the form and spatial coherence of the error distribution, together with a simple model for the observed spatial dependence of the errors, a clear understanding is obtained about current extreme sea level trends. In particular a nonparametric model for the spatial trend is developed, giving a smooth estimate around the entire U.K. coastline.

For the Australian study observations of both annual mean and maximum sea level are available for 78 sites. Findings of Dixon and Tawn (1992) and Tawn and Mitchell (1992) suggest a simplified approach based on mean sea levels trends is appropriate here. The reasons for this choice and the development of the resulting statistical procedures will be discussed.

2 Trends in Extreme Sea Levels

This section summarizes methodology and results from Coles and Tawn (1990) and Dixon and Tawn (1992) for analyses of U.K. extreme sea level data. The data-set, described in detail by Graff (1981), consists of annual maximum sea levels from 62 mainland U.K. sites. In Sections 2.1 and 2.2 respectively the marginal annual maxima analysis and the spatial version of this approach are described and applied.

2.1 Marginal Analysis

At many sites the series of annual maxima appear to have a linear trend in mean level. For a site of interest, the distribution of the annual maximum sea level, Z_t, in year t is taken to be generalized extreme value, GEV(λ_t, γ, k), with $\lambda_t = \alpha + t\beta$, so

$$\Pr\{Z_t \leq z\} = \exp\left[-\{1 - k(z-\lambda_t)/\gamma\}^{\frac{1}{k}}\right], \text{ subject to } \{1 - k(z-\lambda_t)/\gamma\} > 0 \ \ \forall t, \qquad (2.1)$$

for $\lambda_t, \gamma(>0), k$, the location, scale and shape parameters respectively. The justification for using the GEV distribution is that it is the limiting distribution of linearly normalised maxima of stationary sequences (Leadbetter *et al.*, 1983). Therefore applications of this model require the number of observations in a year to be sufficiently large to justify the asymptotic arguments. Under this model the first two moments are easily shown to be

$$E(Z_t) = \alpha + t\beta + \gamma k^{-1}[1 - \Gamma(k+1)] \quad \text{and} \quad \text{Var}(Z_t) = \gamma^2 k^{-2}[\Gamma(2k+1) - \Gamma(k+1)^2]$$

where Γ is the gamma function. Thus it can be seen that β is the trend and that a linear model with independent and identically distributed GEV error structure though time is suitable to describe the annual maxima series for the site.

A statistic of interest for the design of flood defences is the return level, $q_t(p)$, defined as the level exceeded by the annual maximum in year t with probability p, i.e. $\Pr\{Z_t \leq q_t(p)\} = 1-p$. It follows from equation (2.1) that

$$q_t(p) = \alpha + t\beta + \gamma k^{-1}[1 - \{-\log(1-p)\}^k]. \qquad (2.2)$$

Maximum likelihood is used to fit the model to the series of annual maxima and substitution of the maximum likelihood estimates into (2.2) gives an estimate of the return level.

Clearly, accurate information about $q_t(p)$ is required by coastal engineers. In particular, for the design of sea walls, which are intended to be in use for many years, it is essential to have a good estimate for the trend parameter, β, as well as reasonable estimates for the other parameters. Here we concentrate on the trend parameter the others being covered adequately by Coles and Tawn (1990). Trend estimates, obtained by separately applying this method to each of the 62 sites in the study, are shown in Figure 1. Some spatial smoothness in these trend estimates is observed, but this feature is obscured by the estimates obtained at sites with short data series, poor quality data or biased sampling.

2.2 Spatial analysis

Let $Z_{i,t}$ be the annual maxima at site i for year t, so that $Z_{i,t} \sim$GEV$(\lambda_{i,t}, \gamma_i, k_i)$, with $\lambda_{i,t} = \alpha_i + t\beta_i$. A complete description of the marginal and spatial dependence structure of sea level annual maxima is given by $\Pr\{Z_{i,t} \leq z_i : \ i = 1, \ldots, n_s\}$, with n_s the number of sites in the study. The Coles and Tawn (1990) approach is to model this joint distribution of the annual maxima over data sites using a multivariate extreme value distribution accounting for spatial dependence of observations, and modelling the changes in each of the parameters of the

marginal distributions over sites in a way which is consistent with the properties of the underlying generating processes. The advantages of this over that of Section 2.1 are two–fold. First, by identifying the underlying relationship between the marginal parameters over sites increased precision of the GEV parameters is obtained and the estimates are robust to erroneous data typical in historical annual maxima. Secondly, accounting for the spatial dependence in the likelihood analysis allows a transfer of information between any two sites that have non–zero dependence. This occurs if all series are observed in each year (Shi *et al.*, 1992), but is most influential when some sites have missing data or when one site has a longer series of data than another. This information transfer can give a considerable increase in the precision of estimates, and overcomes problems associated with sampling bias.

Since the α_i parameters depend on the tidal characteristics of the site and the base year, these parameters are not spatially modelled. The parameters γ_i and k_i are characteristics of the extreme surges that occur near high tides and so should reflect the spatial continuity of the surge process. On the basis of Figure 1, Coles and Tawn suggest that the trend parameters, β_i, do not display enough spatial stability, and are too sensitive to local effects, to be sensibly modelled spatially. However, since the estimated values for β_i result from a combination of vertical land movements and actual sea level variations, the spatial variation in the β_i parameters will be complex, but smooth.

Dixon and Tawn (1992) use the spatial model and simultaneously estimate the trend parameters within a joint likelihood analysis. As with the marginal trend estimation, this gives a separate estimate of the trend at each site; however in this case the GEV error distribution is modelled spatially. The resulting trend estimates, plotted in Figure 2, have greater spatial stability than the corresponding trend estimates obtained using a separate site analysis.

To accommodate both the observed variation in the trend estimates over data sites and the expected coherence in the underlying trend an estimator which is both smooth and flexible is required. As no simple parametric form fits these trend estimates Dixon and Tawn (1992) use a weighted local linear regression estimator to obtain the estimate displayed in Figure 3. This figure shows significant spatial variation in trends around the coast: large trends occur in the South East and small or negative trends are found around the Scottish coast. Dixon and Tawn (1992) find that this spatial trend consists of two components: one due to vertical land movements and the other a homogeneous term of 1.0 to 1.2mm per year. The latter is consistent with trends in actual mean sea level (Shennan and Woodworth, 1992), hence there is no evidence for trends in extreme surge characteristics.

3 Trends in mean sea level

The results of Section 2 suggest that current information about trends in mean sea level is applicable to trends in extreme sea level and vice-versa. This gives the option of estimating extreme sea level trends using mean sea level data. The main advantage of using mean sea level data is that the inter-annual variation is smaller than in annual maxima data so, for a site with the same length record in both the mean and the maxima series, a more accurate trend estimator is obtained. On the other hand, more data processing is required to obtain mean level data than extreme level data which often leads to there being much more available annual maxima data.

In this section we focus on the mean sea level trend estimation approach. For the estimation of the mean sea level trend at a site, the usual approach is to use standard regression techniques to fit a linear model to the annual mean sea level data. In common with the approach of Section 2.1 this fails to exploit knowledge of the spatial dependence and coherence of the data. Let $W_{i,t}$ denote the annual mean sea level at site i in year t, then a complete description of the marginal and spatial dependence structure of the annual mean sea level is given by

$\Pr\{W_{i,t} \leq w_i : i = 1, \ldots, n_s\}$. Since mean sea levels are obtained as an average of the hourly observations over the year, the central limit theorem suggests that provided the number of observations in a year is sufficiently large an appropriate model for this joint distribution is the multivariate normal distribution. Thus marginally, for site i, the annual mean sea level is taken to follow a normal distribution with mean $\mu_{i,t} = a_i + tb_i$ and variance σ_i^2. If mean sea level and extreme sea level trends are the same then $b_i = \beta_i$.

The improvements in mean sea level trends estimates obtained from using this spatial model are similar to those obtained for extreme sea level trends in Section 2.2. However in this case it is possible to obtain analytical expressions for the trend and its standard error and so quantify the form of transfer of information between sites. Furthermore, since the multivariate normal distribution is a member of the exponential family the standard error is an exact form as opposed to a limiting form in the GEV case.

Here we illustrate one of these results using a simple bivariate case ($n_s = 2$). Let the annual mean sea levels at the sites be $X_{t_1}, \ldots, X_{t_n}$ and $Y_{t_1}, \ldots, Y_{t_m}$ respectively, where $m \leq n$ and the t_i are year indices, so the latter site has missing data in relation to the former. We take (X_t, Y_t) to be bivariate normal with means, variances and covariance given by $\mu_{X,t} = a_X + tb_X$, $\mu_{Y,t} = a_Y + tb_Y$, σ_X^2, σ_Y^2 and $\rho\sigma_X\sigma_Y$ respectively. It can then be shown that

$$\hat{b}_X = \frac{S_{t_n x}}{S_{t_n t_n}} \quad \text{and} \quad \text{Var}(\hat{b}_X) = \frac{\sigma_X^2}{S_{t_n t_n}}$$

and that $\hat{b}_Y$ satisfies

$$\hat{b}_Y = \frac{S_{t_m y}}{S_{t_m t_m}} + r_m \frac{S^*_{tx}}{S_{t_m t_m}} \quad \text{and} \quad \text{Var}(\hat{b}_Y) = \frac{\sigma_Y^2}{S_{t_m t_m}} \left[1 - \rho^2 \left(1 - \frac{S_{t_m t_m}}{S_{t_n t_n}}\right)\right].$$

Here $S_{t_j t_j} = \sum_{i=1}^{j} (t_i - \bar{t}_j)^2$, $S_{t_j x} = \sum_{i=1}^{j} (t_i - \bar{t}_j)(x_{t_i} - \bar{x}_j)$, $S^*_{tx} = \sum_{i=m+1}^{n} (t_i - \bar{t}_m)(x_{t_i} - \hat{\mu}_{X,t_i})$ and $r_m = [\sum_{i=1}^{m} (x_{t_i} - \hat{\mu}_{X,t_i})^2]^{-1} \sum_{i=1}^{m} (x_{t_i} - \hat{\mu}_{X,t_i})(y_{t_i} - \hat{\mu}_{Y,t_i})$, with $\bar{t}_j = j^{-1} \sum_{i=1}^{j} t_i$, $\bar{x}_j = j^{-1} \sum_{i=1}^{j} x_{t_i}$ and $\hat{\mu}_{X,t} = \hat{a}_X + \hat{b}_X t$. The trend estimate and standard error for the longer X series are unaffected by knowledge about the shorter Y series, whereas the trend estimator for the Y series is essentially the usual estimate adjusted, by the second term, for transferred information from the X series. The elements of the second term are of interest: S^*_{tx} measures how unrepresentative the X series is over the additional observation period relative to the overlapping observation period; r_m is a measure of dependence between the two series during the overlapping period; and $S_{t_m t_m}$ a measure of the information from the overlapping period. If either the additional observations are unrepresentative, there is no correlation in the overlapping series, the period of overlapping data is large, or the residual variance of the X series is large relative to the Y series then there is limited information transfer in terms of the trend estimate value. However if there is non-zero correlation then $\text{Var}(\hat{b}_Y)$ is reduced from the variance of the marginal trend estimator. When no data are missing ($n = m$) there is no information transfer. Further improvements are made by using a covariate model for the variance and covariance structure for which the only gain is a reduction in the variance of the trend estimate. In the multivariate case the analytical expressions are more complex, and consequently more difficult to interpret, but information continues to be transferred.

The Australian extreme sea level data-set, analysed by Tawn and Mitchell (1992), consists of both annual maxima and mean sea levels for 78 sites, of which only 30 have 15 or more years of data. Clearly here only a mean sea level trend estimation approach will provide the required information. There is some evidence that trends in extreme and mean sea levels are identical

since for each site Tawn and Mitchell (1992) analysed the $Z_{i,t} - W_{i,t}$ series and found no significant evidence of trends, despite finding evidence for extreme sea level trends. Preliminary numerical results for this example will be discussed in the presentation.

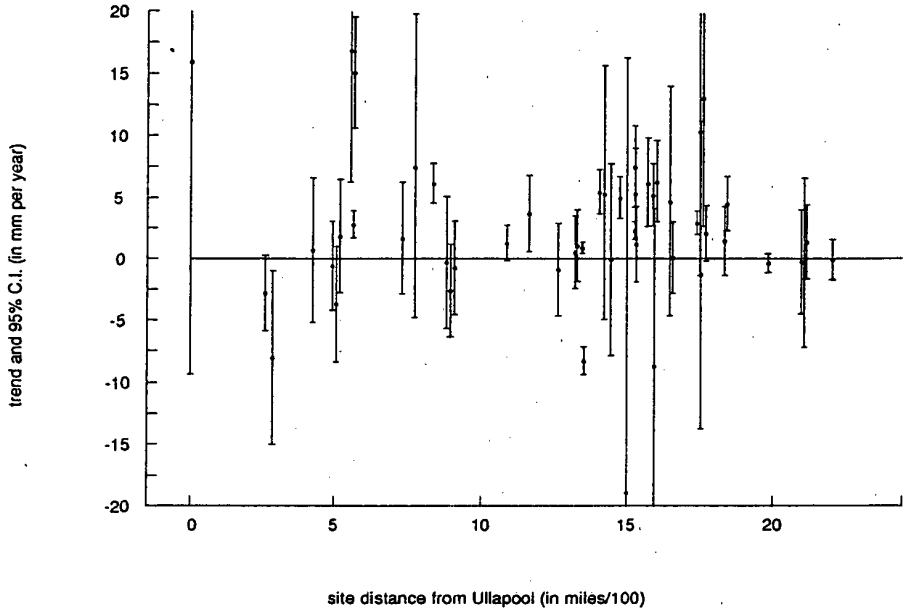


Figure 1: Trend estimates with 95% confidence intervals obtained by the marginal annual maxima method, described in Section 2.1. The abscissa in the plot is the site distance, as defined by Coles and Tawn (1990), measured anticlockwise from Ullapool – a site on the Western Scottish coast. The poor fitting sites have trend estimates which are inconsistent with those at neighbouring sites. As the graph wraps around, site distances of 0 and 2308 miles both correspond to Ullapool.

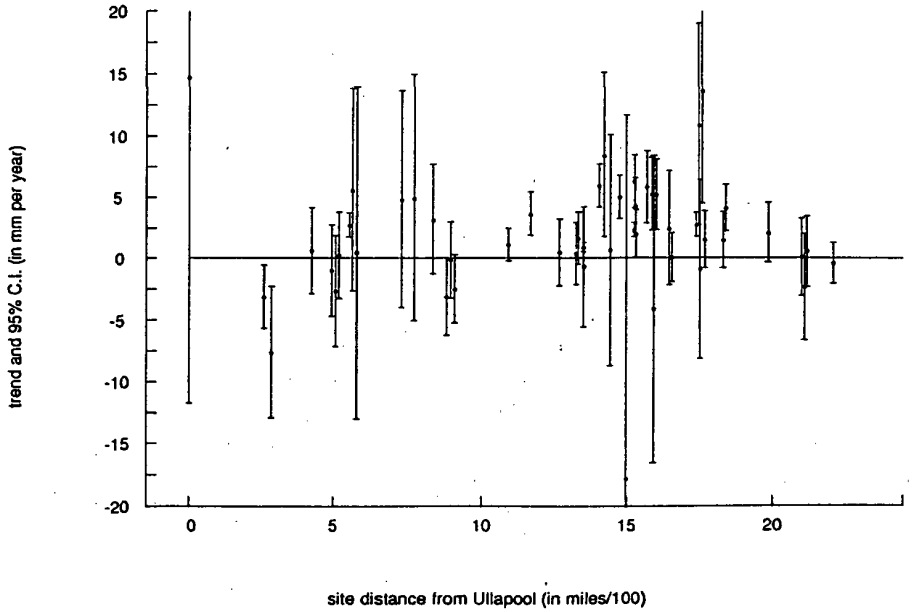


Figure 2: Trend estimates with 95% confidence intervals obtained by the spatial annual maxima method described in Section 2.2. The abscissa is the same as in Figure 1. The improvement in spatial smoothness is seen, particularly at some of the poor fitting sites as these now have trend estimates which are consistent with those at neighbouring sites.

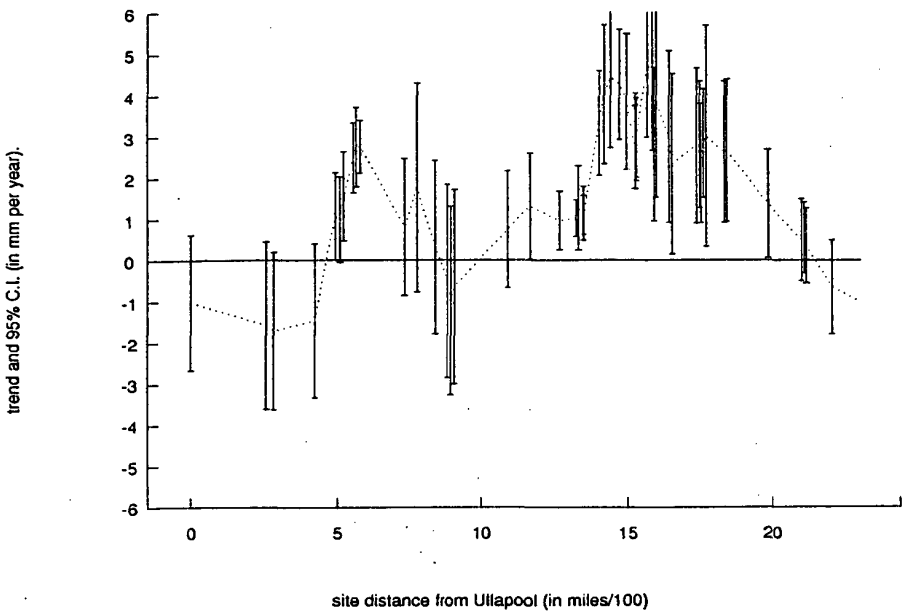


Figure 3: Spatial trend estimates and 95% confidence bounds obtained by the weighted local linear model of Dixon and Tawn (1992). The abscissa is the same as in Figure 1. The estimates have been linearly interpolated to give a clearer impression.

Acknowledgements

The work was partly supported by the Ministry of Agriculture Fisheries and Food, and grants from the Nuffield Foundation and the Royal Society.

References

Barnett, T. P. (1984). The estimation of 'global' sea level change: a problem of uniqueness. *J. Geophys. Res.*, **89**, 7980–7988.

Coles, S. G. and Tawn, J. A. (1990). Statistics of coastal flood prevention. *Phil. Trans. R. Soc. Lond., A*, **332**, 457–476.

Dixon, M. J. and Tawn, J. A. (1992). Trends in U.K. extreme sea levels: a spatial approach. Submitted.

Graff, J. (1981). An investigation of the frequency distributions of annual sea level maxima at ports around Great Britain. *Estuarine Coatal Shelf Sci.*, **12**, 389–449.

Leadbetter, M. R., Lindgren, G. and Rootzen, H. (1983). *Extremes and Related Properties of Random Sequences and Series.* Springer Verlag, New York.

Shennan, I. and Woodworth, P. L. (1992). A comparison of late Holocene and twentieth century sea-level trends from the U.K. and North sea region. To be published in *Geophys. J. Int.*.

Shi, D., Smith, R. L. and Coles, S. G. (1992). Joint versus marginal estimation for bivariate extremes. Submitted.

Tawn, J. A. and Mitchell, W. M. (1992). A spatial analysis of Australian extreme sea levels. Submitted.

Woodworth, P. L. (1987). Trends in U.K. mean sea level. *Mar. Geo.*, **11**, 57–87.

Modeling Daily Minimum Temperatures: An Application of the Threshold Method

Amy Marie Grady
Department of Statistics
University of North Carolina

ABSTRACT

A model using Extreme Value methodology is fitted to the daily minimum temperatures series of Wooster, Ohio. Two different approaches to removing seasonality from the data are presented. The method which blocks the data into homogeneous seasons is adopted on the grounds it provides consistent results with smaller standard errors. Next, a model incorporating a linear trend is fit to the data. The linear trend is not statistically significantly although it is of the same magnitude as the trend parameter, indicating that with proper modifications this procedure may yield significant results.

1. Introduction

It has been established that the global average surface temperature has risen between 0.30° Celsius and 0.60° Celsius in the last 50 years and that this increase can be attributed to the increase in the minimum temperatures(Karl et al, 1991b). What has caused this increase is not yet well defined; that is, many factors including cloud cover, aerosol loads, and the increase of greenhouse gases affect global warming trends and have yet to be modeled satisfactorily. Another concern in recent literature has been in establishing a statistical significance for this trend. This is due in large to the big year-to-year variability in the temperature series. Thus it is possible that an increase in global temperature could go undetected as a result of natural variability of the system(Karl et al, 1991a). The question then posed is as follows: Given that there has been an increase in global temperature and that this trend will continue in the future, how long must we observe the series until we detect a significant increase, and what data will give us the best evidence of this trend?

Much of the research on global warming has centered around the greenhouse effect; which is, that as the world becomes more industrialized, more greenhouse gases, mainly carbon dioxide but also methane, nitrous oxide and chloroflurocarbons(CFCs), are being released into the atmosphere. These emissions absorb infrared radiation and cause the earth's atmosphere to warm. The current models predict that if the carbon dioxide level in the atmosphere were to double then the temperature would increase between 2° C and 4° C(Karl et al 1991a).

Climatologists believe that analyzing the temperatures of nights, especially cloudless nights, will best reflect information on the greenhouse effect. If the nighttime temperatures do show an increase, this would imply that the heat to the day is not being dissipated but is being trapped by the earth's atmosphere. Thus the climatologists advocate analyzing minimum temperature series to best mirror the cloudless nights(Michaels, 1991). The current results do show that the rise of global temperatures can be attributed to a rise in the minimum temperatures. Karl et al(1991b) have analyzed data from the former Soviet Union, the United States of America, and the People's Republic of China and after extracting the rise in temperature due to urbanization have found that while the mean maximum temperatures show no significant increase, the mean minimum temperature show a substantial increase. This asymmetry is most striking when studying the decrease in the mean and absolute temperature ranges. These results support the greenhouse effect as being a substantial factor in the rise of the global temperature(Karl et al., 1991b).

This paper proposes a method of analyzing a daily minimum temperature series using an extreme value methodology. Karl et al(1991b) have demonstrated the importance of studying the minimum temperatures series and since the data represent extreme values, the methodology advocated here is extreme value theory. The method is developed by analyzing one minimum temperature series. Although no significant trend in the data is expected to be established by studying only one station, this methodology can be extended to a multi-site setting. Specifically, the model proposed to is a Possion process approach with a GEV limiting intensity function. The generalized extreme value distribution (GEV) arises naturally as the limiting

distribution of extreme values. In fact, this family represents the entire family of limiting distributions for linearly normalized maxima or minima of identically distribution random variables with weak mixing. This model has been successfully applied to many climatological areas including pollutant levels(Smith, 1989), tide levels(Smith, 1986), and wind speeds(Coles and Walshaw, 1991).

2. Theoretical Background

In this section, a brief outline of the major developments in extreme value theory will be presented. The section concludes with the Poisson process to extreme value theory. This last approach will be the model adopted for this paper.

Classical extreme value theory studies the distribution of the maximum of a sequence of independent and identically distributed (*iid*) random variables. In particular, it has focused on the limiting distribution of the maximum of the sequence.

Suppose $X_1, \ldots, X_n$ are *iid* random variables with a common distribution F(x) and define $M_n = \max(X_1, \ldots, X_n)$. To obtain a non-degenerate limiting distribution function of M_n, one seeks normalizing constants $\{a_n\}$ and $\{b_n\}$ such that, after proper normalization, M_n converges in distribution to a non-degenerate distribution function, H(x), that is :

$$\mathbb{P}\left\{\frac{M_n - b_n}{a_n} \le x\right\} = F^n\left\{a_n x + b_n\right\} \to H(x) \quad \text{as } n \to \infty$$

The above limit holds if and only if (*iff*) for all fixed real values x with $0<H(x)<1$,

$$\lim_{n\to\infty} n\left\{1 - F(a_n x + b_n)\right\} = -\log H(x)$$

The central result in extreme value theory, formulated first by Fisher and Tippett(1928) and Gnedenko(1943), is that H(x) can only belong to three different types of distributions. For statistical purposes, it is convenient to have a single family of extreme value distributions and this can be achieved by rewriting the three types into a single Generalized Extreme Value distribution (GEV):

$$H(x\,;\,k,\sigma,\mu) = \exp\left\{-\left(1 - \frac{k(x-\mu)}{\sigma}\right)^{1/k}\right\}$$

$$\text{where } \left(1 - \frac{k(x-\mu)}{\sigma}\right) > 0 \text{ and } k, \mu \in \mathbb{R} \text{ and } \sigma > 0.$$

The parameter k is the shape parameter and describes the tail behavior of the parent distribution, the distribution of the observations. The case, k=0, should be interpreted as the limit as k tends to zero where H(x) becomes:

$$\lim_{k\to 0} H(x;\,k,\sigma,\mu) = H(x;\,0,\sigma,\mu) = \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}$$

This case describes parent distributions with exponentially decreasing tails. The case, k<0, describes parent distributions with polynomially decreasing tails. Finally, if k>0, this implies that the parent distribution has a finite upper endpoint and thus a short tail.

The scale parameter is σ and relates of the spread to the data. Finally, μ is the location parameter and fitting a linear trend is equivalent to requiring μ to be linear in time. Thus to fit a linear trend, one sets $\mu = \alpha + \beta t$ where α and β are real numbers and t is time.

The standard approach is to fit the annual maxima of a series to a distribution usually via a numerical maximum likelihood. The GEV distribution has the advantage that it represents an entire family of extreme value distribution. Thus fitting the GEV distribution can accommodate a wide variety of applications.

An alternative approach in extreme value theory has been in modeling the exceedances above a high threshold. Here, one defines an exceedance, Y_i, as the amount above the threshold; that is, $Y_i = X_i - u$ where X_i is the observed level and u is the threshold. Threshold methods have been extensively used in hydrology literature under the acronym POT -- peaks over thresholds -- since the 1970's. The simplest approach to threshold methods are based on a Poisson Process of exceedances times with independent excesses over a high threshold. Here, Pickands(1975) has proved that the Generalized Pareto Distribution(GPD):

$$G(y;\,k,\sigma) = 1 - \left(1 - k\frac{y}{\sigma}\right)^{1/k}$$

where k, σ are real numbers with $\sigma > 0$,

is the limiting distribution for the exceedances over a threshold under the same conditions that the annual maximum have the limiting distribution of a GEV. The case, k=0, again should be interpreted as the limit as k tends to zero and the here the GPD reduces to the exponential distribution, $1 - \exp\{-y/\sigma\}$.

This approach of fitting the GPD to extreme values is very flexible and can accommodate situations that the ordinary GEV fitting cannot, such as small data sets,

complex trends, or dependent sequences.

Another way of formulating the threshold methods is through a point process approach. This procedure was first introduced by Pickands(1971) for *iid* random variables and was extended to stationary sequences by Adler(1978) who normalized the process using the constants a_n and b_n from the asymptotic distribution of M_n. Here one is interested in the characteristics of the limiting point process as the number of observation and the threshold tend to infinity.

Here, let $\{X_n\}$ be a stationary sequence of random variables and suppose that the sequences $\{a_n\}$ and $\{b_n\}$ exist such that

$$\mathbb{P}\left\{\frac{M_n - b_n}{a_n} \le x\right\} = F^n\left\{ a_n x + b_n \right\} \to H(x) \quad \text{as } n\to\infty$$

where H is the generalized extreme value distribution.

Now define a point process, P_n, on the plane at the points $\{ i/n, (X_i - b_n)/a_n\}$. Note that the norming constants a_n and b_n incorporate the threshold; that is, the threshold can be written as $u_n = a_n x + b_n$.

Then under suitable mixing conditions, specifically the $D_r(u_n)$ and $D'(u_n)$ conditions of Leadbetter, Lindgren, and Rootzen(1983), then the point process, P_n, converges in distribution to a nonhomogeneous Poisson process whose mean rate of exceedance in a time interval, say (t_1, t_2), is

$$\left(t_2 - t_1 \right) \times \left(1 - \frac{k(x-\mu)}{\sigma} \right)^{1/k} \qquad \text{where } ( ) \text{ is positive and } 0 \le t_1 < t_2 \le 1.$$

The conditions $D_r(u_n)$ and $D'(u_n)$ limit the amount of dependency in the sequence. In particular, the $D_r(u_n)$ condition requires that any two disjoint subsets of the sequence be asymptotically independent of each other and the $D'(u_n)$ limits the amount of clustering that can occur as the process continues.

Then in terms of the classical approach, the $\mathbb{P}[(M_n - b_n)/a_n]$ is equivalent to the probability that the Poisson process has no points in $(0,1] \times (x,\infty)$. The other approaches can be derived, including the GPD.

This approach has all the benefits of a threshold method but also can accommodate missing data more easily. Also, this approach is useful since one can fit the GEV parameters which are easily adapted for fitting a trend.

3. Statistical Methodology

The analysis in this paper begins to model continental United States temperatures using extreme value methodology. The Poisson process approach is taken in this analysis, mixed with a threshold method. The method of estimation is by numerical maximum likelihood. The Poisson process model with a GEV intensity is fitted to the cluster maxima of exceedances above a fixed threshold. Clusters are formed to reduce the dependency in the data. This method is based on the property that if N is Possion, then conditionally on N, maximum of N generalized Pareto random variables has a GEV distribution. Thus the likelihood function to be maximized with respect to the GEV parameters is:

$$\mathcal{L}(\theta; z_1, \ldots, z_n) = \prod_{i=1}^{N} \lambda(z_i, \theta) \exp\left\{ - \int_A \lambda(z, \theta) dy \right\}$$

for $\theta = (k, \sigma, \mu)$, $z_i = (y_i, t_i)$, where y_i is excess and t_i is the exceedance time. A is the region of observation. Note that A may be different than the time interval due to missing data. In this approach, one can actually model the period of observation, excluding missing values. This is significant advantage of this approach. Here,

$$\lambda = \frac{1}{\sigma}\left\{ 1 - \frac{k(y_i - \mu)}{\sigma} \right\}^{1/k} \qquad \text{for } \{ \} \text{ positive}$$

The program used here which computes the maximum likelihood estimates also produces the estimates' standard errors along with the 100 year return level and its standard error. The 100 year return level represents that temperature which is expected once in 100 years. The formula is derived as $\mu + k(1 - 0.01^k)/\sigma$ and its standard error is calculated via the delta method.

The important issues which must be resolved when fitting the GEV distribution are selecting the cluster separation, addressing seasonality, and choosing the appropriate threshold.

The issue of cluster separation is important since like most climatological data, there is a strong degree of dependency in the daily minimum temperature data. Therefore, a cluster separation must be established so that each cluster is relatively independent of the other clusters. A cluster separation is defined as the number of days needed between two exceedances before they are considered independent excedances.

Seasonality is important due to the identically distributed assumption underlying the GEV approach. If the characteristics of the observations from one part of the year to the next are not similar, then modeling them together would mix different distributions and would bias any model fit. Indeed, the model fit would not describe what is happening at any time. Thus it is important to eliminate the seasonal effect. One way this can be accomplished is to prewhiten the data set by estimating the seasonal effect and subtracting it from the data. Another approach is to break up the data into different seasons, to block the data, and then model each season separately.

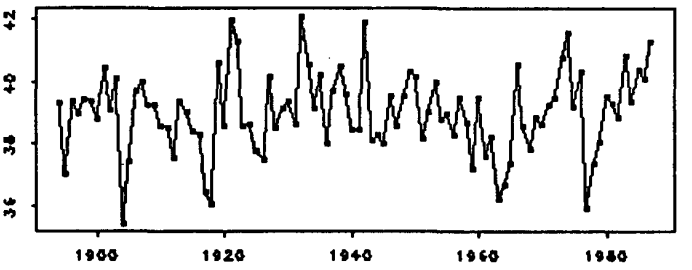
The final critical issue is selecting an appropriate threshold. If the threshold is set too low, then the data modeled will not represent extreme values and will bias the fit of an extreme value distribution. If the threshold is selected too high, there will not be enough exceedances to model.

4. The data

The data analyzed comes from the National Oceanic and Atmospheric Administration (NOAA). They have selected 138 weather stations in the continental United States on the basis of length of record and quality of data. They have also cleaned up the data set by estimating many of the missing data and labeling any changes which would affect the data set. From this data set, I have selected Wooster Ohio to analyze. It is a representative site for a number of reason. The series dates between June of 1893 to December of 1987 and consists of 34,518 days of which only 17 days are missing. Due to its length it should insure a closer fit to the limiting distribution. Also, Wooster is located in the central part of Ohio and is not close to any significant body of water or mountains which could affect temperatures. Finally, Wooster is the center of an agricultural zone and although more urbanized than in 1893, it will not be greatly effect by the rise of temperatures due to an increased population, the urban effect.

A plot of the daily minimum temperature series revealed that the warmer temperatures in the series are more compact, rarely exceeding 70° Fahrenheit. On the other hand, the lower temperatures, the extreme minimum temperatures, show a great variability. In other words, the temperatures in the winter nights in Ohio can be very different, especially from one year to another. This variability can best be seen in Figure 1, a plot of the annual average minimum temperatures. The spread of the average minimum temperatures is

Figure 1
Annual Average Minimum Temperatures



from 35.5°F to 42.0°F and more importantly a lage fluctuation can take in just a few years. Although there does not appear to be any significant overall trend, the annual minimum temperatures in the middle of the data set, from approximately 1920 to 1960, tend to be higher than at the beginning or the end of the series. A linear trend, representing either a warming or a cooling, does not appear to be present.

5. The Data Analysis

The intent of this analysis on the daily minimum temperatures using an extreme value distribution is to investigate the standard errors associated with the parameter estimates, especially the trend parameter. The analysis is broken up into three parts. First, a model which utilizes the prewhitening approach to removing the seasonality is presented and then a model which uses the block version is discussed. Finally, after deciding which approach is most successful, a model incorporating the trend parameter is fitted. The goal is to obtain the "smallest" standard error as possible for the trend.

The focus of this analysis is to model the extreme minimum temperatures. Therefore, the exceedances of interest will be the minima of the data set or the lowest minimum temperatures. Although extreme value theory was developed for exceedances which refer to the maxima of a series, the theory can be applied to minimum exceedances. In this analysis, an exceedance should be thought of as how far a temperature falls below the prescribed threshold.

In the prewhitening approach, a daily average minimum temperature was calculated using a 5 day moving average. This moving average was then subtracted away so that now the data being modeled are differences from the "normal". Note that the minima of this series could represent either a winter night which is colder than normal or a summer night which is colder than normal.

Table 1
Prewhitening Model

Month	k	σ	μ	100 year return
1	0.33(0.07)	9.77(0.97)	26.87(3.23)	26.87(3.23)
2	0.28(0.05)	9.07(0.80)	3.04(0.85)	26.84(3.22)
3	0.11(0.11)	5.78(0.87)	-0.09(0.59)	20.93(18.85)
4	-0.01(0.12)	2.29(0.45)	-0.75(0.31)	10.13(93.63)
5	0.24(0.07)	3.19(0.39)	0.05(0.32)	8.86(1.79)
6	0.54(0.14)	5.72(1.45)	-2.69(0.92)	7.05(0.75)
7	0.34(0.15)	4.17(1.59)	-3.78(1.36)	5.99(2.12)
8	0.31(0.11)	3.39(0.87)	-2.24(0.71)	6.03(1.49)
9	0.23(0.11)	3.04(0.48)	-0.25(0.33)	8.37(2.96)
10	0.33(0.11)	3.68(0.59)	-0.34(0.59)	8.35(1.72)
11	0.10(0.15)	4.18(1.02)	-1.66(0.63)	13.86(20.64)
12	0.35(0.13)	8.78(0.99)	2.02(0.83)	22.11(3.44)

(Standard errors in parenthesis)
NOTE: The 100 year return refers to how far below normal the temperature can be expected to drop once in 100 years.

The first problem presented was that simply subtracting away the seasonal average did not remove all the seasonality in the data. Both the number of exceedances expected in each month and the spread of these exceedances depends on the month and has not been accounted for in the prewhitening. Thus it is not appropriate to group the months and each month should be and is analyzed separately.

Both selecting a cluster separation and a threshold were straightforward. Neither varying the threshold nor the cluster separation greatly affected the model. Any level of cluster separation and threshold fit gave consistent 100 year return levels, the lowest temperature expected in 10years. Thus a cluster separation of 3 days and a threshold of 15° F below normal were adopted since these reduced the standard error of the parameters by allowing more observations to be fit.

The parameter estimates, including the 100 year return level, for this model are in Table 1. The shape parameter, k, is positive indicating a finite lower bound on the minimum temperature. The scale parameter, σ, is larger in the winter months representing a greater variability in the winter months' temperatures. Also, the location parameter, μ, and the 100 year return level are larger in the winter indicating that the temperature in the winter can be expected to drop farther below normal than in the summer.

Clearly, there is a big variation between the months. This strengthens the argument to divide the data set into monthly seasons. Simply subtracting the seasonal mean does not eliminate the seasonality in the data. The deviations away from normal are not homogeneous throughout the year.

Table 2
Blocking Model

Month	k	σ	μ	100 year return
1	0.52(0.04)	14.52(0.69)	17.12(1.31)	42.44(0.97)
2	0.56(0.05)	15.12(0.84)	14.72(1.37)	39.81(0.99)
3	0.32(0.06)	8.62(0.56)	6.57(0.81)	27.20(2.77)
4	0.10(0.21)	4.34(2.21)	-5.10(2.17)	10.77(25.50)
10	1.00(0.03)	38.68(0.94)	-34.78(0.94)	3.61(0.04)
11	0.18(0.07)	6.23(0.69)	1.01(0.59)	20.56(6.54)
12	0.44(0.05)	12.78(0.70)	12.07(1.17)	37.02(1.50)

(Standard errors in parenthesis)
NOTE: The 100 year return level refers to how far below 20°F the temperature can be expected to drop once in 100 years.

Next, a model based on blocking the data so as to create homogeneous seasons was established. Here, a low threshold is fixed and the data consist of exceedances below that threshold In fact, only the colder months will have exceedances to model.

To decide seasonal length, models were fitted based on monthly and bi-monthly seasons. The results were clear. Two-month periods, except maybe for January and February, are not homogeneous and should not be grouped. The model fit closely adheres to the colder month and gives no relevant information on the warmer month.

As for cluster separation and threshold selection, again the model was not sensitive to various levels. Given the stability of the model, levels were chosen so as to minimize the standard error. A cluster separation of 3 days was again adopted. An exceedance was defined as any temperature below 20°F.

The results for this model are in Table 2. The characteristic of the parameters are similar to the prewhitening approach. The parameter, k , is positive and the others are largest in the winter season.

Care should be taken when comparing the models since although both eventually model each month separately, the moving average approach models deviations from the normal while the blocking method models the exceedances below a fixed threshold. In particular, the 100 year return levels represent different values. What is evident is that the standard errors for

the parameters are consistently smaller for the block style method. This may be attributed to the block style modeling more observations each month although at this time the difference between the standard errors is not fully understood.

Due to this substantial decrease in the standard error by using the block method, the block style model was incorporated when fitting the trend parameter. Although one does not expect the trend to be significant given Figure 1, this approach will provide the lowest possible standard error obtainable from the models presented here. The estimates and their standard errors of this model are in Table 3, except for the k parameter which is almost identically the same as for the model without the trend. Using the rule of thumb that a parameter is significant if its standard error is less than one half is magnitude, one observes that although the trend parameter is not significant (and negative!), the standard error is at least of the same magnitude. This is exciting due to the limited scope of the data; that is, the above analysis focused only on one site. It would be surprising if the trend was significant due to the variability in one series. These results give hope that when this model is adapted to a multi-site version, there will be enough information to obtain a significant trend.

Table 3
Blocking Model with Trend

Month	σ	α	β	100 year return
1	3.81(0.38)	38.62(1.46)	-0.03(0.03)	41.78(2.22)
2	3.52(0.37)	36.58(1.57)	-0.03(0.03)	39.54(1.47)
3	3.90(0.45)	21.75(1.63)	-0.01(0.03)	30.12(3.06)
4	3.35(0.61)	6.05(1.74)	-0.04(0.03)	15.22(32.49)
11	4.03(0.54)	14.93(1.44)	-0.03(0.02)	24.73(7.51)
12	4.12(0.41)	32.83(1.87)	0.04(0.04)	36.81(2.81)

(Standard errors in parenthesis)

6. Conclusions

Although the one station analyzed here does not provide significant results for a trend in the data, the methodology used appears reasonable. The need to analyze minimum temperatures has clearly been defined by Karl et al(1991b). Since this series represents extremes in the data, one should be fitting an extreme value distribution.

In fitting the model, what needs further investigation is why different methods of eliminating seasonality should affect the standard errors so dramatically and how best to modify the model to incorporate many different stations. Given the flexibility of the model proposed and the size of the standard errors produced, this model once extended to a multi-site version may provide a effective extreme value approach to global warming. It is expected that a multi-site model will provide a significant trend.

Acknowledgements: I wish to express my gratitude to Professor Richard Smith for his continuous help and encouragement in the preparation of this project. I would also like to thank Professor Walter Smith for his help in formatting the document.

REFERENCES

Adler, R.J. Weak convergence results for extremal processes generated be dependent random variables, *Ann. Prob.*, **6**,660-667.

Coles, S.G. and Walshaw, D. (1991). Directional Modeling of Extreme Wind Series. Research Report no. 383/91. Department of Probability and Statistics, University of Sheffield, UK.

Fisher, R.A. and Tippett, L.H.C.(1928). Limiting Forms of the Frequency Distributions of the Largest of Smallest Member of a Sample, *Proc. Camb. Phil. Soc.* **24**, 180-190.

Gnedenko, B.V.(1943). Sur la Distribution Limite du Terme Maximum d'une Serie Aleatoire, *Ann. Math.* **44**, 423- 453.

Karl, T.R., Heim, R.R. and Quayle, R.G(1991a). The Greenhouse Effect in Central North America: If Not Now, When?, *Science*, **251**, 1058-1061.

Karl, T.R. et al(1991b). Global Warming: Evidence for Asymmetric Diurnal Temperature Change, *Geophys. Res. Lett.*, **18**, 2253-2256.

Leadbetter, M.R., Lindgren, G. and Rootzen, H(1983). *Extremes and Related Properties of Random Sequences and Series.* Springer Verlag, New York.

Michaels, P.(1991). Global Pollution's Silver Lining, *New Scientist*, **23**, 40-45.

NOAA(1991). Historical Climatological Network--Daily stations, United States Department of Commerce, Asheville, North Carolina.

Pickands, J.(1975). Statistical Inference Using Extreme Order Statistics, *Ann. Statist.*, **3**, 119-131.

Smith, R.L.(1986). Extreme Value Theory Based on the r Largest Annual Events. *J. Hydrology*, **86**, 27-43.

Smith, R.L.(1989). Extreme Value Theory for Environmental Time Series: An Example Based on Ozone Data, *Statistical Science*, **4**, 367-393.

Weissman, I.(1978). Estimation of Parameters and Large Quantiles Based on the k Largest Observations, *J. Amer. Statist. Assoc.*, **73**, 812-815

FURTHER STUDIES OF SKILL CHARACTERISTICS OF U.S. LONG-RANGE FORECASTS. PART I: SURFACE TEMPERATURE AND PRECIPITATION FORECASTS

by

Robert E. Livezey and Jonathan D. Hoopingarner

Climate Analysis Center/NMC/NWS/NOAA

A program of modernization and automation of the U.S. Climate Analysis Center's Prediction Branch has led to the creation of accessible digital files of monthly and seasonal forecasts. These include forecasts of monthly U.S. temperature and precipitation categories (in three classes) and hemispheric 700 mb height anomalies produced 24 times a year since 1974, and of seasonal U.S. temperature and precipitation categories produced four times a year from 1958 through 1981 and 12 times a year thereafter. Probability forecasts are also available for all surface forecasts since 1982 and surface forecasts for Alaska/Canada and mid- and high-latitude Eurasia are available since 1987.

These data form the basis for a program of detailed documentation of the variability over time, and by season, location, parameter, class, regime, etc., of the skill of U.S. monthly and seasonal prediction practices. Ideally the information generated will not only prove useful for management of forecast operations and for the forecasters themselves, but also to potential users in deciding whether to include the forecasts as part of their particular decision-making process (Livezey, 1990). In this first part attention will be focused on temperature and precipitation forecasts and highlights of their analysis. In part II a number of aspects of the skill characteristics and history for monthly mean N.H. 700 mb height anomaly forecasts will be presented.

For the categorical forecasts examined here a skill score that measures the percent of hits above that expected from random forecasts is most heavily employed. Time series and annual cycles of map skills, maps of local skills, and summary contingency tables are produced for all surface forecasts and examined carefully for meaningful variability.

One example of variability that will be presented is a large increase in the difference in skill between official U.S. monthly mean temperatures forecasts and forecasts of persistence between the 1970's and 1980's. This increase in the advantage of the official forecasts over persistence is statistically significant and can be partially attributed to improvements in global numerical weather prediction (Kalnay and Livezey, 1985). Examination of skill climatologies reveals that the advantage is

entirely realized during the cold half of the year when persistence skill is particularly low and official skill particularly high.

Other characteristics of the skill of monthly temperature forecasts will be illustrated through contingency tables and skill maps that have been tested for field significance. The discussion will than focus on the less skillful seasonal temperature forecasts and finally on the forecasts with least skill of all, monthly and seasonal precipitation forecasts.

FURTHER STUDIES OF SKILL CHARACTERISTICS OF U.S. LONG-RANGE FORECASTS. PART II : N.H MONTHLY MEAN 700 MB HEIGHT ANOMALY FORECASTS

by

Jonathan D. Hoopingarner and Robert E. Livezey

Climate Analysis Center / NMC / NWS / NOAA

Forecasts of monthly mean Northern Hemisphere sub- and extratropical 700mb height anomalies have been routinely made and archived in the Prediction Branch of the Climate Analysis Center (CAC) since the early 1970's. The addition of new equipment and personnel allowed CAC to institute operational digitization and archival of these height forecasts in August of 1989. A grant from the U.S. Forest Service made the wholesale digitization of the historical backlog feasible and work was completed in May 1990.

Both operational and historical digitization is accomplished in an interactive process which allows for extensive quality control. Verifications are performed on a 545-point diamond grid for compatibility with existing observational data sets.

Forecast skill is measured by the percent reduction in squared error from that of a climatology (zero-anomaly) forecast. This score is -1 for random forecasts, zero for climatology forecasts and +1 for perfect forecasts. Murphy and Epstein (1989) have shown that it can be decomposed into four terms that describe the error in phase of the forecast, the error due to difference of amplitude, the overall or unconditional bias of the forecast anomalies, and the error due to mismatched climatology. For large samples the latter two mostly compensate each other. The square root of the first term is a correlation, either anomaly or temporal, depending on whether the skill is for a forecast map at a particular time or a forecast history at a particular gridpoint respectively. The other terms, of course, also have spatial and temporal analogs.

Both types of skills, map and local history, and their decompositions have been computed for forecasts from 1973 through 1990. Thus, two types of non-summary presentations are possible: time histories of map skills and their constituent terms and maps of local history skills and their constituent parts.

The time series show clear trends in both the phase error (a decrease) and in the bias term (from too high forecast heights in the 1970's to too low in the 1980's). The former result is

obviously related to the trend in advantage of the operational U.S. monthly mean surface temperature forecasts over persistence forecasts described in the companion talk. Mean annual cycles of both terms over the last decade show that the skill in forecasting phase resides in cold-season forecasts and that the cold bias is a result of spring and fall forecasts when signal-to-noise ratios are large and observed warm biases cannot be confidently anticipated by forecasters.

Maps of winter local history (last decade) skills reveal that overall skill at this time of the year is a result of skill in both the timing and amplitude of events and is concentrated over the high-latitude oceans (including the Arctic) and to a lesser extent over the eastern half of North America. Forecasts almost mimic random forecasts over eastern Asia.

Results for other terms and other seasons will also be presented.

VERIFICATION OF MONTHLY MEAN OUTLOOKS FOR FIRE WEATHER ELEMENTS IN THE CONTIGUOUS UNITED STATES

William H. Klein and Joseph J. Charney
Department of Meteorology
University of Maryland
College Park, MD 20742

Morris H. McCutchan and John W. Benoit
USDA Forest Service
4955 Canyon Crest Drive
Riverside, CA 92507

In a recent paper, Klein and Whistler (1991) described a system for specifying monthly mean anomalies of midday temperature (T), dewpoint (D) and wind speed (W) at 127 surface stations in the contiguous United States. Multiple regression equations containing from 1 to 6 terms were derived for each element and month from concurrent fields of 700 mb height anomaly (H) observed over North America and adjacent areas, plus the previous month's observed local anomaly of T, D or W, during the 20-year period: 1964-1983. The specification equations are now being applied routinely by the U. S. Forest Service (Fujioka and McCutchan, 1990).

In this paper, we present the results of testing the above forecast system on two types of independent data: 1) observed values of H during the seven years from 1984 to 1990 and 2) prognostic values of H prepared operationally twice a month at the NOAA Climate Analysis Center (CAC) during the 18 years: 1973-1990. Note that we verified at 122 (instead of 127) stations due to missing data and converted all anomaly forecasts into percentiles (P). Also, we included mid-month to mid-month forecasts in our test in order to increase the sample size, and we pooled the data for any one period with its four adjacent periods to reduce the noise.

Our primary statistic was the Heidke skill score (HSS). It was computed for 2 classes (above or below the 50th P), 3 classes (based on the 30th and 70th P used by CAC), 4 classes used operationally by the Forest Service (P below 50, 50-74.9, 75-89.9, and 90 or above) and 5 classes (as above but with the 4th class subdivided at P>97.5). Fig. 1 shows the four different values of HSS obtained for each weather element by pooling the data for all months and all stations for prognostic heights from 1973 to 1990. A similar chart was prepared for observed heights from 1984 to 1990. The results were remarkably consistent in showing a rapid fall of all scores from 2 to 3 classes and a slow decrease from 4 to 5 classes. In all cases, the skill was positive, higher for observed heights, highest for T and lowest for W. The relative rank of the elements is the same as that obtained in the original derivation.

Gilman (1986) previously found that CAC monthly forecasts have more skill for "above" and "below" normal classes than for near normal (see also Livezey, 1990, and Van den Dool and Toth, 1991). We found a similar result when we computed the values of HSS separately for each class. For example, Fig. 2 shows the scores for T specified from 18 years of prognostic heights in a 4-class system for January, April, July, October and the entire year. In every case, the 1st and 4th classes outperformed the middle two classes. It is especially encouraging that forecasts of P > 90, the most critical category for fire outbreaks, consistently had greater skill than that for any other class. It is also noteworthy that the scores were generally highest in January and lowest in October, in agreement with CAC experience. Similar results were obtained for D and W and for use of observed, rather than prognostic, 700 mb heights.

The positive values of HSS discussed thus far indicate that the forecasts based on prognostic 700 mb heights were more skillful than climatology or a random guess. However, Table 1 shows that they improved over local month-to-month persistence only for T and D but not for W. This was evident for all five statistics tabulated except bias. Part c) of the table verifies specifications made by assuming that H could be forecast perfectly. In this case, even the wind forecasts were superior to persistence (except for bias); while the skill scores for T and D were 2-3

times as large as those for prognostic heights and about 5 times as large as the persistence scores.

The geographical distribution of the mean annual HSS for temperatures specified into four classes from 18 years of prognostic heights is mapped in Fig. 3. Scores range from over 13 in the western Great Lakes region and Ohio Valley to under 6 in parts of New England and the west coast states.

Fig. 4 shows that specifications of T and D from prognostic heights were more skillful from 1984 to 1990 than from 1973 to 1990. Furthermore, this improvement occurred in a period of *lower*, rather than higher, persistence. Therefore, the CAC prognostic heights on which the specifications are based were probably more accurate during the second period, as implied in a recent paper by Hoopingarner and Livezey (1992).

Acknowledgments

This research was supported by the U.S. Department of Agriculture, Forest Service, Riverside, California under Cooperative Agreement No. PSW900081CA. We are grateful to Francis Fujioka of the Forest Service, Alan Robock of the University of Maryland, and Paul Dallavalle and Ed O'Lenic of the National Weather Service for help in various aspects of this project.

REFERENCES

Fujioka, F. M. and M. H. McCutchan, 1990: Long-range fire weather forecasting. *Proc. 14th Annual Climate Diagnostics Workshop*, La Jolla, CA, Oct. 16-20, 1989. Climate Analysis Center, NOAA, Washington, DC, pp. 410-415.

Gilman, D. L., 1986: Expressing uncertainty in long-range forecasts. *Namias Symposium*. John O. Roads, Ed., Scripps Institution of Oceanography reference series 86-17, 174-187.

Hoopingarner, J. D. and R. E. Livezey, 1992: Further studies of skill characteristics of U. S. long-range forecasts. Part II: N. H. monthly mean 700 mb height anomaly forecasts. Proc. 16th Annual Climate Diagnostics Workshop, Los Angeles, CA, Oct. 28-Nov. 1, 1991. Climate Analysis Center, NOAA, Washington, DC, pp. 434-440.

Klein, W. H. and B. T. Whistler, 1991: Specification of monthly mean anomalies of fire weather elements in the United States. *Agricul. Forest Meteor.*, **56**, 145-172.

Livezey, R. E., 1990: Variability of skill of long range forecasts and implications for their use and value. *Bull. Amer. Meteor. Soc.*, **71**, 300-309.

Van den Dool, H. M. and Z. Toth, 1991: Why do forecasts for "near normal" often fail? *Wea. Forecasting*, **6**, 76-85.

Table 1. Verification of three methods of estimating monthly percentiles for midday temperature, dewpoint and wind speed at 122 stations in the contiguous United States during each month for seven years from 1984–1990.

Statistic	Temperature	Dewpoint	Wind Speed
a) Simple month–to–month persistence:			
Mean absolute error	30.5	31.9	30.5
Bias (forecast–observed)	–0.3	–0.0	–0.1
Correlation coefficient	0.16	0.13	0.28
Skill score (2 classes)	10.3	9.8	19.0
Skill score (5 classes)	4.7	5.3	10.5
b) Specification equations applied to prognostic 700 mb heights:			
Mean absolute error	27.1	27.3	30.1
Bias (forecast–observed)	–3.7	–1.3	–5.1
Correlation coefficient	0.23	0.19	0.13
Skill score (2 classes)	17.0	15.2	11.0
Skill score (5 classes)	9.5	10.2	7.1
c) Specification equations applied to observed 700 mb heights:			
Mean absolute error	17.9	21.1	27.3
Bias (forecast–observed)	–1.6	–3.3	–3.0
Correlation coefficient	0.63	0.52	0.29
Skill score (2 classes)	45.6	36.2	21.6
Skill score (5 classes)	28.4	23.1	12.3

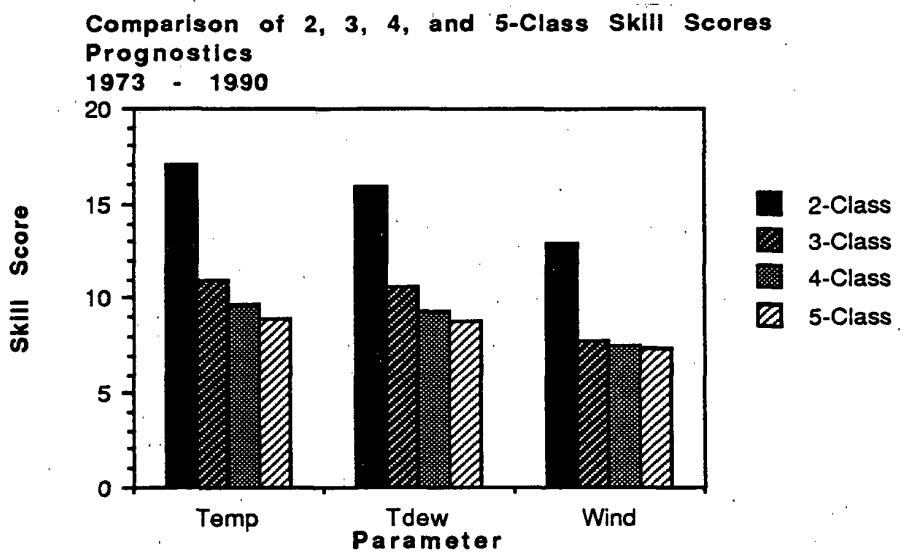


Fig. 1 Heidke skill scores for 2, 3, 4, and 5 classes for specifications of midday monthly mean temperature, dewpoint and wind speed from prognostic 700 mb heights. Data were averaged for 122 stations, two forecasts per month and 18-year period.

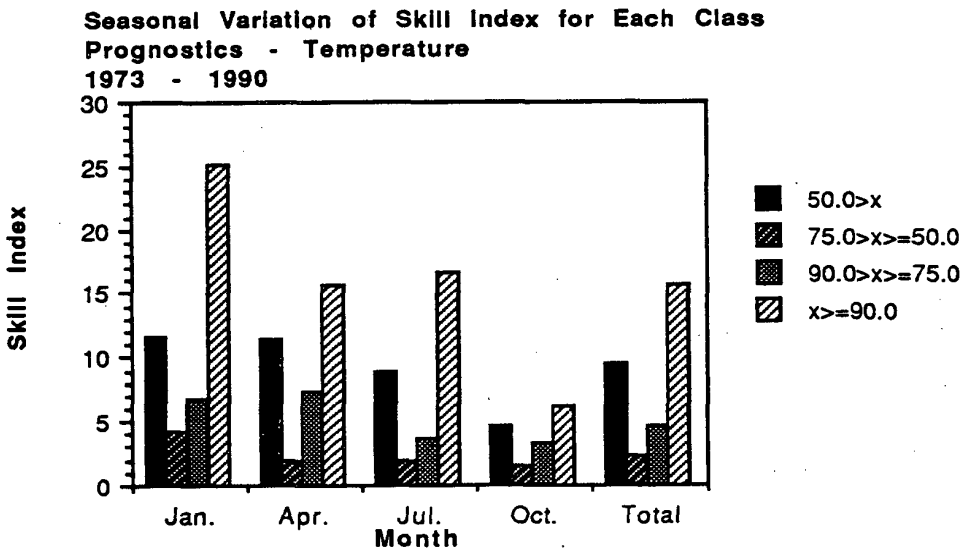


Fig. 2 Same as Fig. 1 but for skill scores computed separately for each of four classes defined (in percentiles) on the right. Scores are given for temperature forecasts during four individual months and the year as a whole (Total).

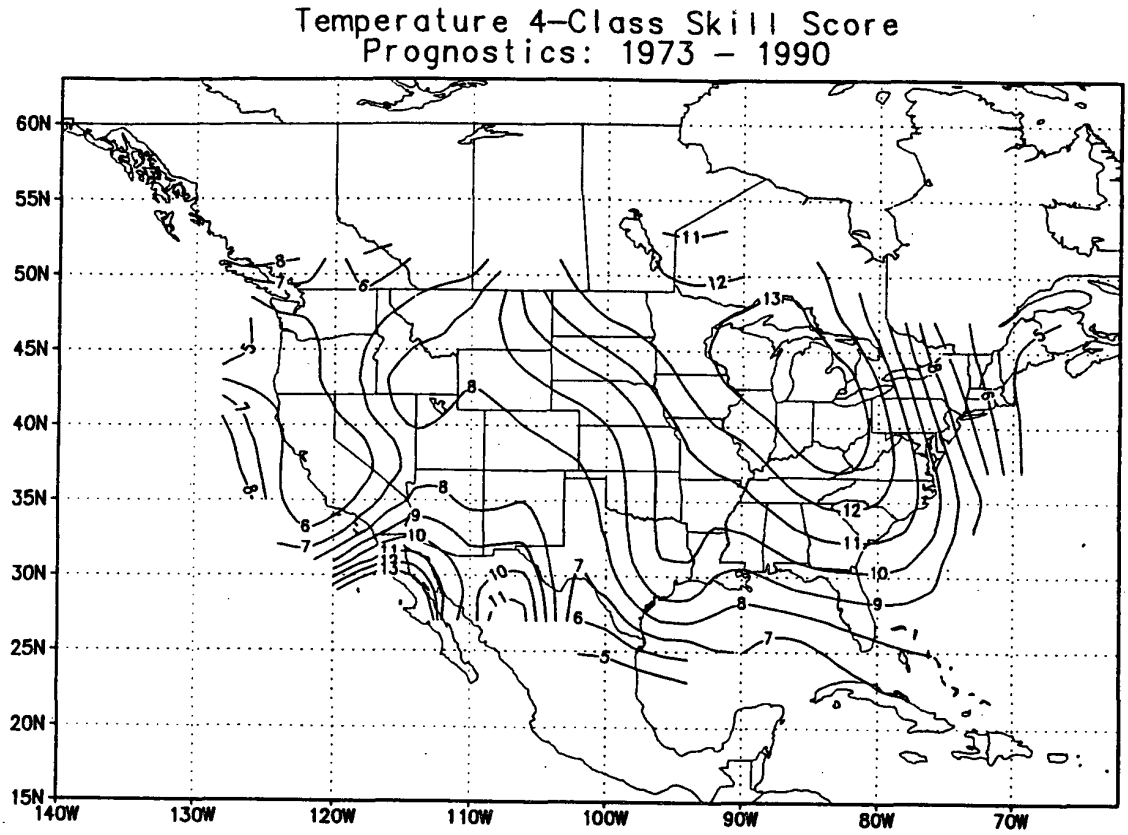


Fig. 3 Geographical distribution of mean annual skill scores in four classes for temperature forecasts based on prognostic heights from 1973 to 1990.

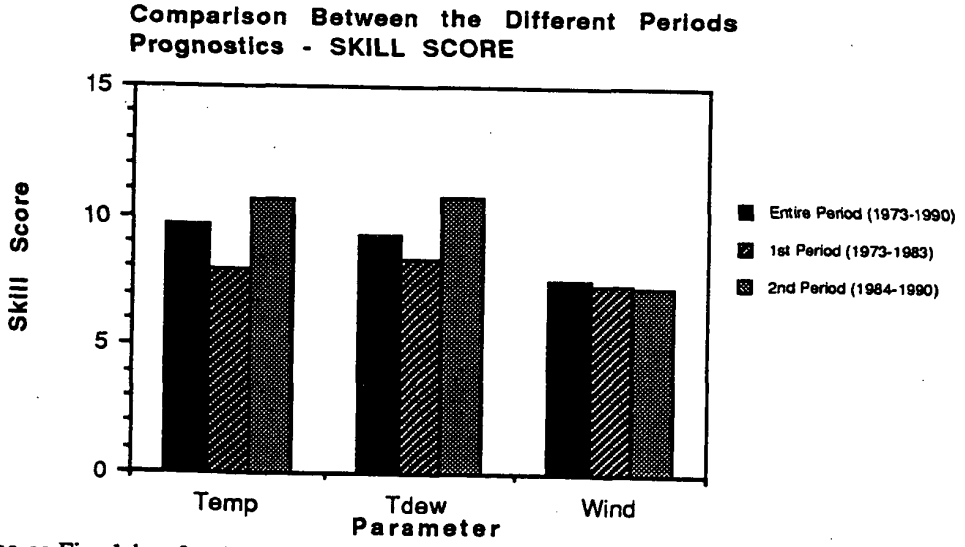


Fig. 4 Same as Fig. 1 but for 4-class skill scores computed for the entire period and two sub-periods defined on the right.

AN INDEX OF THE TROPICAL 30- TO 60-DAY OSCILLATION AND ITS PREDICTION

Hans von Storch, Max-Planck-Institut für Meteorologie
Hamburg, Germany

The *Principal Oscillation Pattern* (POP) analysis of daily equatorial velocity potential at 200 hPa yields a bivariate index of the tropical 30- to 60-day oscillation. The index' potential to monitor and to predict the oscillation are described by [1,2,3] - the following is a brief summary of some of the highlights of these papers.

The result of the POP analysis are pairs of patterns $\underline{e}_1$ and $\underline{e}_2$ whose coefficients $z_1(t)$ and $z_2(t)$ vary coherently but 90° out of phase. The analysis of equatorial velocity potential yields one significant pair of patterns (Fig. 1). The two patterns have a spatial scale of wave number 1 and are about 90° out of phase. The coefficients, for a five year period from 1984 through 1989, vary coherently on the 30 to 60 day band (Fig. 2; cross spectral analysis not shown). z_1 leads z_2 so that the reconstructed signal, $z_1(t)\underline{e}_1+z_2(t)\underline{e}_2$, propagates smoothly around the globe in eastward direction.

The coefficients z_1 and z_2 form a bivariate index (z_1,z_2) of the 30-60 day oscillation. This index can be used to determine the connection between the phase and the amplitude of the oscillation and other atmospheric or oceanic processes. An example is the frequency of tropical storms in the West Pacific (Fig. 3). The 2-dimensional plane is cut in eight 45° degree sectors labelled A to H. Each tropical cyclone is attributed to that sector in which the (z_1,z_2)-index of the day of the storm's first appearance is located. There is a tendency of tropical storms (not) to form in an area where the minimum (maximum) of the upper air velocity potential field is atop or has just passedby. This effect is not strong but statistically significant (risk $\leq$ 0.5 % in the SW Pacific and $\leq$ 1% in the NW Pacific).

The POP analysis yields also a predictive scheme for the index which is sometimes very powerful. An example is the 30. January 1985 when the POP prediction skillfully forecasted the phase of the oscillation for more than 20 days (Fig. 4). For the same initial day an experimental long-range forecast was made with the NCAR forecast model; this dynamical model had a comparable skill as our POP scheme. The overall forecast skill is monitored by means of the correlation skill score (Fig. 5), and compared to that of persistence. The skill scores were derived from 5 years of daily data; about 60% of the data were independent. The POP forecast passes the 50% line after about 9 days

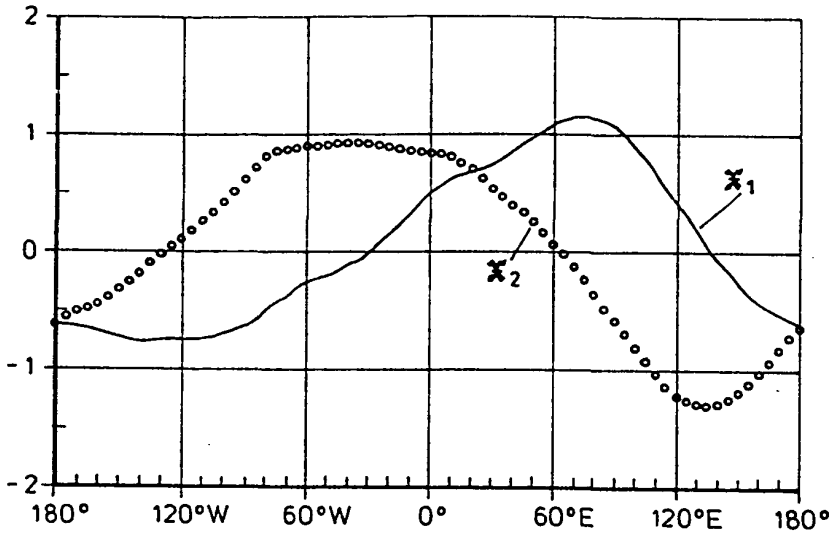
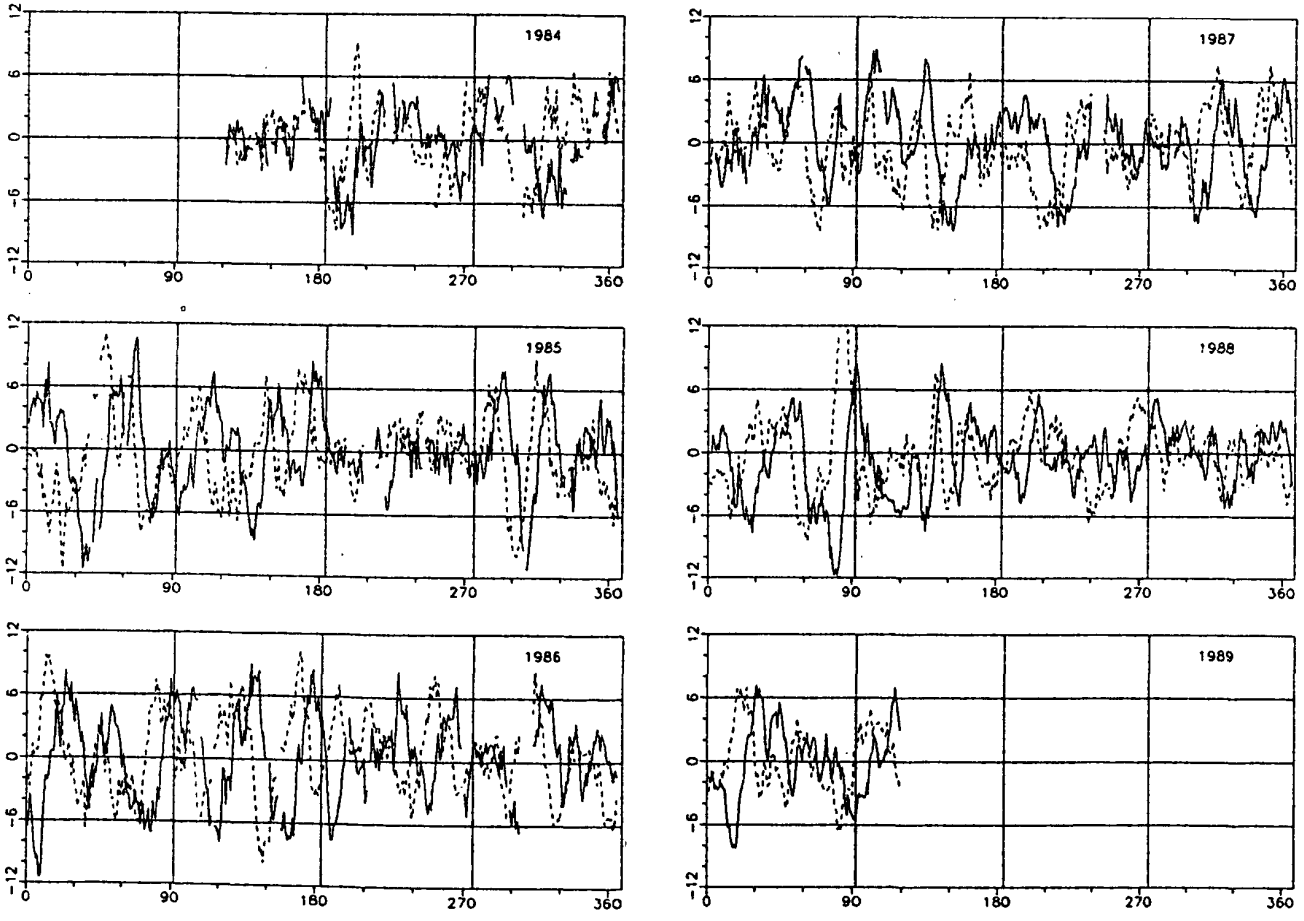


Figure 1
The POP patterns $\mathcal{E}_1$ and $\mathcal{E}_2$ that go with the coefficients in Fig. 2. [1]

Figure 2
The time series of the POP coefficients z_1 and z_2 from 1984 through 1989. [1]



whereas the persistence yields useless results already after about 4 days. If only those case are considered when the initial signal of the 30- to 60-day oscillation is strong, the power of the POP forecast increases substantially (in Fig. 5 the number μ_ℓ represents the mean length of the bivariate vector (z_1,z_2).

A first check of operational weather forecast models indicated that these should be improved with respect to the 30- to 60-day oscillation (not shown).

[1] Storch, H. von; J.S. Xu, 1990: Principal Oscillation Pattern Analysis of the Tropical 30- to 60-day Oscillation: Part I: Definition of an Index and its Prediction. - Climate Dyn. 4, 175-190

[2] Storch, H. von; D. Baumhefner, 1991: Principal Oscillation Pattern analysis of the tropical 30- to 60-day oscillation. Part II: The prediction of equatorial velocity potential and its skill. - Climate Dynamics 5, 1-12

[3] Storch, H. von; A. Smallegange, 1991: The phase of the 30-60 day oscillation and the genesis of tropical cyclones in the western Pacific. - Max-Planck-Institut für Meteorologie Report 66 (Bundesstrasse 55, 2000 Hamburg 13, Germany)

Figure 3

Frequency distribution of the formation of tropical cyclones in two different areas conditional to the phase of the 30-60 day oscillation [3].

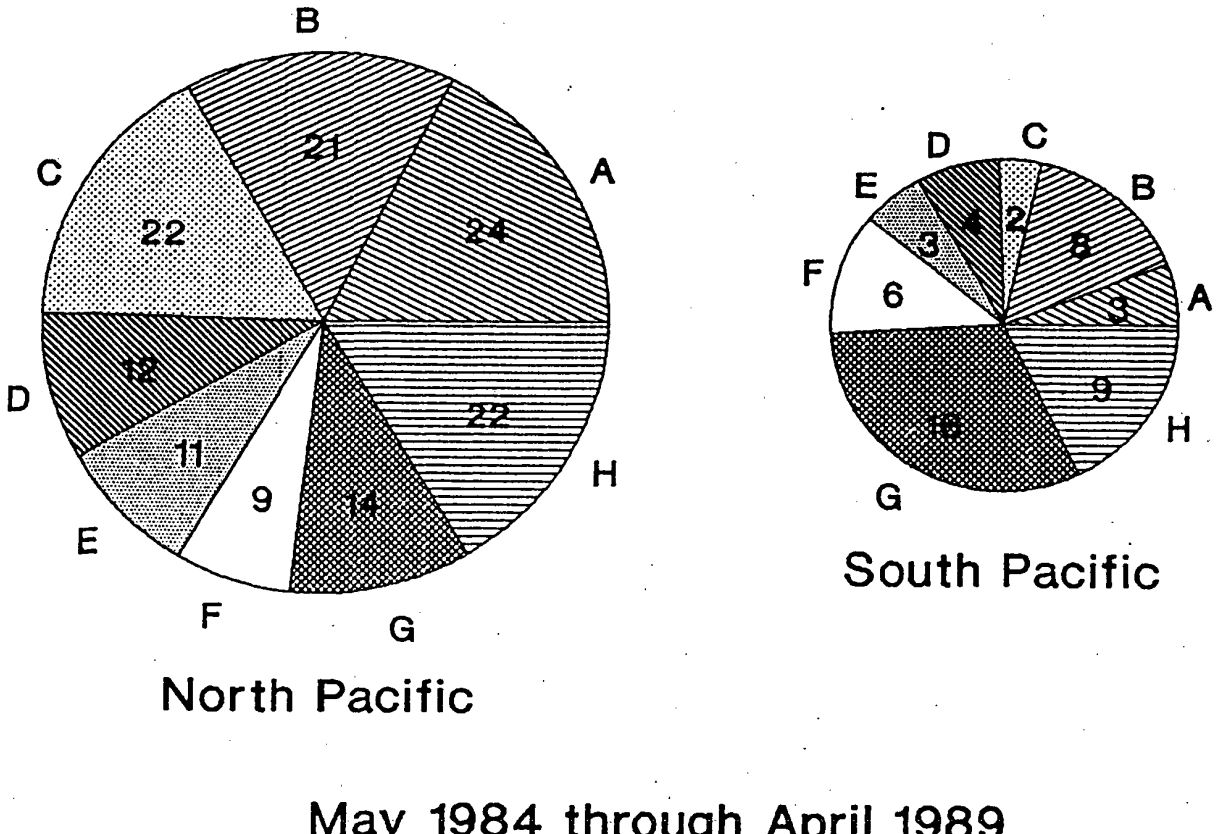


Figure 4

Forecast of the POP index of the 30- to 60-day oscillation from January 30, 1985. For comparison a forecast prepared by the NCAR extended-range dynamical forecast is shown. [2]

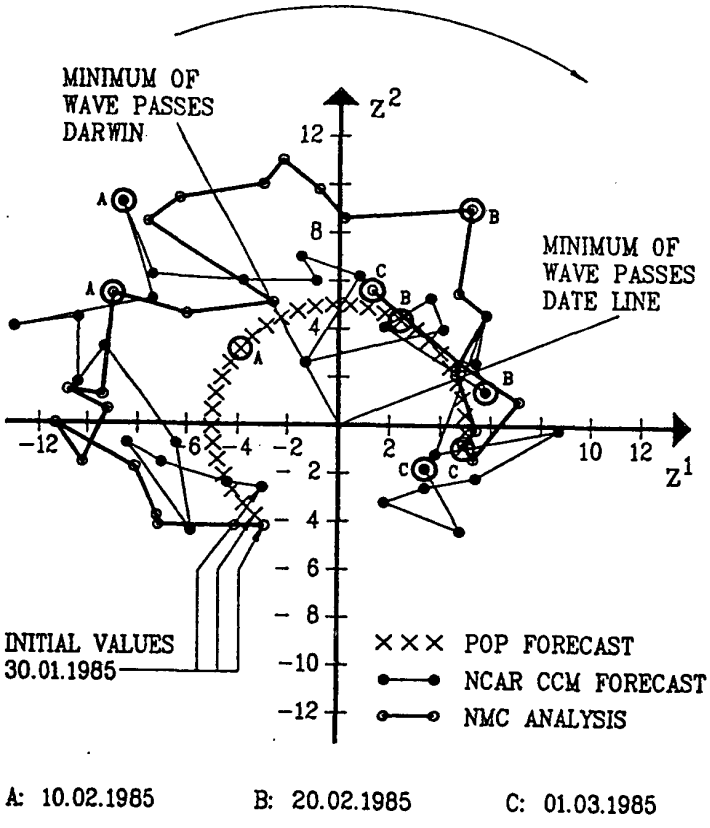
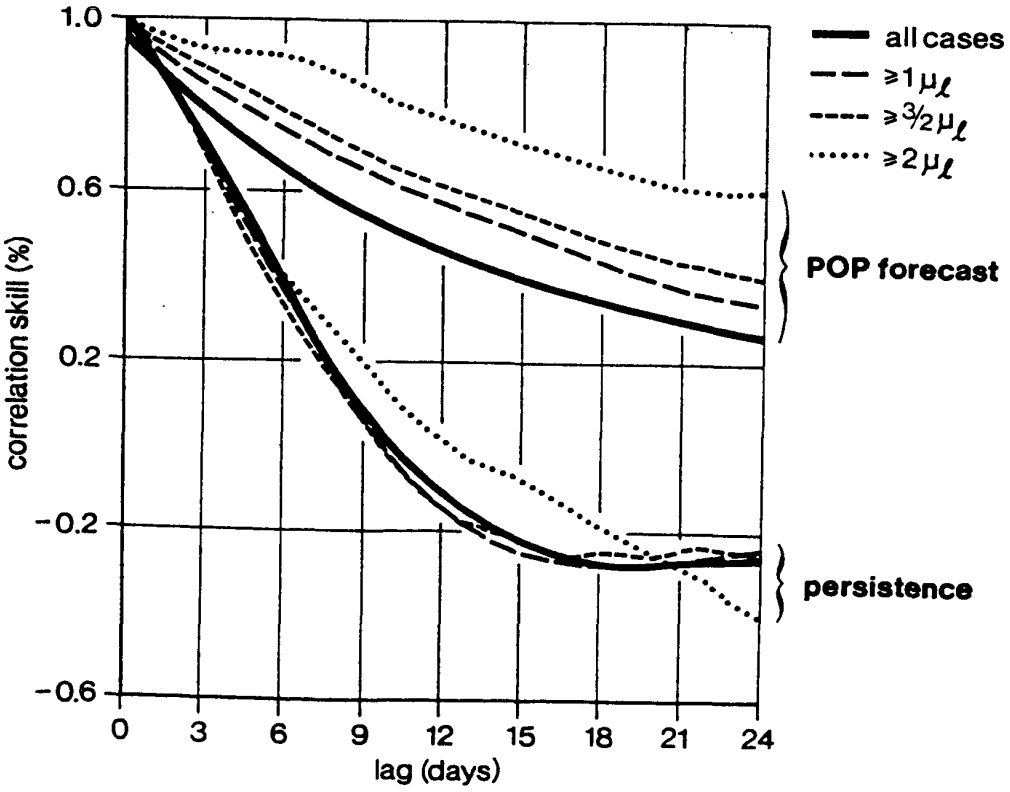


Figure 5

Correlation skill score of the POP forecast of the POP index stratified after the relative strength of the initial signal. μ_ℓ is the average length of the POP index. The performance of persistence is added for comparison. [1]



Estimating forecast statistics from dynamic extended-range forecast ensembles.

Stan Woronko and Peter H. Lee,
Extended-Range Forecast Research Division,
Canadian Climate Centre

Monte-Carlo ensembles of numerical forecasts at medium and extended ranges can, in principle, be used to estimate moments and distributions of forecast variables. In practice, small ensemble sizes and non-representative sampling make derived forecast statistics vulnerable to outliers. In this study, a set of 166 monthly integrations of a T20 version of the Canadian Climate Centre general circulation model is used to study alternative approaches for estimating forecast statistics from ensembles. The model was initialized with global analyses for time-lagged initial conditions for 31 different months from 1979 to 1986, with ensemble sizes ranging from 2 to 6.

Resistant and robust measures of centrality and spread were applied to the distributions of systematic errors and forecasts, to reduce the sensitivity to outliers and departures from normal distributions. The results of these measures are compared to those obtained using conventional estimators and to some obtained with forecast trajectory clustering algorithms.

POTENTIAL PREDICTABILITY ASSESSMENT FOR ANALOG LONG-RANGE PREDICTION SYSTEM

BY G.V. GRUZA AND E.YA. RAN'KOVA

Institute for Global Climate and Ecology, U.S.S.R.

A predictability is defined as skill scores forecasts, based on the historical analog, chosen in archive to the predictand.

The choice of the preferable forecast quality scores and also of the metrics for analogy degree assessment is substantiated.

A predictability is analyzed depending on a target region, season, member of analogs used, etc.

A prior statistical scores are considered relative to a posterior forecast skill and a length of the meteorological time series used. A priori quality scores value and usefulness are discussed.

Empirical estimates are presented for monthly and season air temperatures predictability over Northern Hemisphere and also U.S.S.R. and U.S.A. region.

An Examination of Potential Predictability of Precipitation in Canada

by

Amir Shabbar
Canadian Climate Centre
4905 Dufferin Street, Downsview
Ontario, Canada M3H-5T4

In order to assess the climatic predictability of monthly precipitation, the magnitude of the natural variability is estimated. Following Madden (1976), we define natural variability as that variability which results from the variance and autocorrelation associated with daily weather fluctuations and would be present even within an unchanging climate. Its magnitude is closely dependent on the persistence of daily weather fluctuations. The standard error of time-averaged precipitation resulting from daily fluctuations is only predictable at lead times less than dynamic predictability limits and therefore can be referred to as "climatic noise." We also calculate interannual variability that may exist in excess of the climate noise. We presume that it is at least potentially predictable at long lead times.

Estimates of natural variability are made by first computing spectra of station daily precipitation for time scales of a season. These seasonal spectra are the average of the individual spectra of each of the 40 years from 1946 to 1985. Near-zero frequency spectral estimates are approximated by a low frequency white noise extension of these seasonal spectra as suggested by Leith (1975).

Data

Since precipitation amounts are positively skewed and not normally distributed, we first transformed daily precipitation by taking cube-roots (Freeman and Tukey, 1950) in an attempt to find a better form of precipitation distribution. Daily data from 12 Canadian stations for the period 1946-85 are studied. The data were standardized at each station with

$$x(i,j) = \frac{p(i,j) - \overline{P(j)}}{\overline{\sigma(j)}} \qquad (1)$$

where $x(i,j)$ = standardized precipitation in year i on day j
$p(i,j)$ = daily precipitation in year i on day j.
$\overline{P(j)}$ = normal precipitation (smoothed) and
$\overline{\sigma(j)}$ = normal standard deviation (smoothed).

The normal daily values $\overline{P(j)}$, $\overline{\sigma(j)}$ were smoothed by applying a centered 5 point (5 day)

binomial filter to the average precipitation for the years 1946-85. The resulting standardized data have a mean of zero and a standard deviation of one.

Natural Variability

The variance of the monthly averages σ_T^2, where T represents the monthly averaging time, is given by

$$\sigma_T^2 = \int_{-\infty}^{\infty} S(f) H^2(f)_T df \tag{2}$$

where S(f) is the input seasonal power spectral density function and $H(f)_T$ is the monthly averaging transfer function

$$H^2(f)_T = \frac{\sin^2(2\pi fT/2)}{(2\pi fT/2)^2} \tag{3}$$

The standardized data were separated into individual series for each of the four seasons with each series consisting of 96 days. The mean of each seasonal series was subtracted and using 10% of the data, the series was cosine tapered at both ends. A Fourier transform was then used to obtain 48 harmonic coefficients. A seasonal spectrum for each station was obtained by averaging each of 40 spectra for individual years. All spectra S(f) were normalized so that

$$\int_0^{0.5} S(f) df = \text{total variance} \tag{4}$$

Subtracting seasonal means removes all low frequency (i.e. frequency less than 1/96) information from the spectra. White noise extension is then used to approximate the low frequency ends of the spectra. That is, spectral estimates at 1/96 cycles day^{-1} are extended to zero frequency. This is similar to modelling the near-zero spectral estimates with a Markov process. The resulting spectra for Vancouver, Winnipeg, Toronto and Chatham (N.B.) are shown in Fig. 1. The actual variance for each month (σ_A^2) is determined as

$$\sigma_A^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left[\bar{x}(i) - \bar{\bar{x}}\right]^2 \tag{5}$$

where $\bar{x}(i)$ is the monthly average for each year and $\bar{\bar{x}}$ is the average of all monthly averages from 1946 to 1985.

Potential Predictability

Potential long-range predictability is defined as any variability that exists in excess of our estimates of natural variability which is obtained by low frequency white noise extension of the spectra. Factors such as change in climate regime owing to the almost intransitive nature of the atmosphere (Lorenz, 1968), correlations at lags of a year or more or changes in external conditions may contribute to the interannual variability and thus give rise to potential predictability.

The ratio of the actual variance to the variance due to natural variability σ_A^2 / σ_T^2 was computed. Potential predictability in monthly precipitation is expected to be the greatest where this ratio is the largest. The ratio σ_A^2 / σ_T^2 follows the F-distribution and the null hypothesis: $\sigma_A^2 = \sigma_T^2$ is submitted to an F-test. Potential long-range predictability is most likely to be found at stations where the ratio is significantly larger than unity. Since 40 individual monthly means were used in calculating σ_A^2, the number of degrees of freedom for σ_A^2 is 40. Following Blackman and Tukey (1958), the number of degrees of freedom for σ_T^2 is evaluated to be approximately 100.

Assuming that at each station the ratio σ_A^2 / σ_T^2 is distributed as $F_{n1,n2}$, where $n_1 = 40$ and $n_2 = 100$ and that the null hypothesis is true, there would then be only about a 5% chance that the ratio would exceed 1.52. Table 1 shows the F ratio for each of the four seasons. Based on the limited number of stations that we analyzed, we tentatively conclude that there is more potential predictability of monthly precipitation in winter than summer. Although predictability seems to be the highest on the prairies, potential predictability is also found over the Great Lakes and the East Coast in winter.

REFERENCES

Blackman, R. B., and J. W. Tukey, 1958: *The Measurement of Power Spectra from the View of Communications Engineering*. New York, Dover, 190 pp.

Freeman, M. F., and J. W. Tukey, 1950: Transformations related to the angular and the square root. *Ann. Math Stat.* **21**, 607-611.

Leith, C. W., 1975: The design of a statistical-dynamical climate model and statistical constraints on the predictability of climate. Appendix 2.2 in *The Physical Basis of Climate and Climate Modelling*, No. 16, Geneva, World Meteorological Organization.

Lorenz, E. N. 1968: Climatic determinism, *Meteor. Monogr.*, **8**, 1-3.

Madden, R. A. 1976: Estimates of the natural variability of time-averaged sea-level pressure. *Mon. Wea. Rev.*, **104**, 942-952.

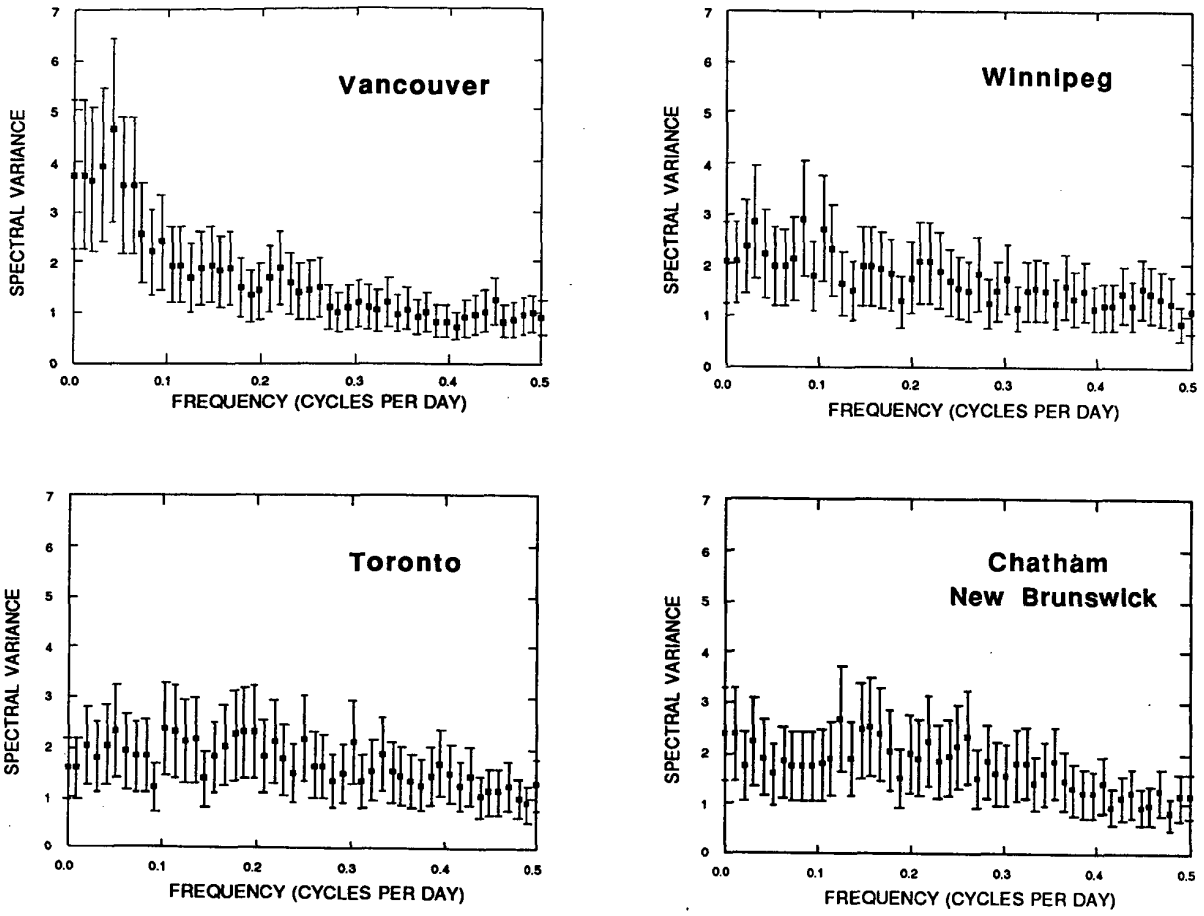


Fig. 1. Winter season spectrum of precipitation. The 99% confidence limit based on 100 dof are indicated by bars.

Table 1. Tabulation of potential predictability (σ_A^2 / σ_T^2) for 12 Canadian stations for each of the four climatological seasons. The ratio is distributed as $F_{n1,n2}$ with $n_1 = 40$ and $n_2 = 100$. A ratio of one indicates no long-range predictability. A ratio of two suggests that one-half of the variance is potentially predictable. Values exceeding 1.52 are underlined. At these stations, the null hypothesis can be rejected with 5.0% chance of error.

Station	Winter	Spring	Summer	Autumn
Whitehorse	1.38	<u>1.54</u>	1.00	1.46
Vancouver	<u>1.78</u>	1.45	1.03	<u>1.55</u>
Edmonton	1.40	1.37	0.80	<u>1.54</u>
Prince Albert	<u>1.89</u>	<u>1.55</u>	1.18	1.09
Winnipeg	<u>2.07</u>	<u>1.54</u>	1.03	1.12
Churchill	1.41	<u>2.31</u>	1.25	<u>1.65</u>
North Bay	<u>1.94</u>	0.93	1.20	1.36
Toronto	<u>1.53</u>	1.04	1.03	1.42
Montreal	1.34	1.00	1.00	<u>1.74</u>
Chatham (N.B.)	<u>1.77</u>	1.47	1.32	1.19
Shearwater	<u>1.74</u>	1.10	1.00	1.33
St. John's	1.34	<u>1.89</u>	<u>1.70</u>	0.98

Skill scores based on mean square error and correlations: Some further relationships and insights

IAN T. JOLLIFFE* and JACQUELINE M. POTTS†
University of Kent at Canterbury, UK.

ABSTRACT

In assessing the accuracy of meteorological forecasts it is often necessary to make comparisons between meteorological fields over a set of grid points. Similar comparisons may also be necessary when comparing current climatological fields with those of the past, in order to detect possible climatic change. The most frequently used measures of similarity are based on either mean square error or correlations. Correlations may take a number of different forms; three are considered in the present paper. Relationships between skill scores based on mean square error and correlations have been investigated by Murphy (1988) and Murphy and Epstein (1989). A further result is given in the present paper and some new insights are obtained.

1. Introduction

There are various contexts in which it is desirable to make comparisons between meteorological fields over a set of grid points. The fields to be compared may be a forecast field and the corresponding observed field, or two members of an ensemble of forecasts. In the former case any measure of similarity between the two fields can be viewed as a skill score for the forecast. Such scores can, however, also be used for other purposes, such as a measure of the extent of climatic change.

The best known skill scores are based on either mean square error or correlation, with correlation defined in a number of ways. Three such definitions, together with that of mean square error, are given in Section 2. Some relationships between mean square error and correlation are already well known; Section 3 gives two of these and adds a third.

In Section 4, it is shown that the mean square error may be artificially reduced using an unskilled method of modifying the forecast, which takes a linear combination of the original forecast and climatology. Reductions in mean square error may also be obtained by taking a linear combination of the forecast and its spatial mean, or by doing the same to the forecast anomaly.

There has been some discussion in the literature regarding the merits or demerits of skill scores based on mean square error or correlation. This discussion is taken further in Section 5 in the light of our new results.

2. Definitions of mean square error and correlation

Suppose that x_i, y_i, $i = 1, 2, \cdots, n$ denote the observed and forecast values respectively of a variable which is recorded at each of a set of grid points. Denote the climatological average at the ith grid point by c_i and let

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i, \quad \bar{c} = \frac{1}{n}\sum_{i=1}^{n} c_i.$$

The mean square error (MSE) is

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2$$

and the standard form for the correlation is

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2\right]^{1/2}}.$$

Because each grid point has a different climatological average c_i, the coefficient r_{xy} has the undesirable property that its expected value is not in general zero when forecast and observed fields are independent. To overcome this problem, the correlation can be calculated between the anomalies $x_i - c_i$ and $y_i - c_i$ instead of between x_i and y_i, giving r'_{xy}. Either form of correlation is invariant under linear transformations and so is unable to detect systematic errors due to the addition of the same constant at each grid point, or the multiplication of the variable (for r_{xy}), or of the anomaly (for r'_{xy}), by the same scale factor at each grid point.

A third form of correlation is obtained by taking deviations of $(x_i - c_i)$, $(y_i - c_i)$ about

* Present address: Department of Mathematical Sciences, University of Aberdeen, Aberdeen AB9 2TY, UK.

† Present address: Department of Statistics, Rothamsted Experimental Station, Harpenden, Herts AL5 2JQ, UK.

their expected value, zero, rather than their sample averages $(\bar{x} - \bar{c})$, $(\bar{y} - \bar{c})$, giving

$$r''_{xy} = \frac{\sum_{i=1}^{n}(x_i - c_i)(y_i - c_i)}{\left[\sum_{i=1}^{n}(x_i - c_i)^2 \sum_{i=1}^{n}(y_i - c_i)^2\right]^{1/2}}.$$

Like r'_{xy} this correlation has an expectation of zero for independent patterns, but it has the advantage that it is sensitive to the addition of the same constant at each grid point. However, it is still unable to detect when the anomaly at each point is multiplied by the same scale factor. Yet other forms of correlation have been suggested, but we restrict attention to these three.

3. Relationships between mean square error and correlations

Murphy (1988) demonstrated the following relationship between MSE and r_{xy}:

$$\text{MSE} = (\bar{x} - \bar{y})^2 + s_x^2 + s_y^2 - 2 s_x s_y r_{xy} \quad (1)$$

where $s_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ and $s_y^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2$. By similar reasoning, using anomalies instead of raw data, Murphy and Epstein (1989) obtained the following result:

$$\text{MSE} = (\bar{x} - \bar{y})^2 + s'^2_x + s'^2_y - 2 s'_x s'_y r'_{xy} \quad (2)$$

where $s'^2_x = \frac{1}{n}\sum_{i=1}^{n}(x_i - c_i - \bar{x} + \bar{c})^2$ and $s'^2_y = \frac{1}{n}\sum_{i=1}^{n}(y_i - c_i - \bar{y} + \bar{c})^2$. Similarly, we can obtain a third result:

$$\text{MSE} = s''^2_x + s''^2_y - 2 s''_x s''_y r''_{xy} \quad (3)$$

where $s''^2_x = \frac{1}{n}\sum_{i=1}^{n}(x_i - c_i)^2$ and $s''^2_y = \frac{1}{n}\sum_{i=1}^{n}(y_i - c_i)^2$. Thus there are equations linking the MSE and all three forms of the correlation coefficient defined in Section 2.

4. Unskilled reduction of mean square error

Using the relationship of the previous section, it is possible to show that an unskilled strategy can be used to reduce the expected value of the MSE, and hence apparently increase skill. Consider a system which produces forecasts $f_1, f_2, \cdots, f_n$ and assume that, from past experience, expected values of the correlation and anomaly correlation coefficients between the observed and forecast fields are known. Now construct modified forecasts of the form

$$y_i = a f_i + (1 - a)\bar{f}, \quad (4)$$

where a is some constant.
Then $\bar{y} = \bar{f}$, $s_y^2 = a^2 s_f^2$ and $r_{xy} = r_{xf}$. So from (1) the MSE is given by

$$\text{MSE} = (\bar{x} - \bar{f})^2 + s_x^2 + a^2 s_f^2 - 2 a r_{xf} s_x s_f$$

which is minimised by setting

$$a = r_{xf}\frac{s_x}{s_f}.$$

Next, consider a forecast of the form

$$y_i = b f_i + (1 - b)(c_i + \bar{f} - \bar{c}) \quad (5)$$

or equivalently,

$$y_i - c_i = b(f_i - c_i) + (1 - b)(\bar{f} - \bar{c}),$$

for some constant, b.
Then $\bar{y} = \bar{f}$, $s'^2_y = b^2 s'^2_f$ and $r'_{xy} = r'_{xf}$. This forecast has

$$\text{MSE} = (\bar{x} - \bar{f})^2 + s'^2_x + b^2 s'^2_f - 2 b r'_{xf} s'_x s'_f$$

and the value of b which minimises the MSE is

$$b = r'_{xf}\frac{s'_x}{s'_f}.$$

Finally, consider a forecast of the form

$$y_i = d f_i + (1 - d) c_i, \quad (6)$$

for some constant, d.
Then $s''^2_y = d^2 s''^2_f$ and $r''_{xy} = r''_{xf}$. This forecast has

$$\text{MSE} = s''^2_x + d^2 s''^2_f - 2 d r''_{xf} s''_x s''_f$$

which is minimised when

$$d = r''_{xf}\frac{s''_x}{s''_f}.$$

Thus all three strategies lead to some reduction in the expected MSE. If s_x and s_f, s'_x and s'_f, and s''_x and s''_f are of similar magnitude then a, b and d are approximately equal to the three forms of correlation coefficient r_{xf}, r'_{xf} and r''_{xf} respectively. So assuming that the mean values of these coefficients are known from past experience, it is possible to improve the apparent skill of a forecast using a strategy which does not require any skill.

Of the three modified forecasts (4), (5), (6), it can be shown (Potts, 1991) that (6)

should provide the maximum reduction in MSE, although the resulting modified forecasts may be less realistic than the original ones because $s''_y{}^2$ is less than $s''_f{}^2$.

The phenomenon of reduction of MSE by an unskilled modified forecast also occurs when a persistence forecast is replaced by a linear combination of persistence and climatology (Daan, 1984; Lorenz, 1973). The apparent skill of a forecast, as measured by r_{xf}, can also be increased by a similar strategy, but r'_{xf} and r''_{xf} are unaffected by such modifications to the forecasts.

5. Decompositions of skill scores based on MSE

Skill scores are often constructed by taking some measure of the accuracy of the forecast and comparing it with that of a corresponding reference forecast. For MSE, with climatology as the reference forecast, a skill score is

$$\mathrm{SS} = 1 - \frac{\mathrm{MSE}(x,y)}{\mathrm{MSE}(x,c)},$$

where $\mathrm{MSE}(x,y)$, $\mathrm{MSE}(x,c)$ are MSEs between the forecasts and the observations, and between climatic averages and the observations respectively. Using decomposition (1), Murphy (1988) showed that SS involves terms $r_{xy}{}^2$,

$\left[r_{xy} - \frac{s_y}{s_x}\right]^2$ and $\frac{(\bar{y}-\bar{x})^2}{s_x{}^2}$, which are measures, respectively, of the strength of linear relationship between forecasts and observations, conditional bias in the forecast, and unconditional bias. Murphy and Epstein (1989) carry out a similar exercise using decomposition (2) and find three terms with similar interpretations,

namely $r'_{xy}{}^2$, $\left[r'_{xy} - \frac{s'_y}{s'_x}\right]^2$ and $\frac{(\bar{x}-\bar{y})^2}{s'_x{}^2}$.

In addition, there is a term $\frac{(\bar{x}-\bar{c})^2}{s'_x{}^2}$, which is the squared coefficient of variation of the anomalies.

A similar, but somewhat simpler, decomposition can be made in terms of r''_{xy}, namely

$$\mathrm{SS} = r''_{xy}{}^2 - \left[r''_{xy} - \frac{s''_y}{s''_x}\right]^2.$$

Murphy and Epstein (1989) argue that $r'_{xy}{}^2$ can be viewed as a measure of the *potential* skill which would be achieved if conditional and unconditional bias could be eliminated. However, the apparent advantage of SS over correlation coefficients in allowing for bias is somewhat diluted when it is realised that adopting the unskilled strategies of (4), (5) or (6) eliminates the conditional bias term. In particular, if forecast (6) is used with d chosen to minimise MSE, then SS reduces to $r''_{xy}{}^2$.

Finally, it is of interest to re-express (1), (2) and (3) by dividing through by $s_x s_y$, $s'_x s'_y$ and $s''_x s''_y$ respectively. For example, (1) can be written as

$$\frac{\mathrm{MSE}}{s_x s_y} = \frac{(\bar{x}-\bar{y})^2}{s_x s_y} + \frac{(s_x - s_y)^2}{s_x s_y} + 2(1 - r_{xy}),$$

where the three terms are measures of the difference in the sample mean, the difference in spread about the sample mean and the difference in pattern as measured by the correlation coefficient respectively. The three terms in this decomposition are similar to the trinity statistics used by Preisendorfer and Mobley (1982) to compare time series of data over a grid of points except that averaging takes place over space rather than time and there is only one time interval involved. By a similar argument based on (2), we obtain

$$\frac{\mathrm{MSE}}{s'_x s'_y} = \frac{(\bar{x}-\bar{y})^2}{s'_x s'_y} + \frac{(s'_x - s'_y)^2}{s'_x s'_y} + 2(1 - r'_{xy}).$$

Here the first term relates to the difference in the sample mean; the second term relates to the difference in the spread of anomalies about the sample mean anomaly and the third term relates to the difference in pattern as measured by r'_{xy}. Alternatively, using (3), we obtain

$$\frac{\mathrm{MSE}}{s''_x s''_y} = \frac{(s''_x - s''_y)^2}{s''_x s''_y} + 2(1 - r''_{xy}).$$

The first term is a measure of the difference in the magnitudes of the departure from the climatological average while the second corresponds to the difference in pattern as measured by r''_{xy}.

As an example of the meaning of this kind of decomposition in practice, we consider the pairs of patterns shown in Figures 1 and 2. These represent half monthly mean sea-level pressure (MSLP) for an area around the UK. If a comparison is made between the two patterns in Figure 1, then the term corresponding to the difference in spread about the sample mean is large although the difference in the sample mean is quite small and the correlation between

the patterns is high. In Figure 2, however, the correlation is low as the patterns are almost mirror images of one another. The sample means of these two patterns are very similar and the spread of the raw data about the sample mean is also fairly similar but there is a large difference in the spread of the anomalies about the sample mean anomaly.

Figure 1(a). *MSLP for the UK region in the second half month of January 1967.*

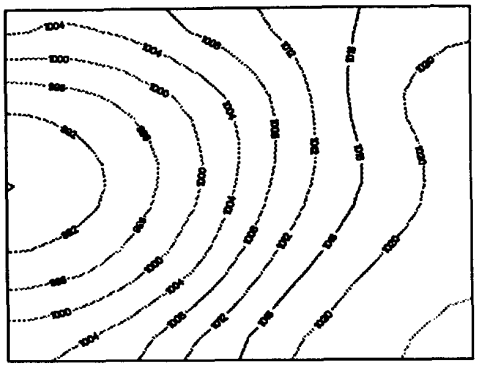


Figure 1(b). *MSLP for the UK region in the second half month of January 1980.*

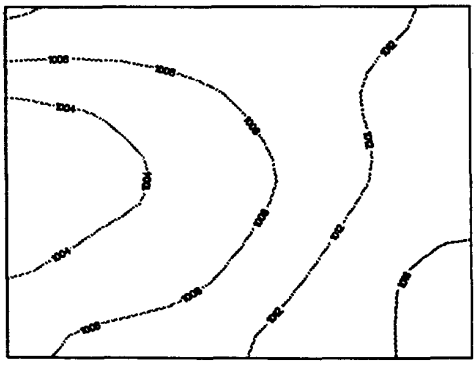


Figure 2(a). *MSLP for the UK region in the first half month of January 1953.*

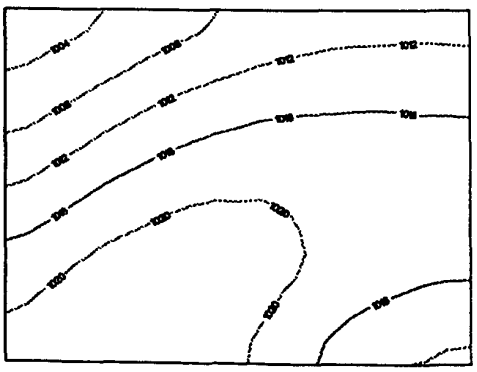
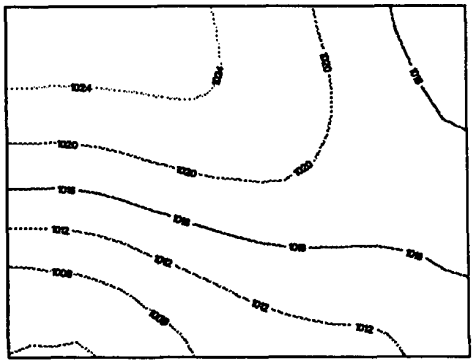


Figure 2(b). *MSLP for the UK region in the first half month of January 1963.*



Acknowledgements

The work of J.M. Potts was supported by a CASE studentship from SERC and the Meteorological Office.

References

Daan, H., 1984: *Scoring rules in forecast verification*. World Meteorological Organisation.

Lorenz, E.N., 1973: On the existence of extended range predictability. *J. Appl. Meteorology*, **12**, 543-547.

Murphy, A.H., 1988: Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev.*, **116**, 2417-2424.

————, and Epstein, E.S., 1989: Skill scores and correlation coefficients in model verification. *Mon. Wea. Rev.*, **117**, 572-581.

Potts, J.M., 1991: *Statistical methods for the comparison of spatial patterns in meteorological variables*. Unpublished Ph.D. thesis, University of Kent at Canterbury.

Preisendorfer, R. and Mobley, C., 1982: *Data intercomparison theory II. Trinity statistics for location, spread and pattern differences*. NOAA Technical Memorandum ERL PMEL-39, Pacific Marine Environmental Laboratory, Seattle, Washington.

On the Prospects for Long-Range Prediction of Temperature over the Globe

Roland A. Madden and Wilhelm May
National Center for Atmospheric Research*
Boulder, Colorado 80307

1. Introduction

We must expect a certain amount of interannual variability in time-averaged meteorological data due to statistical sampling variations. That is, one monthly or seasonal 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 unpredictabale 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 means could reflect the effects of slowly changing external conditions such as sea surface temperature or it could reflect intransitivity. 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 at stations over the globe and then compare them with actual variance of time-means. The more the actual variance exceeds the noise variance the greater are the prospects for long-range predictability.

The comparisons are handled as an analysis of variance problem. The variance within seasons is assumed to be very nearly entirely the result of the weather systems whose details are unpredictable at long range. The null hypothesis is that values from each season come from a population with the same mean and that there are no prospects for long-range predictability because all variations between seasons are simply sampling variations of unpredictable weather, or climate noise. If the variance between seasons exceeds levels that the variance within seasons can explain, then we reject the null hypothesis and conclude that there is, at least, a potential for long-range predictability. Experience has shown that the spectra of daily data, upon which our test for stationarity in the means is based, can vary from one period to another. That is, the climate noise is not necessarily stationary. For this reason, even though our comparisons are framed as a rigorous statistical test, we choose to interpret the results a little more qualitatively. We conclude only that the likelihood of potential long-range predictability is roughly proportional to the ratio of actual variance to climate noise variance.

2. Estimating Climate Noise

To estimate the climate noise we used daily temperature $((T_{max} + T_{min})/2)$

* The National Center for Atmospheric Research is sponsored by the National Science Foundation

data from about 3000 stations over the world for the period January 1979 through June 1989. These data were originally prepared by the Climate Analysis Center (NOAA). Further analysis was done at NCAR (May et al., 1992). Spectra, S(f), were estimated for each station and, season by:

1) subtracting the seasonal variation from the time series
2) breaking the time series up into separate, slightly overlapping, 96-day seasonal segments
3) Fourier transforming each seasonal segment and squaring the coefficients
4) averaging the squared coefficients over all available seasonal segments

The time period allows for a maximum of 10 summer, fall, and winter segments, and for 11 spring segments. There were considerable missing data in some locations, but no spectra were based on less than three segments.

Figure 1 is an example of the resulting spectra for DJF and JJA at Toronto. They depict the within-season variance as a function of frequency from 0.5 cycles/day to 1/96 c/d. The spectra give no certain information about variance on time scales exceeding 96 days. If we make the reasonable assumption that the unpredictable (at long range) weather variations are uncorrelated at lags this long, we expect that the unknown low-frequency variance is nearly constant. We refer to this as the low-frequency white noise assumption (LFWN). As a point of comparison, the spectrum of a first order autoregressive process is close to white at periods longer than 96 days for one-day lag correlations less than 0.8. The LFWN has the same effect on the total variance that dividing by N-1 rather than N does when making an unbiased estimate of variance for uncorrelated data (white noise).

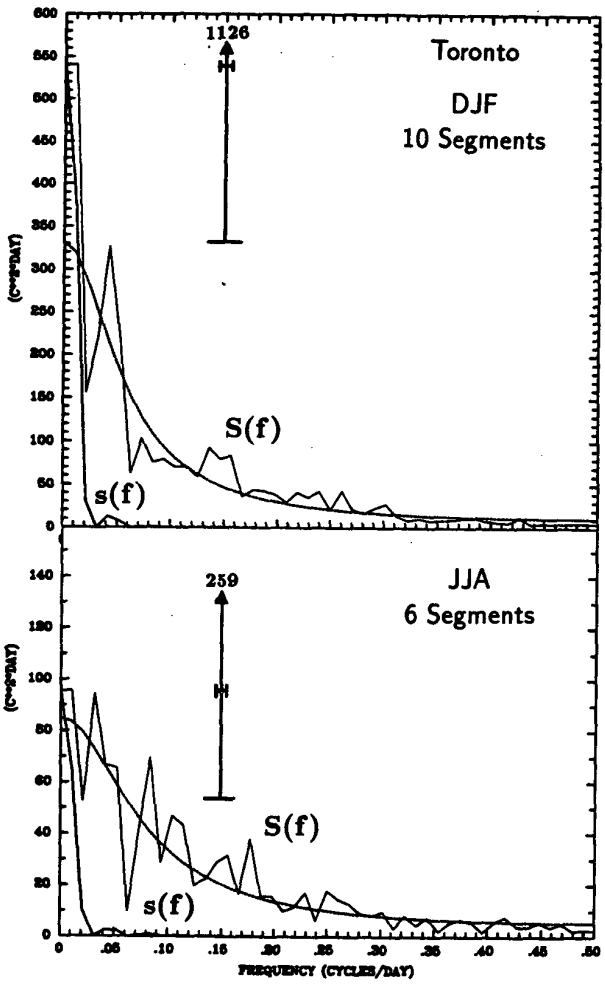


Figure 1. Spectra of daily temperatures at Toronto for summer (bottom) and winter (top) are indicated by the thin line. The dotted lines are the spectra for a first order autoregressive process with a one-day lag correlation of .57 and .68, which are ones computed for summer and winter, respectively. The LFWN is indicated by the horizontal line between 0.0 and 1/96 c/d. There is only a 5% chance that the true spectral value at 1/96 c/d falls outside of the indicated error bars. The upper limit for JJA is 259 and for DJF is 1126. Error bars assume 12 and 20 df. The thick lines represent the expected spectra for 31-day averages determined from (1).

The climate noise for monthly averaged data is then

$$\sigma_N^2 = \int_0^{0.5} s(f)df \qquad (1)$$

where $s(f) = S(f)H_N^2(f)$. The $H_N^2(f)$ is the power transfer function of the monthly, or N-day average. Figure 1 also shows s(f), the spectra of 31-day averages (N=31) determined from the within-season spectrum of daily data. It is clear, but not surprising since monthly averaging is a low-pass filtering operation, that the climate noise comes entirely from the low frequency end of the spectrum of the daily weather variations. For our example, the climate noise of a 31-day average during JJA and DJF at Toronto is given by the areas under the s(f)'s of Fig. 1 which are 1.4 and 7.3°C^2, respectively. Similar estimates of climate noise were made at all stations.

3. A Measure of the Prospects for Long-Range Prediction

To learn where and during which seasons the prospects for long-range prediction are best (and worst), we compute the ratio of actual variance of monthly means to that of the climate noise. The actual variance σ_A^2 of monthly means is that determined from the 30-year period 1950–1979 (Shea 1986). The ratios are presented in Figs. 2 and 3 for January and July. If the null hypothesis were true then the ratios should be distributed as an F-distribution associated with the degrees of freedom (df) in estimating the numerator and the df in estimating the denominator. The numerator provides 29 df. Determining df of the denominator is not so straightforward. They can be estimated from the number of segments used to compute S(f) and from the shape of the s(f) spectra. They vary from station to station but df=20 is a representative value. It is close to the df of one spectral estimate (number of segments × 2 for the real and imaginary components of the Fourier transform). These df suggest that only 5% of the estimates of σ_A^2/σ_N^2 would exceed 2.0 if $\sigma_A^2 = \sigma_N^2$. Similarily, only 5% of ratios should be less than about 0.5. If in fact $\sigma_A^2 < \sigma_N^2$ at a location then the true low frequency spectrum is at a level lower than the LFWN.

There is a tendency for relatively large ratios to occur near the coasts of the continents. Large ratios are present in polar latitudes during July and in the tropics during July and January. Ratios exceed 3.0 over Peru and northern Chile. This is a result of the fact that the climate noise is only about 0.2–0.3°C^2, while, probably influenced by relatively large, local sea temperature changes, the interannual variance of surface temperature remains near 1.0–2.0°C^2 during the year.

4. Summary

Estimates of climate noise and the potential for long-range predictability of surface temperature have been extended to a data set of stations located over the globe. These estimates are based on relatively short records of daily data. We plan to make careful comparisons between these results and available ones that were determined from longer records. We also plan to make comparisons with studies of slow atmospheric variability.

References

May, W., D. J. Shea, and R. A. Madden, 1992: The Annual Variation of Surface Temperatures over the World. NCAR/TN-372+STR, in press.

Shea, D. J., 1986: Climatological Atlas: 1950–1979, Surface Air Temperature, Precipitation, Sea-Level Pressure, and Sea-Surface Temperature (45°S–90°N). NCAR/TN-269+STR, 193 pp.

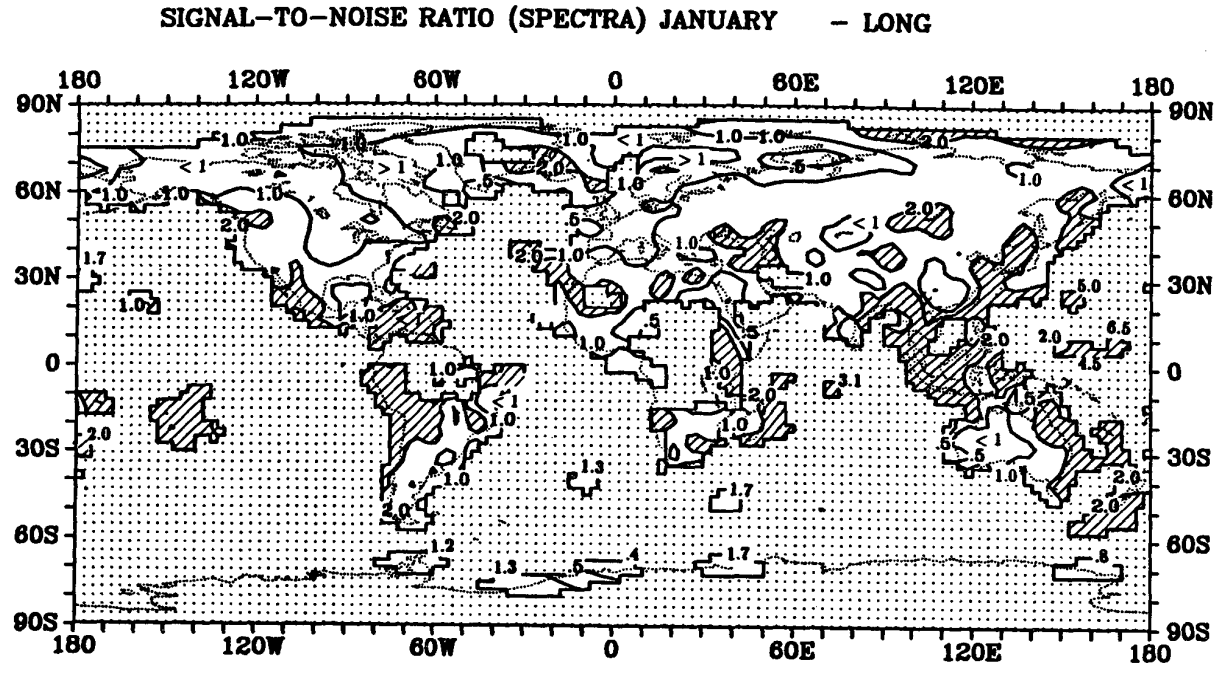


Figure 2. Ratio of σ_A^2/σ_N^2 for January. Contour intervals are 0.5, 1.0, 2.0.

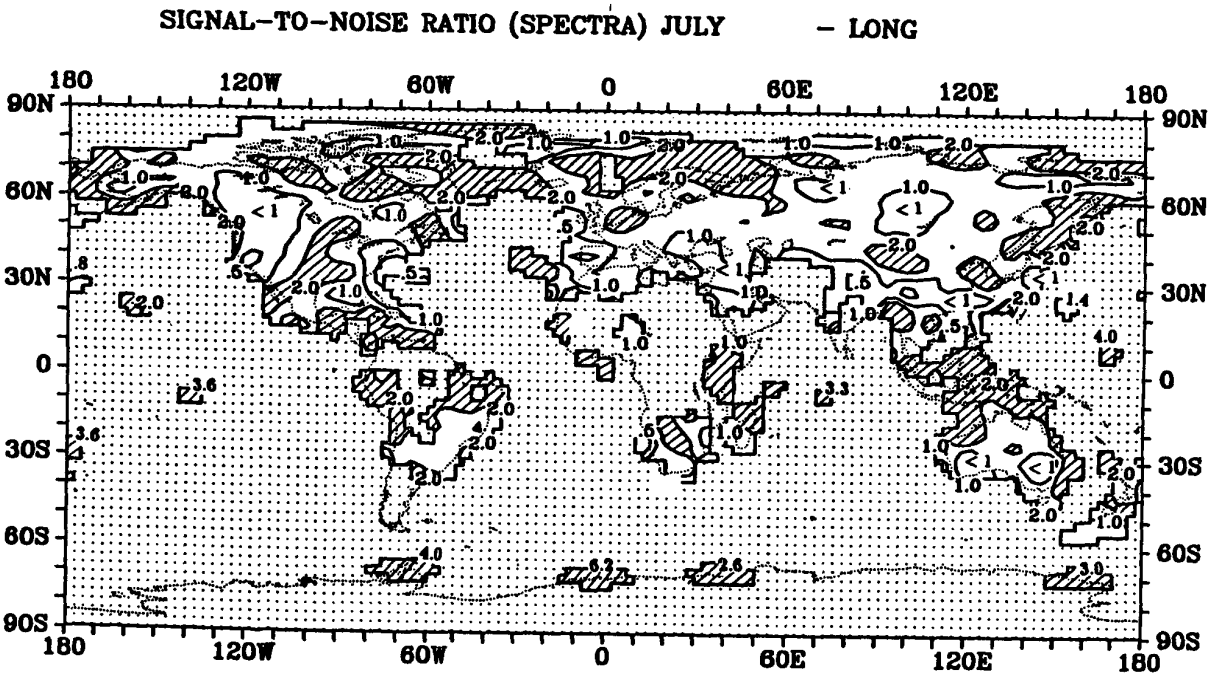


Figure 3. Like Fig. 2 for July.

Reconstructing the Phase Space of the Climate System in Recent 134 Years and Its Complexity and Predictability

Weiyu Yang
Institute of Atmospheric Physics, Academia Sinica
P. O. Box 2718, Beijing, 100080, China.

Abstract

Based on the time series of the northern hemisphere monthly mean surface air temperature in recent 134 years, the complexity and predictability of the climate system are analysed with the technique of reconstructing the phase space and the nonlinear chaos theory. The characters of the reconstructed phase spaces which use different values of the embedding dimensionality m and the time delay parameter τ are analysed. The range of the τ is from 3 to 40, in which just the 12 values, which make the coordinate components of the reconstructed phase spaces be close to linearly independent, are used. The range of the m is from 4 to 120 with the inteval of 2. It is found that the m and the τ can affect the characters of the reconstructed phase spaces.

The correlation dimension and the order–2 Renyi entropy and their temporal variation of the climate system are calculated, and then the average predictable time scale of the system is found. The results show that the climate system is a chaotic system which has limited and fractal dimension, and the complexity of the climate system in recent 134 years is increase gradually. The correlation dimension is 9.6. The saturated embedding dimensionality m_∞ is 36. Although the data number is limited, through detailed analyses, the results may be reliable. The average predictable time scale is less than 9.9 years for monthly average climate system. Such information is very important for designing climate models and predicting climate.

In order to simplify the climate models and lengthen predictable time scale, I treat the original air temperature time series by a filter process and found that, in the phase space which be reconstructed with the new filtered time series, the correlation dimension value is decreased greatly and the predictable time scale is lengthened. It suggests that, if predicting climate in this new phase space, we may get more satisfactory results.

Submitted to the Oral or poster sessions, the Fifth International Meeting on Statistical Climatology, June 22–26, 1992, Toronto, Ont., Canada.

A NEW MULTIDIMENSIONAL TIME SERIES FORECASTING METHOD BASED ON THE EOF INTERATION SCHEME

BANG-LIN ZHANG AND JIE LIU
Institute of Atmospheric Physics, Beijing, China

ZHAO-BO SUN
Nanjing Institute of Meterology,Nanjing,China

ABSTRACT

In this paper a new multidimensional time series forecasting scheme is introduced, which is based on the empirical orthogonal function (EOF) stepwise iteration process. The scheme is tested in a series of forecast experiments of Nino3 SST anomalies and Tahiti-Darwin SO index. The results show that the scheme is feasible and ENSO predictable.

1. Introducton

As a pattern of short-term climatic fluctuations with a characteristic period of 2-7 years, the ENSO is closely associated with the global climatic anomaly. An overall relationship between the ENSO and the clmate has been relatively extensively studied(Chen,1981;Rasmusson and Carpenter 1982; Ropelewski and Halpert, 1987), and many of the ascribed effects are widely accepted as being statisticaly significant, physically plausible, and of sufficent magnitude to be of considerable economic concern in a number of regions. Therefore,the reseaches of the predictability and long-range forecasting technique have been considerable interest and are paid more attention by meteorologists. A lot of papers have been pubilished on the predictability and prediction of the ENSO. For example, Fraedrich(1988) estimated predictability time scales of two SO time series with the conclusion that predictions are possible for at least one to two years. Inoue and O'Brien(1984) demonstrated the feasibility of a forecasting scheme for predicting the onset of a major ENSO, which is based on simple physics and statistics. Using a purely statistical technique – the linear prediction theory, Barnett (1981) was also able to show that SST anomalies off Peru are predictable one year in advance for the period 1976-79. Recently, Xu and Storch (1990) introduced a forecasting scheme based on principal oscillation pattern(POP) analysis. A series of hindcast experiments show that the POP prediction scheme is skillful for a lead time of two to three seasons. Thus, the ENSO is predictable.

The purpose of this paper is to introduce a new statistical forecasting scheme based on empirical orthgonal function (EOF) iteration process and

illustrate its use by applying the scheme to some of the ENSO indices.

A description of the scheme based on EOF iteration process is given in section 2. This is followed by an application of the scheme to the forecasts of ENSO indices in section 3. The paper is concluded with discussion in section 4.

2. Time series forecast scheme based on EOF iteration process

Given a discretely-sampled multi-dimensional time series X_i of length $1 \le i \le N$, the scheme based on EOF iteration process for predicting X_{N+1} at time N+1 is as follows: at first we construct an initial augumental matrix of the form

$$F^{(0)}_{(N-M+2)\times LM} = \begin{bmatrix} X_1 & X_2 & \cdots & X_{M-1} & X_M \\ X_2 & X_3 & \cdots & X_M & X_{M+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ X_{N-M+1} & X_{N-M+2} & \cdots & X_{N-1} & X_N \\ X_{N-M+2} & X_{N-M+3} & \cdots & X_N & X^{(0)}_{N+1} \end{bmatrix}$$

where LM is dimension of embedding space, $X^{(0)}_{N+1}$ denotes the initial L-dimensional vector. EOF-expanding of the matrix $F^{(0)}$ yields its fitting field $\hat{F}^{(0)}$ in the form

$$\hat{F}^{(0)}_{(N-M+2)\times LM} = \sum_{k=1}^{K^{(0)}} T^{(0)}_k V^{(0)}_k$$

$$= \begin{bmatrix} \hat{X}^{(0)}_1 & \hat{X}^{(0)}_2 & \cdots & \hat{X}^{(0)}_{M-1} & \hat{X}^{(0)}_M \\ \hat{X}^{(0)}_2 & \hat{X}^{(0)}_3 & \cdots & \hat{X}^{(0)}_M & \hat{X}^{(0)}_{M+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \hat{X}^{(0)}_{N-M+1} & \hat{X}^{(0)}_{N-M+2} & \cdots & \hat{X}^{(0)}_{N-1} & \hat{X}^{(0)}_N \\ \hat{X}^{(0)}_{N-M+2} & \hat{X}^{(0)}_{N-M+3} & \cdots & \hat{X}^{(0)}_N & \hat{X}^{(0)}_{N+1} \end{bmatrix}$$

in which $K^{(0)}$ represents the truncated order number of the iteration at the zero-th step, $T^{(0)}$ and $V^{(0)}$ are time expansion coeffients and empirical orthogonal functions of the matrix $F^{(0)}$. After replacing $X^{(0)}_{N+1}$ in $F^{(0)}$ with $\hat{X}^{(0)}_{N+1}$ in $\hat{F}^{(0)}$, we have the matrix of the first approximation

$$F^{(1)}_{(N-M+2)\times LM} = \begin{bmatrix} X_1 & X_2 & \cdots & X_{M-1} & X_M \\ X_2 & X_3 & \cdots & X_M & X_{M+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ X_{N-M+1} & X_{N-M+2} & \cdots & X_{N-1} & X_N \\ X_{N-M+2} & X_{N-M+3} & \cdots & X_N & \hat{X}^{(0)}_{N+1} \end{bmatrix}.$$

The EOF-expansion of $F^{(1)}$ yields its fitting matrix $\hat{F}^{(1)}$ of the form

$$\hat{F}^{(1)}_{(N-M+2)\times LM} = \sum_{k=1}^{K^{(1)}} T^{(1)}_k V^{(1)}_k$$

$$= \begin{bmatrix} \hat{X}^{(1)}_1 & \hat{X}^{(1)}_2 & \cdots & \hat{X}^{(1)}_{M-1} & \hat{X}^{(1)}_M \\ \hat{X}^{(1)}_2 & \hat{X}^{(1)}_3 & \cdots & \hat{X}^{(1)}_M & \hat{X}^{(1)}_{M+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \hat{X}^{(1)}_{N-M+1} & \hat{X}^{(1)}_{N-M+2} & \cdots & \hat{X}^{(1)}_{N-1} & \hat{X}^{(1)}_N \\ \hat{X}^{(1)}_{N-M+2} & \hat{X}^{(1)}_{N-M+3} & \cdots & \hat{X}^{(1)}_N & \hat{X}^{(1)}_{N+1} \end{bmatrix}.$$

Likewise, we obtain the matrix of the second approximation which is also expanded by EOFs for the fitting field $\hat{F}^{(2)}$. The procedure is repeated until the n-th iteration is performed, resuliting in

$$F^{(n)}_{(N-M+2)\times LM} = \begin{bmatrix} X_1 & X_2 & \cdots & X_{M-1} & X_M \\ X_2 & X_3 & \cdots & X_M & X_{M+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ X_{N-M+1} & X_{N-M+2} & \cdots & X_{N-1} & X_N \\ X_{N-M+2} & X_{N-M+3} & \cdots & X_N & \hat{X}^{(n-1)}_{N+1} \end{bmatrix}$$

which is then expanded by EOFs, giving

$$\hat{F}^{(n)}_{(N-M+2)\times LM} = \sum_{k=1}^{K^{(n)}} T^{(n)}_k V^{(n)}_k$$

$$= \begin{bmatrix} \hat{X}^{(n)}_1 & \hat{X}^{(n)}_2 & \cdots & \hat{X}^{(n)}_{M-1} & \hat{X}^{(n)}_M \\ \hat{X}^{(n)}_2 & \hat{X}^{(n)}_3 & \cdots & \hat{X}^{(n)}_M & \hat{X}^{(n)}_{M+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \hat{X}^{(n)}_{N-M+1} & \hat{X}^{(n)}_{N-M+2} & \cdots & \hat{X}^{(n)}_{N-1} & \hat{X}^{(n)}_N \\ \hat{X}^{(n)}_{N-M+2} & \hat{X}^{(n)}_{N-M+3} & \cdots & \hat{X}^{(n)}_N & \hat{X}^{(n)}_{N+1} \end{bmatrix}$$

For $\left\| \hat{X}^{(n)}_{N+1} - \hat{X}^{(n-1)}_{N+1} \right\| \le \varepsilon, \hat{X}^{(n)}_{N+1}$ is the forecast value at time N+1. A lot of numerical experiments show that the scheme

developed in this paper is in nature of convergent.

After the forecast value $\hat{X}^{(n)}_{N+1}$ at time N+1 is gotted, we construct an initial augumental matrix

$$F^{(0)}_{(N-M+3)\times LM} = \begin{bmatrix} X_1 & X_2 & \cdots & X_{M-1} & X_M \\ X_2 & X_3 & \cdots & X_M & X_{M+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ X_{N-M+2} & X_{N-M+3} & \cdots & X_N & \hat{X}^{(n)}_{N+1} \\ X_{N-M+3} & X_{N-M+4} & \cdots & \hat{X}^{(n)}_{N+1} & X^{(0)}_{N+2} \end{bmatrix}$$

and repeat the EOF iteration process. Then the forecast value at time N+2 can be obtained. Clearly, we may get the forecast values up to time N+τ. Especially ,when L=1, the above scheme may be used for forecasts of single variable time series.

3. Applications of the forecast scheme to the ENSO indices

Using the above forecasting scheme,we have done forecasts of Nino3 SST index

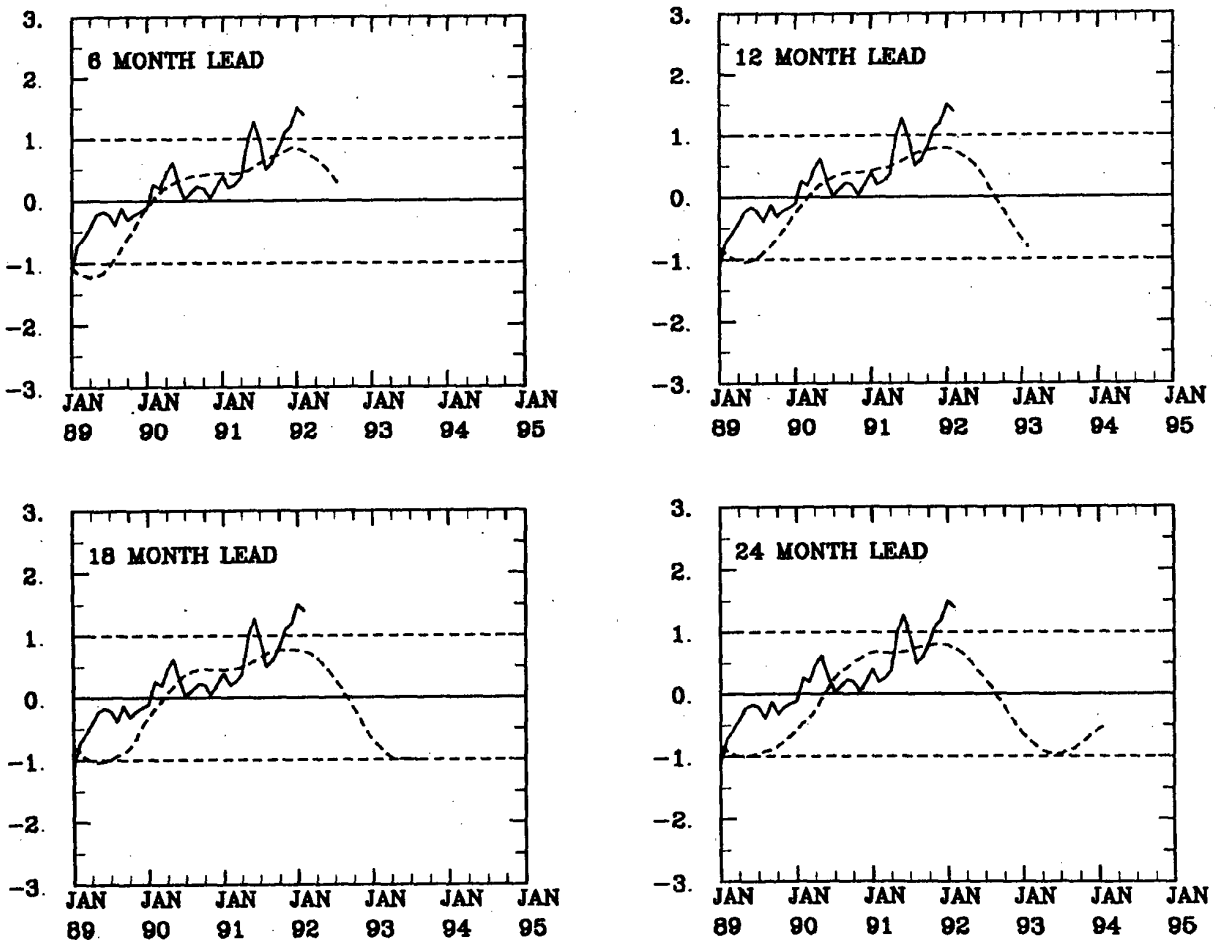


FIG.1. The forecasts for four different lead times versus observations of Nino3 SST index. Solid curve: observed. Dashed curve: predicted.

and Tahiti-Darwin SO index based on the data set of atmospheric and SST indices provided by Climate Analysis Center of USA. Forecasts of Nino3 SST anomalies for a few differnt lead times are given in Fig.1. From the figure we know that the ENSO could be predicted well up to about 24 months.

To assess the scheme's predictive capabilities, the outcome of forecast experiments can be compaired with a traditional reference-persistence forecast by defining forecasting skill (Xu and Storch,1990)

$$S^p = \frac{(O, P_\tau)}{\sqrt{O^2 P_\tau^2}}$$

where $P_\tau(t)$ and $O(t)$ are predicted for lead time τ and the observed variables.If those variables are perfectely predicted, $S^p = 1$, whereas $S^p = 0$ for a useless prediction. The forecasting skills of Nino3 SST and Tahiti-Darwin SO indices at each lead time $\tau = 1,24$ by using our forecasting scheme and persistence are respectively given Fig.2-3. From these figures, we may know: Persistence of Nino3 SST anomalies is better than our forecasts for $\tau \leq 5$ months. After that, our scheme is superior.For Tahiti-Darwin SO index, skills of persistence are almost equal to our forecasts at $\tau \leq 6$ months. After that, the scheme is also superior. Thus the forecasting scheme based on EOF iteration process is feasibility.

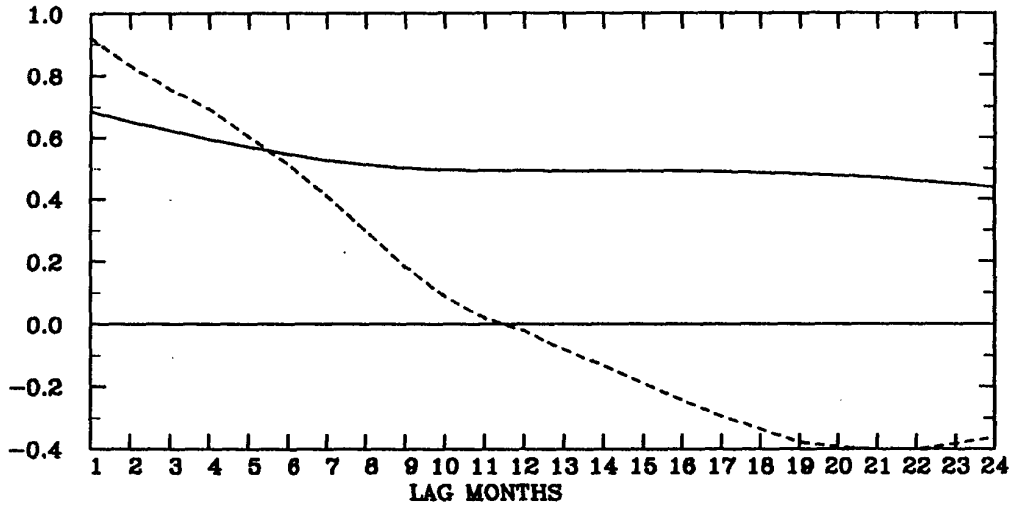


FIG.2. Correlation skill S^p of prediction of the Nino3 SST index. Solid curve: prediction based on EOF iteration process. Dashed curve: persistence.

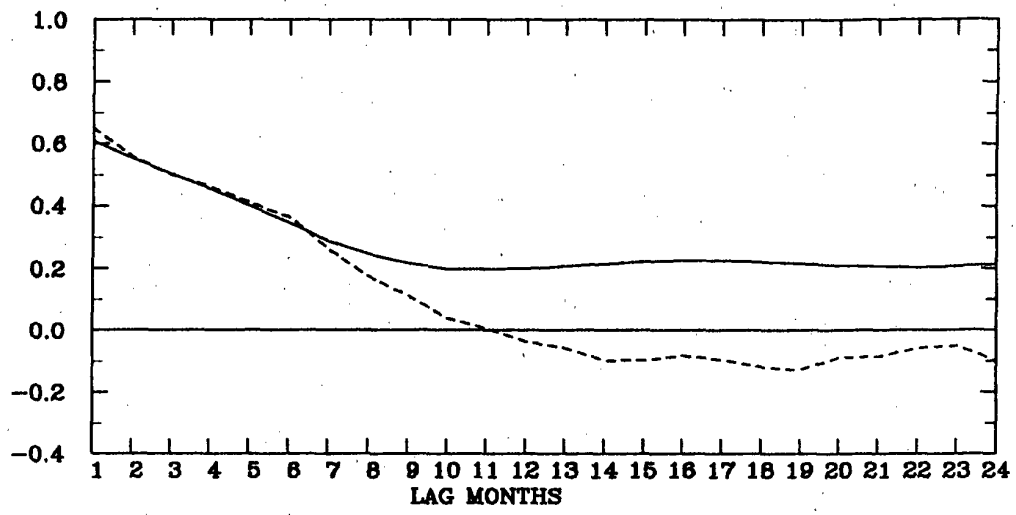


FIG.3. Same as in FIG.2, but for the Tahiti-Darwin SO index.

4. Discussion

We have outlined a new time series prediction scheme based on the EOF stepwise iteration process and demonstrated its usefulness by conducting the forecasting experiments of the ENSO indices. This method, no doubt, applicable to the prediction of other climatic elements, e.g., precipitation.For operational purpose much more work, including choice of N,L,M and definition of $K^{(i)}, \varepsilon$ for EOF iteration scheme will have to be done. Greater skill scores could be undoubtedly reached if these aspects were considered in practice.

REFERENCES

Barnett, T.P.,1981: Statistical relations between ocean/atmosphere fluctuations in the tropical Pacific. J. Phys. Oceanogr., 38,1043-1058.

Chen, W.Y., 1981: Fluctuations in Northern Hemisphere 700mb height field associated with the Southern Oscillation, Mon.Wea. Rev., 110,808-823.

Fraedrich, K.,1988: El Nino /Southern Oscillation predictability. Mon. Wea. Rev., 116,1001-1012.

Hasselman,K., and T.P. Barnett, 1981: Techniques of linear prediction for systems with periodic statistics. J. Atmos. Sci., 38, 2275-2283.

Ionue,M., and J. J. O'Brien,1984: A forecasting model for the onset of El-Nino, Mon. Wea. Rev., 116, 120-136.

Rasmusson, E.M., and T.H. Carpenter, 1982: Variations in tropical seasurface temperature and surface wind fields associated with the Southern Oscillation/ El Nino, Mon. Wea. Rev., 110, 358-384.

Ropelewski, C.F., and M.S. Halpert, 1987: Global and regional scale precipitation patterns associated with the El Nino/ Southern Oscillation. Mon. Wea. Rev., 115, 1606-1626.

Xu, J., and H.V. Storch, 1990: Predicting the state of the Southern Oscillation using principal oscilation pattern analysis. J. Climate, 3,1316-1329.

A Forecasting Model of Vector Similatity in Phase Space for Southern Oscillation

Zhou Jiabin

(Institute of Atmopheric Physics, Academia Sinica, Beijing 100029, China)

I .Introduction

Since the study on the deterministic nonperiodic flow by Lorenz in 1963 , there have been great advances in the theory on chaos. The study on predictability (See Yang et. al., 1990)makes the theory on chaos possese perspect of practical use. However, the study on predictability only provides a deadline of avaliable forecast and the problem how to make forecast is another kind of problem. Recently a new approach for prediction based on deterministic chaos has been proposed by Farmer and Sidorowich (1987) and Sugihara and May (1990). This technique was used to forecast monsoon rainfall by Kulkarni and Verma (1991). Lin Zhenshan et. al.(1991) suggested a method of predicting rainfall also. In this paper we propose a forecasting model of vector similarity in phase space for southern oscillation.

II . Reconstruction of M–Dimensional Phase Space

In practical problem we frequently study the univariate time series

$$x(t_i) = x(t_o + i\triangle t) \qquad (1)$$
$$(i = 0,2,...,N-1)$$

where $t_0$is initial moment, $\triangle$t is time interval of sample, N is total number of sample. Although x only is one in all variables representing the state of system, it contains information in all variables because they interact.

Using the method of Packard et. al.(1980) and introducing a time lag τ , a m–dimenssional phase space $R_m$is reconstructed as follows.

$$X_m(t_i) = \{x(t_i), x(t_i - \tau), x(t_i - 2\tau),...,x(t_i - (m-1)\tau)\} \quad (2)$$
$$(i = k,k+1,...,N-1; k = \frac{(m-1)\tau}{\Delta t})$$

It can be proved that the nature of geometry and information of reconstructed systerm are of equal value with that of origional system as long as m is large enough. Here τ should not less than decorrelation time of series to ensure that all coodinates in space $R_m$are independent each other.

III . Forecasting Rule I. The Distance between Predicted Point and Nearest Adjacent Point at Present is Equal to that at Next Moment

Lin Zhenshan (personal communication) proposed a forecasting model in which the

distance between predicted point and nearest adjacent point at present is equal to that at next moment. Here we show the forecasting formula for this model.

The phase point at predicted is

$$X_m(t_i) = \{x(t_i), x(t_i - \tau), x(t_i - 2\tau), \ldots, x(t_i - (m-1)\tau)\}$$

Evaluating the distances between predicted point and other points in phase space and selecting the minimum distance among them, we have

$$L_i = \{\sum_{l=1}^{m} [x(t_i - (l-1)\tau) - x(t_b - (l-1)\tau)]^2\}^{\frac{1}{2}} \quad (3)$$

where $x(t_b-(l-1)\tau)$ $(l=1,2,\ldots,m)$ is the coordinates of point $X_m(t_b)$ which is nearest to $X_m(t_i)$ in phase space.

At next monment we can evaluate the distance L_{i+1} between point $X_m(t_{i+1})$ and $X_m(t_{b+1})$ in the same way. Supposing $L_i = L_{i+1}$, we get a quadratic equation for unknown number $x(t_{i+1})$ as follows

$$y^2 + by + c = 0 \quad (4)$$

where y refers to the predicted value of $x(t_{i+1})$ and

$$b = -2x(t_{b+1}) \quad (5)$$

$$c = x^2(t_{b+1}) + \sum_{l=2}^{m} [x(t_{i+1} - (l-1)\tau) - x(t_{b+1} - (l-1)\tau)]^2 - \sum_{l=1}^{m} [x(t_i - (l-1)\tau) - x(t_b - (l-1)\tau)]^2 \quad (6)$$

Based on Exp.(4) we can get two values for $x(t_{i+1})$. Which of them satisfies our purpose? we will discuss this problem in section V.

IV . Forecasting Rule II. The Angle between Predicted Vector and the Vector Corresponding to Nearest Adjacent Point at Present is Equal to that at Next Moment.

A phase point corresponds a vector from origional point to that point. The cosine of angle between predicted vector $V_m(t_i)$ and the vector $V_m(t_b)$ corresponding to nearest adjecent point $X_m(t_b)$ can be evaluated as follows.

$$\cos\theta_i = \frac{(V_m(t_i) \cdot (V_m(t_b))}{|(V_m(t_i)||(V_m(t_b)|} \quad (7)$$

where $(V_m(t_i) \cdot V_m(t_b))$ is the inner product of vector $V_m(t_i)$ by vector $V_m(t_b)$ and $|V_m(t_i)|$ and $|V_m(t_b)|$ is the length of vector $V_m(t_i)$ and $V_m(t_b)$ respectively.

At next moment we can evaluate the angle θ_{i+1} between vectors $V_m(t_{i+1})$ and $V_m(t_{b+1})$ in the same way. Supposing $\theta_i = \theta_{i+1}$ we get a quadratic equation for unkown number $x(t_{i+1})$ as follows

$$dy^2 + ey + f = 0 \quad (8)$$

where y refers to the predicted value of $x(t_{i+1})$ and

$$d = Z_1 Z_4 - x^2(t_{b+1})$$

$$e = -2Z_2 x(t_{b+1})$$

$$f = Z_1 Z_3 Z_4 - Z_2^2$$

$$Z_1 = \cos^2\theta_i$$

$$Z_2 = \sum_{l=2}^{m} [x(t_{i+1} - (l-1)\tau) \; x(t_{b+1} - (l-1)\tau)]$$

$$Z_3 = \sum_{l=2}^{m} x^2(t_{i+1} - (l-1)\tau)$$

$$Z_4 = \sum_{l=2}^{m} x^2(t_{b+1} - (l-1)\tau) \qquad (9)$$

Based on Exp.(8) we can get two values for $x(t_{i+1})$. We also should select one of two values. This problem will be discussed in next section.

V . The Method to Determine the Predicted Value

Here we proposed the following rules to determine the predicted value.

1. As indicated before two values of predicted values of $x(t_i)$ satisfy Rule I which are represented by $y_1$and y_2. We evaluate the angle θ_1 between vector $V_m(t_{b+1})$ and the vector corresponding $y_1$and the angle θ_2 for y_2. Then we have $\theta_1-\theta_i$ and $\theta_2-\theta_i$ and get $\left|\theta_{j_0} - \theta_i\right| = \min_j \left|\theta_j - \theta_i\right|$.

If $j_o=1$ we select $y_1$as predicted value of $x(t_{i+1})$, if $j_o=2$, we select y_2.

2. Two values of predicted values of $x(t_i)$ satisfy Rule II which are represented by $y_3$and y_4. We evaluate the distance L_3 between point $X(t_{b+1})$ and the point corresponding y_3 and the distance L_4 for y_4. Then we have $L_3- L_i$ and $L_4-L_i$and get $\left|L_{j_0} - L_i\right| = \min_j \left|L_j - L_i\right|$. If $j_o=3$, we select y_3 as predicted value of $x(t_{i+1})$; if $j_o=4$, we select y_4. In order to improve forecast more points the distance between which and point $x(t_i)$ are shorter can be used. Futhermore an ensemble approach can be adopted for more predicted values.

VI. The Application of the Method to Forecast of Southern Oscillation

The mothly mean sea level pressure at Darwin is used to represent southern oscillation. The data during Jan. 1882–Feb.1992 are used and the forecast for March 1992–Feb. 1993 is made.

REFERENCES

Farmer J. D. and J. J. Sidorowich, 1987. Predicting Chaotic Time Series, Phys. Rev. Lett., 59, 845–848.

Kulkarni J. R. and R. K. Verma, 1991. Deterministic Chaos and Long Range Prediction of India Summer Monsoon Rainfall, Extended Abstracts Submitted to the ICTP / WMO Inter. Tech. Conf. on LRWF Res., WMO / TD No. 395, 147–150.

Lin Zhenshan and Li Xiangru, 1991. A Method of Anomaly Sign for Long–Term Forecast–A Similar Sign Propagation Model of Phase Space, Plateau Meteor., 10, 241–247. (In Chinese)

Lorenz E. N., 1963. Deterministic Nonperiodic Flow, J. Atmos. Sci., 20, 130–141.

Packard N. J. et. al., 1980. Geometry from a Time Series, Phys. Rev. Lett., 45, 712–716.

Sugihara G. and R. M. May, 1990. Nonlinear Forecasting as a Way of Distinguishing Chaos from Measurement Error in Time Series, Nature, 344, 734–741.

Yang Peicai and Chen Lieting, 1990. El Nino / Southern Oscillation Predictability, Chinese J. Atmos, Sci., Vol.14, No.1.

Climatological method of forecasting the severity of winter in the region of Gulf of Gdańsk

Mirosław Miętus, Institute of Meteorology and Water Management
Marine Branch, Waszyngtona 42, 81-342 Gdynia, Poland

Abstract: Spectral analysis was used to estimate the period of natural variation of mean winter air temperatures for two Polish stations, Gdynia and Hel located at the Gulf of Gdansk, southwestern Baltic Sea. Results confirm the observation that severe and warm winters appear with the recurrence of 7 or 8 years (exact period of oscillation is equal to 7.35 years). A linear operator L was constructed based on computed values of autocorrelation function. Using this operator the degree of severity for some nearest winters has been computed for Gdynia. The forecasted value (0.0°C) for Jan-Mar 1992 allows to classify this winter period as a moderate one.

1. Introduction

There are two main reasons for which climatological winter conditions at the coast of Gulf of Gdansk are of great interest for the Polish economy. The first one are two big Polish ports, Gdynia and Gdańsk, situated there, which demand reliable forecast of the intensity of winter to protect their efficient work, and the other is the nearby depression of Zulawy - an agricultural region in the delta of Vistula river - threatened in winter by floods, brought both by storm surges or by ice jamming in the Vistula estuary.

Lately several attempts were made in Marine Branch Meteorological Office of IM&WM in Gdynia (Malicki 1959, 1990, Drużyński and Paszkiewicz 1990) to gain an improved tool of forecasting the degree of winter severity in the above region. One of them is the method based on natural oscillations of thermal conditions in the region. To determine these oscillations, data of two meteorological stations of long observation series were used, in the southwestern coast of the Gulf of Gdansk: Gdynia (ϕ =54°31'N, λ = 18°34'E), and Hel (ϕ = 54°36'N, λ = 18°49'E), both operating since the early twenties of this century.

2. Natural oscillations of winter severity

To find the period of long-term thermal oscilations of the degree of winter intensity data for the winter as a whole were considered, by averaging the mean monthly temperatures of the three winter months, January, February and March, gaining in this way the mean temperature of the season for each of the both stations. The series T(t) of mean winter season temperature for the years 1923-1990 was taken as the base, what makes a population of 68 seasons. In the case of Gdynia the amplitude of variations of winter mean temperatures is 11.7°C and it is the difference between the coolest period Jan-Mar 1942 and the warmest Jan-Mar 1990 (Fig.1).

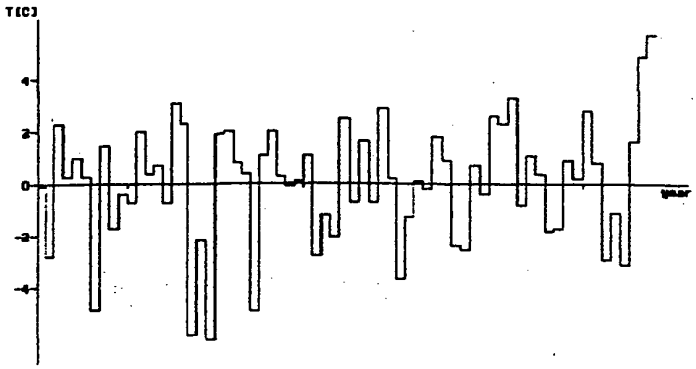


Fig.1 Variation of winter (Jan-Mar) mean air temperature for Gdynia 1923-1990

The mean temperature for all winter periods is 0°C, for the station Gdynia. The oscillations with the period 7-9 seasons are easy to see on the diagram of temperature variation. Similar characteristics we observe in the data for Hel.

The homogeneity of data creates a big problem for studies time series analysis. Results presented by Malcher and Schonwiese (1987) show that the time series of a stations may be homogeneous or not according to the criterion used. The results of the Abbe criterion, the Bartlett test and the Hartley test (Mitchell et al.,1966) allow to say that the series for Gdynia and Hel seem to be homogeneous.

For the ground series linear trend was estimated. The value of the trend for Gdynia amounted to 1.3°C in these 68 years (for Hel the trend is equal to 1.0°C). The trends removed, further calculations were carried out under the assumption that each of the series was a stationary one (the mean value is independent of time). The resulting time series were smoothed by moving overlapping means, using averaging over three, five, seven, nine and eleven seasons. For all time-series gained in this way the autocorrelation functions were calculated, too.

In both the ground, the trend removed and the smoothed series significant peaks can be observed for lags of 7 or 8 seasons. This means that the oscillations of 7 or 8 seasons are the dominating ones.

After the trend analysis showing winters temperature data sets behavior on the scale of the period of observation the data were processed using spectral analysis technic. The conventional Blackman and Tukey (1958), Parzen methods were used. The results are similar and show that there are two statistical significant peaks coresponding to lags 9 and 30 (Fig.2). Periods of oscillations are 7.56 and 2.28 seasons.

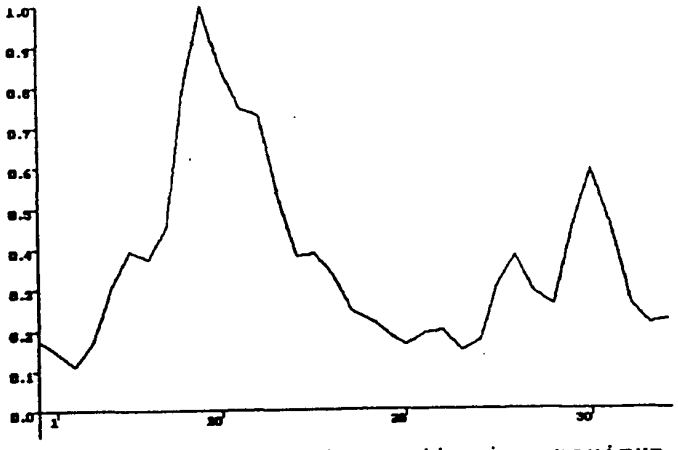


Fig.2 Normalized to unit in maximum Parzen's spectrum for Gdynia, 1923-1990

This result means that the natural oscillation of defined thermal conditions for winter in the region of Gulf of Gdansk amounts to about 7.56 years and is modulated by oscillations with the period 2.28 years. Similar results are true for Hel, too.

After Thomson (1982) we know that no single method is suited ideally for the study of climatological variables. Due to this reason the "maximum entropy" method (Burg, 1967), which allows to find the pseudo periodicities and to estimate better the amplitude relationship was applied additionally (Fig.3).

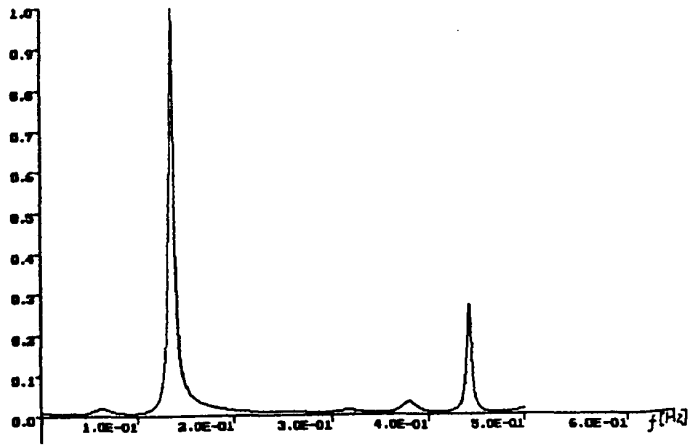


Fig.3 Normalized to unit in maximum "maximum entropy" spectrum

Testing the spectra by means of Markov red noise and χ^2 tests (Mitchell 1966) we received that two peaks coresponding to the periods of 7.35 and 2.27 are significant with respect to 95% confidence level.

3. Forecasting the severity of winter

The results gained as described in chapter 2 can be applied to foresee the degree of winter severity in the region. The reasoning is as follows.
Assuming that the realization of elements X(t) is known in the time $t-\tau_o$ and in the time t, we could compute the realization of X in the time $t+\tau$.
For this purpose we assume that there exists a linear operator L, such that

$$X(t+\tau)=L(X(t))=aX(t)+bX(t-\tau_o) \qquad (1)$$

which satisfies the condition

$$F(a,b)\equiv E(\left|X(t+\tau)-L(X(t))\right|^2)=\min \qquad (2)$$

where F(a,b) is a function of coefficients a and b and condition (2) is connected with the stationarity of the data set.
Based on the definition of autocorrelation function R and due to the reason that F(a,b) is positively defined condition (2) can be taken in equivalent form

$$\frac{\partial F}{\partial a}=0 \qquad \frac{\partial F}{\partial b}=0 \qquad (3)$$

To facilitate the calculations we put mean temperature $T_{mean}=0^{\circ}C$ (this is true for the series of mean winter temperature in Gdynia, as calculated before), and receive

$$X(t+\tau)=\frac{R_\tau + R_{\tau+\tau o}R_{\tau o}}{1-R^2_{\tau o}}X(t)+\frac{R_{\tau+\tau o}-R_{\tau o}R_\tau}{1-R^2_{\tau o}}X(t-\tau_o) \qquad (4)$$

what makes the desired tool and enables the computation of the mean temperature of winters ahead.

Using this formula, the values for winters of 1991 and 1992 have been computed (Tab.1) for Gdynia

Table 1. T_f-forecasted, T_a- actual, $\Delta=T_f-T_a$, $^{\circ}C$.

		Mean temperature		
Season	Year	T_f	T_r	Δ
Jan-Mar	1991	1.9	1.8	0.1
Jan-Mar	1992	0.0		

4. Closing remarks

The series of 68 years which we had to our disposal forms a set which must not necessarily be sufficiently long and some other oscillations may come to evidence with the time growing (and the series growing). We hope, however that the estimated oscillations of 7.35 and 2.27 years in the thermal conditions as a new means of long-term forecasting will contribute to better meteorological protection of the region.

The forecasted mean temperature for the winter season 1992 ($0.0^{\circ}C$) is close to mean values of Jan-Mar temperatures of the years, 1950-52, 1962, 1965-66, 1971, 1978, 1982 which were, according to classification by Druzynski and Paszkiewicz (1990) classified as moderate winters.

References

Blackman, R.B. and Tukey,J.W.,1958. The measurement of power spectra, Dover Publ. NY.

Burg J.P.,1967. Maximum entropy spectral analysis, Paper presented at 34th meeting Soc. Explor. Geophys., Oklahoma City.

Drużyński B., Paszkiewicz Cz.,1990. Method of the winter severity forecasting on the Polish coast, News of Marine Branch of IM&WM, (in Polish)

Malcher J.,Schonwise Ch.D.,1987. Homogeneity, spatial correlation and spectral variance analysis of long European and North American air temperature records, Theor. Appl. Climatol. 38,157-166.

Malicki, J.,1959. Contribution to long-term forecast of winter severity in the Baltic region, Przeglad Geofizyczny, vol.4 (XII), Z.3-4, 231-236

Malicki, J.,1990. Periodicity of thermal conditions and the resulting prospect of forecasting the severity of approaching winter time, News of Marine Branch of IM&WM, (in Polish)

Mitchell J.M., et al.,1966. Climatic change, WMO - No.195. TP.100: Geneva

Thomsom D.J.,1982. Spectrum estimation and harmonic analysis, Proc. IEEE, 70 (9), 1055-1096.

THE IMPACT OF BIASES IN PRECIPITATION TIME-SERIES

David R. Legates
College of Geosciences
University of Oklahoma

INTRODUCTION

With the concern over global climate change, variations in the spatial and temporal distribution of precipitation have become an important topic. Detecting and predicting fluctuations in global precipitation is useful in the evaluation of global warming since precipitation is a measure of the vertically-integrated diabatic heating field. Changes in precipitation and its intra-annual variability also are important in agricultural and economic applications, the detection of the impact of land-use changes, and the design of culverts and stream channels.

Unfortunately, precipitation measurement is subject to several influences which may considerably degrade its accuracy and reliability. Systematic errors in precipitation gage measurement can be attributed to six sources — the effect of the wind, wetting and evaporative losses, splashing, blowing snow, and automatic recording techniques — while random errors add an unsystematic component (*cf.*, Legates, 1987). These errors have been widely recognized as potential biases in precipitation measurements although their impact largely has been ignored in most climatological analyses. This is due, in part, to a misconception that although precipitation measurements are biased, these biases can be removed by applying a correction factor that varies by month. While these corrections are greater in winter than summer due to the deleterious effects of the wind on snowfall, many proposed corrections do not take into account year-to-year climate variability that affects precipitation measurement. When considering precipitation variability, it therefore appears that these biases can be ignored since, once the seasonal cycle is removed,

$$\sigma_c^2 = \sigma_{o+b}^2 = \sigma_o^2 \qquad (1)$$

where σ_c^2 and σ_o^2 are the variance in the "true" and observed precipitation time-series, respectively, with b as the bias in gage measurement. If the bias is not a constant but varies from year to year, however, this assumption is necessarily incorrect.

THE INTRA-ANNUAL VARIABILITY IN GAGE BIASES

Legates (1992) considered a hypothetical station where significant winter snowfall occurs and summer air temperatures rise well above freezing. For this station, the precipitation time-series exhibits no adverse effects caused by changes in instrumentation and recording practices, location, and siting characteristics.

Mean monthly air temperatures and wind speeds, however, would fluctuate from year to year due to intra-annual variability. In particular, air temperature variations would vary the frequency of snowfall during spring and autumn and, since gage catch decreases with increasing wind speed, the measured precipitation time-series would still exhibit intra-annual variability even if the actual precipitation was the same year after year.

This paper evaluates the applicability of Legates (1992) using climatological time-series of air temperature, precipitation, and wind speed from five stations: Allentown, PA, Des Moines, IA, Duluth, MN, Portland, ME, and Syracuse, NY. These stations were selected because a continuous record from 1950 through 1987 exists and significant winter snowfall does occur on average. Gage corrections were estimated using the procedure outlined by Legates (1987). Gage errors range from three percent to about thirty-six percent and are greater in winter than summer (Table 1). They also are commensurate with the monthly maps of Legates (1987).

As expected, these gage measurement biases exhibit a strong seasonal variation and increase as the monthly precipitation increases. This should result, therefore, in a decrease in the observed variance (Table 2). However, these biases do not always result in an enhancement of the seasonal cycle. The variance of the seasonal cycle in the observed time-series, σ^2_{so}, is greater than that of the corrected time-series, σ^2_{sc}, for Allentown, Des Moines, and Duluth while it is less for Portland and Syracuse. This occurs since Portland and Syracuse exhibit a winter precipitation maxima while the other three stations have significant summer precipitation. The variance in the gage error also is larger for Portland, Syracuse, and Duluth owing, in part, to a greater variability in wind speed.

CLIMATE CHANGE IMPLICATIONS

An assessment of climate change often involves the ability to detect the signal from the noise (*cf.*, Leith, 1973; Madden, 1981; Trenberth, 1984). In the case of the 'true' precipitation time-series, the decomposition can be written as

$$\sigma^2_c = \sigma^2_{sc} + \sigma^2_{nc} + \sigma^2_{pc} \qquad (2)$$

where σ^2_{sc} is the variance explained by known signals (*e.g.*, the seasonal cycle), σ^2_{nc} is the climate noise or the effect of persistence, and σ^2_{pc} is the residual or that proportion of the variance which can be explained by potentially predictable components (Madden, 1981). Shukla (1981) has pointed out, however, that the term 'climate noise' may be somewhat misleading "since it has not yet been established that changes due to internal dynamics are unpredictable" (p. 2551). Often climate noise can be estimated by assuming that persistence can be modelled by a Markov model and thus depends on the lag 1 autocorrelation of the detrended series (Madden, 1979) although higher order autocorrelation models may be used (*cf.*, Jones, 1975). Therefore, the power of detecting climate change depends on the ratio of the remaining variance attributable to climate change to the climate noise. Leith (1973) indicates that this ratio must be at least unity before significant conclusions about climate change can be drawn.

Biases in the precipitation record, however, adversely affect this variance decomposition. If the observed time-series is used, then equation (2) is really

Table 1: Estimated mean monthly gage measurement errors (in percent) for the five stations

Month	Allentown	Des Moines	Duluth	Portland	Syracuse
January	20.0	32.0	36.0	23.4	24.5
February	18.7	27.2	33.5	22.0	22.5
March	11.5	18.1	30.3	16.4	15.8
April	6.2	8.7	15.7	8.5	7.0
May	4.5	6.5	7.6	5.8	4.8
June	4.0	5.8	6.1	4.7	4.3
July	3.4	5.0	5.2	4.3	4.1
August	3.4	5.0	5.3	4.2	3.9
September	3.5	5.4	6.3	4.5	4.2
October	4.5	6.3	9.2	5.6	5.2
November	7.3	13.5	25.0	9.3	9.2
December	15.0	26.0	32.5	18.8	18.9

Table 2: Variance estimates for the five stations (in mm^2)

Variance of the ...	Allentown	Des Moines	Duluth	Portland	Syracuse
Observed record, σ_o^2	2465.6	2886.1	2100.3	2534.8	1383.1
Corrected record, σ_c^2	2818.6	3154.0	2333.8	2490.8	1717.7
Gage error, σ_e^2	68.9	46.6	89.1	176.9	91.3
Seasonal cycle in the observed record, σ_{so}^2	120.9	984.8	853.9	234.2	78.8
Seasonal cycle in the corrected record, σ_{sc}^2	79.2	871.6	640.9	498.2	112.0
Covariance between the observed record and gage error, $\sigma_{o,e}$	142.1	110.6	72.2	389.5	121.6
Correlation between the observed record and gage error	0.345	0.302	0.167	0.582	0.342
Lag 1 autocorrelation of the observed record	0.0426	-0.0115	0.1088	0.0462	0.1271
Lag 1 autocorrelation of the corrected record	0.0464	0.0078	0.1144	0.0482	0.1212

$$\sigma_o^2 = \sigma_{so}^2 + \sigma_{no}^2 + \sigma_{po}^2 \quad (3)$$

but it is not clear that the ratio of the signal to the noise is similar. This is important since the correlation between the observed time-series and the gage bias is significant (Table 2). Thus, the relationship between the observed record and the gage error can be written as

$$\sigma_c^2 = \sigma_{o+g}^2$$

$$= \sigma_{so}^2 + \sigma_{no}^2 + \sigma_{po}^2 + \sigma_e^2 + 2\sigma_{o,e} \quad (4)$$

where σ_e^2 is the variance of the gage biases and $\sigma_{o,e}$ is the covariance between the observed time-series and the gage biases. Thus,

$$\sigma_c^2 - \sigma_o^2 = \sigma_e^2 + 2\sigma_{o,e} \, . \quad (5)$$

However, the error does not just deflate the observed variance in precipitation but masks some of the potentially predictable variance. This occurs since the lag 1 autocorrelation is not significantly affected (Table 2) and thus the magnitude of the climate noise is similar for both the corrected and observed time-series. Thus, the signal-to-noise ratio is decreased by the gage biases which hampers the detection of any climate change signal.

RECOMMENDATIONS

It is apparent that the biases in precipitation gage measurement have a significant impact on the variability in precipitation time-series and thus cannot be treated as a seasonally-varying constant. Therefore, it is imperative that reliable correction procedures for estimating and removing these biases be developed to enhance the efficacy of research which utilizes precipitation time-series.

ACKNOWLEDGEMENTS

I would like to thank Ms. Tracy DeLiberty for assistance with the data compilation. This research was funded under NSF grant ATM-9015848.

REFERENCES

Jones, R.H., 1975: Estimating the variance of time averages. *J. Appl. Meteor.*, **14**, 159–163.

Legates, D.R., 1987: A climatology of global precipitation. *Publ. in Climat.*, **40**(1), 84pp.

Legates, D.R., 1992: The need for removing biases from rain and snow-gage measurements. *Proceedings, Snow Watch '92*, Niagara-on-the-Lake, Ontario, forthcoming.

Leith, C.E., 1973: The standard error of time-average estimates of climatic means. *J. Appl. Meteor.*, **12**, 1066–1069.

Madden, R.A., 1979: A simple approximation for the variance of meteorological time averages. *J. Appl. Meteor.*, **18**, 703–706.

Madden, R.A., 1981: A quantitative approach to long-range prediction. *J. Geophys. Res.*, **86**, 9817–9825.

Shukla, J., 1981: Dynamical predictability of monthly means. *J. Atmos. Sci.*, **38**, 2547–2572.

Trenberth, K.E., 1984: Some effects of finite sample size and persistence on meteorological statistics. Part II: Potential predictability. *Mon. Wea. Rev.*, **112**, 2369–2379.

Credibility of the Measured Precipitation at Ocean Weather Station "P"

K. Higuchi, W. Wong*, M.A. Jenkins*, and J.L. Knox
Atmospheric Environment Service, 4905 Dufferin Street, Downsview, Ontario M3H 5T4
*Department of Earth and Atmospheric Sciences, York University, North York, Ontario

1. Introduction

Knox (1991) examined the credibility of the precipitation record at Ocean Weather Station "P" (50N, 145W) in the northeast Pacific, with respect to "instrumentation, exposure, observer expertise, and relative consistency with contemporaneous synoptic conditions." In his article, Knox indicated not only the need for further investigation of the OWS "P" measured precipitation, but also suggested a method of re-examination of the data. In this paper, we extend the study of Knox (1991) by investigating the data in terms of their statistical and physical consistency. We present results of 3 methods: (1) Application of the Tucker technique (Tucker, 1961), using the values of X,Y,Z as derived by Tucker, (2) Application of the Tucker technique, but deriving the values of X,Y,Z from the 3-hourly "ww" observations during the 1954-1966 period of the OWS "P" record, and (3) Product of vertical velocity multiplied by mixing ratio. Because it is the cold season precipitation measurements after 1969 that are questionable, we will examine winter precipitation only in this paper.

2. Tucker Method 1

In this first method, we employ the technique developed by Tucker (1961). We also use his calculated values of the X,Y,Z parameters (X=1.85, Y=5.66, Z=8.13). These parameters give estimated rainfall rates in mm hr^{-1} related to light rain (X), moderate rain (Y), and heavy rain (Z). Table 1 in Tucker's paper is used to relate the present weather code numbers to precipitation rate in terms of X,Y,Z. The result is shown in Fig. 1(a). Estimated winter mean (average of December,January, February) for each year from 1954 to 1980 is shown against the measured winter mean. The calculated precipitation using the X,Y,Z values derived by Tucker consistently overestimates the observation, even prior to 1970. The extreme minimum observed in 1969 is also observed at Victoria (Knox, 1991). After 1969, calculated and observed values do not agree, even in trend. The result is consistent with those obtained by Reed and Elliot (1973), who performed similar calculations.

3. Tucker Method 2

In this method, we again use the technique developed by Tucker (1961), but we calculate the values of X,Y,Z using the observed weather at OWS "P" from 1954 to 1966, to determine the most likely hourly precipitation rate associated with a given weather "ww". Table 1 in Tucker's paper is again used to generate for each year estimated mean winter precipitation. The minimum in 1969 is quite well simulated. However, unlike the observation, the estimated precipitation recovers in the following winter and tracks the observation in trend but remaining about 15 to 20 mm above the measured values (Fig. 1(b)). There appears to be some systematic undermeasurements of precipitation after 1969.

4. (Vertical Velocity) X (Mixing Ratio) Method

In this method, we estimate precipitation rate by vertically integrating the product of calculated vertical velocity and observed mixing ratio over OWS "P". The vertical velocity over the station is calculated following the method outlined by Penner (1963). Using the 500 mb and the 1000-500 mb thickness fields on a 455 northern hemisphere grid from the Canadian Meteorological Centre, we calculate the vertical velocity for each day from 1948 to 1987 as a sum of absolute vorticity advection and thickness temperature advection. These vertical velocity values are then multiplied by daily values of the mixing ratio. The result from 1960 to 1976 is shown in Fig. 1(c). The calculated values have been shifted upward by 30 mm for easy comparison with the observation. Again, the calculated precipitation estimate recovers after 1969. It is also interesting to note that the trend of the calculated values is very similar to that seen in the measured values.

5. Summary

We feel that the Method 1 is not really applicable at OWS "P". The results obtained from Tucker Method 2 and the (vertical velocity) X (mixing ratio) method suggest that the measured precipitation at OWS "P" during the winter season is a gross underestimate of the "real" precipitation amount after 1969. However, the trend seems to be real.

6. References

Reed, R.K. and W.P. Elliot, 1973: Precipitation at ocean weather stations in the North Pacific. J. Geophys. Res., 78, 7087-7091.

Knox, J.L., 1991: An assessment of the 27-year record of measured precipitation at Ocean Weather Station "P" in the Northeast Pacific Ocean. Clim. Bull., 25, 65-80.

Penner, C.M., 1963: An operational method for the determination of vertical velocities. J.Appl.Met.,2, 235-241.

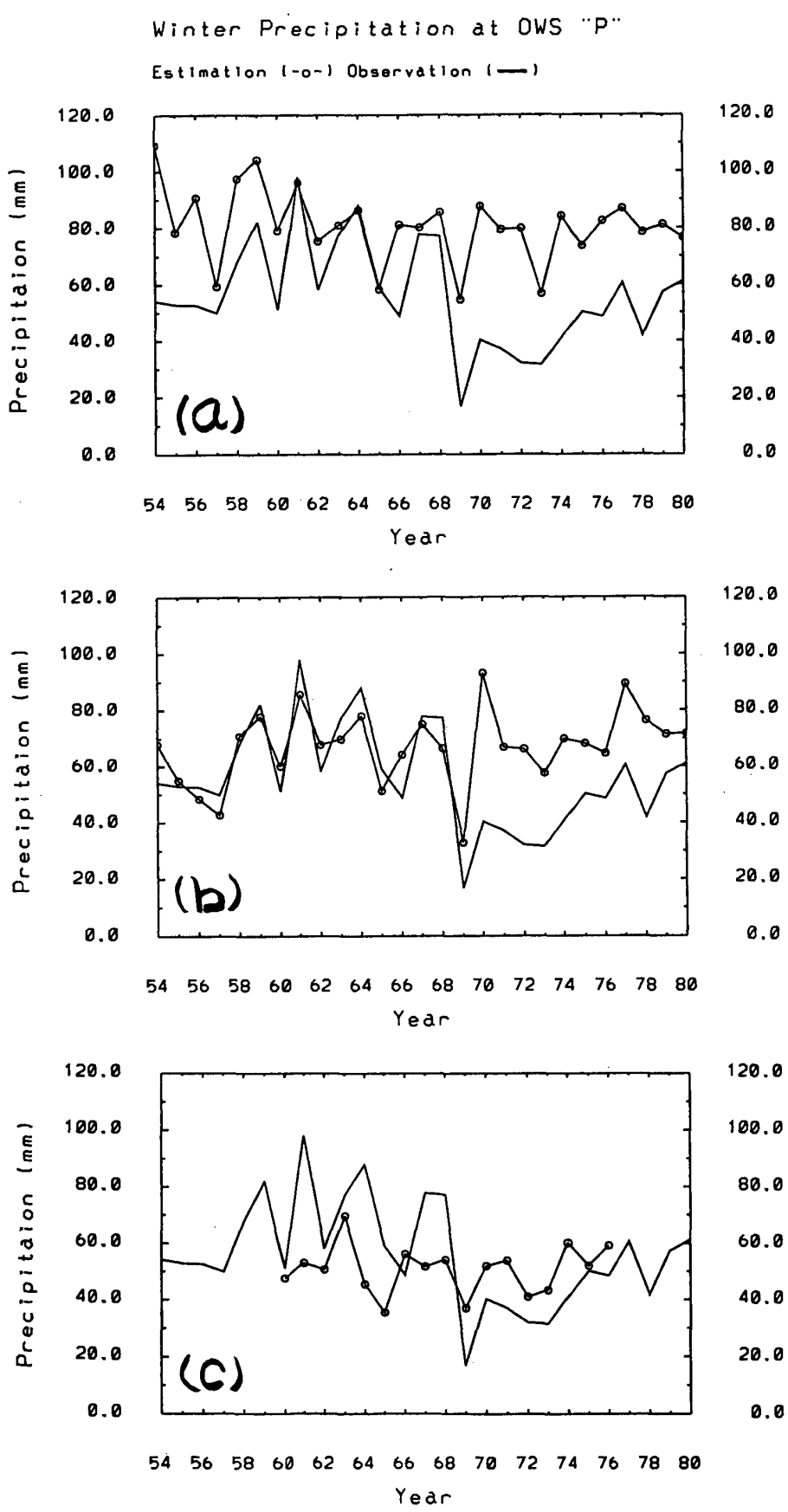
Tucker, G.B., 1961: Precipitation over the North Atlantic Ocean. Quart.J.R.Met.Soc., 87, 147-158.

7. Figure Captions

Figure 1(a): Measured mean winter precipitation values are shown in the solid line. The connected circles indicate calculated values obtained using the Tucker's method and his X,Y,Z values.

Figure 1(b): Same as Figure 1(a), but calculated values (connected circles) are obtained from the application of the Tucker's method with X,Y,Z values derived from the 1954-1966 OWS "P" record.

Figure 1(c): Same as Figure 1(a), but calculated values (connected circles) are obtained from the product of vertical velocity multiplied by mixing ratio.


Winter Precipitation at OWS "P"
Estimation (-o-) Observation (——)
Precipitaion (mm)
120.0
100.0
80.0
60.0
40.0
20.0
0.0
(a)
54 56 58 60 62 64 66 68 70 72 74 76 78 80
Year
Precipitaion (mm)
(b)
Year
Precipitaion (mm)
(c)
Year

Semi-automatic Quality Control of Daily Precipitation Measurements

Henning Madsen

The Danish Meteorological Institute

Introduction.

The existing procedure for quality control in Denmark is based on a subjective inspection of precipitation data by comparisons of some 10 neighbouring stations for error detection in consideration of weather situations.

In several places over the world, attempts have been made to develope quality control routines (Bryan, 1979 and Abbot, 1986). Semi-automatic systems for analyzing daily precipitation values have been developed in Norway and Sweden (Dahlström et al, 1980 and 1988).

A few years ago a semi-automatic quality control system with a technique based on a statistical model was developed in Denmark (Madsen, 1989). This paper is especially concentrated on a non-parametric technique, in which the daily standardized median value for the measurements occur as a reference value for the other measurements.

Data.

Daily precipitation data from 12 neighbouring stations including the period 1976 - 1980 are used in this investigation. The stations represent an area of roughly 1000 km^2, and are situated in a gently undulating landscape indicating, that the terrain will not give rise to special effects influencing the precipitation pattern.

The available data series consist of [1] raw data and [2] the same data subjected to manuel quality control. The procedure for such a quality control includes error detection, error diagnosis and correction and interpolation.

Now we will get possibility of comparing the corrected and interpolated data series to the raw data series, which have been controlled by means of an automatic quality control method.

For convenience we have fixed the length of the data series to 1800 days (nearly 5 years). The number of precipitation days in Denmark are 159 in a year, and the total number of precipitation days in 1976-80 for the 12 stations on the conditions, that $\geq$ 1 station has recorded precipitation are 1.321 days corresponding to 73% of the total number of days.

Test statistics.

To avoid the influence of outliers, the test statistics

$$(1) \qquad T_{it} = \frac{x_{it} - M_t}{q_{t,75} - q_{t,25}}$$

are studied. Here x_{it} is the observation at station no i for day no t, M_t the median value of the observations from the 12 stations, and $q_{t,75}$ and $q_{t,25}$ are quantile values of the 12 observations.

The advantage of using (1) is, that erroneaus data will often be extreme, and therefore leave the median value as probably "true". In that way (1) will not be dependent on a single or a few outliers.

By "trial and error" technique following condition for T_{it} concerning marking of erroneaus data has been fixed:

(2) $|T_{it}| > 2{\cdot}00$, and $x_{it} > 4$ (tenth of mm) which is indicated as "1" (see example no. 1 and no. 4). In case of $q_{t,75} - q_{t,25} = 0$, which especially occurs in dry weather situations, (1) must be complementet with

(3) $x_{it} / \sum x_{it} > 0{\cdot}60$, and $x_{it} > 4$ (tenth of mm) indicated as "2" (ex. no. 2 and no. 5).

Finally in situations where observations are missing (0) we have:

(4) $x_{it} = 0$, i = 1, 2 or 3, and $x_{it} \neq 0$ for the other observations, (ex. no. 3 and no. 6).

Results.

Comparison of the two data series controlled formerly by a manual quality control and by the above mentioned non-parametric technique respectively, has pointed out some important things presented in the table and the examples below.

Table 1 show results from the comparison between the two data series after controlling all days in the period 1976-80.

		Manual control		
		not marked	marked	Σ
Automatic control	not marked	18.433	725	19.158
	marked	637	636	1.273
	Σ	19.070	1.361	

Table 1. Numbers of marked and not marked data of daily precipitation from 12 stations in the period 1976-80 in relation to manual and automatic quality control.

Superficially the two methods give nearly the same number of marked data. However the number of data marked by both control methods only amounts to about the half, 636. When looking at cases where the manual control has marked the data and the automatic not, and vice versa, we get nearly the same numbers, 725 and 637 respectively. To give an explanation of this, a consideration of the markings on the basis of the single days in 1976-80 was decided. After examining the data series to see how the observations have been marked by means of the test statistics, and after having noticed how these markings correspond with the markings of the manual control, some conclusions can be drawn concerning the results of the two control methods. Some examples will illustrate the situation.

The first 3 examples show cases, where the markings of the two control methods are corresponding. Each of the 3 test statistics are represented by an example. In the examples the upper coloumn contains observations from the 12 stations corrected and/or interpolated by means of the manual control, wheras the coloumn beneath are the raw data marked by the automatic control.

In ex. 1, the observation no. 3 has been corrected to 30 mm according to the manual control and the analogous raw value of 13.0 mm has been marked by "1", since $T_{it} > 2$ (cf (2)). A typical example of noting rainfall on a wrong day is illustrated in ex. 2. Here the observation no. 10 is corrected to 0, and the original value 2.1 mm has been marked by "2", in which case the test statistics (3) is used. It can be noticed, that only in few cases has it been necessary to apply (3). Ex. 3 shows a situation with missing observations for 3 stations, and they are all interpolated. The same observations are marked according to (4). The three examples represent situations not very complicated with respect to performing a quite good quality control, manual as well as automatic but, as mentioned, these situations only include about half of the selected suspect data. Example 4-6 demonstrates situations where no correspondence occur between the markings for the two control methods.

In ex. 4 observation no. 2 is marked by the automatic control due to the test statistics $T_{it} > 2$. However, no correction has been done after the manual inspection, since the weather situation is classified as showers, so that the amount 3.1 mm has been accepted as true.

Y	M	D	Precipitation (tenth of mm) for station no. 1	2	3	4	5	6	7	8	9	10	11	12	
			43	42	30	30	32	32	18	27	22	60	20	50	
78	2	14	43	42	130	30	32	32	18	27	22	60	20	50	Ex.1
		mark.			1										
			0	0	0	0	0	0	0	0	0	0	0	0	
76	1	14	0	0	0	0	0	0	0	0	0	21	0	0	Ex.2
		mark.										2			
			14	15	20	15	18	11	12	17	20	15	13	10	
76	11	2	14	15	0	15	18	11	12	17	0	15	13	0	Ex.3
		mark.			3						3			3	
			0	31	0	7	25	0	13	0	0	12	0	0	
76	6	11	0	31	0	7	25	0	13	0	0	12	0	0	Ex.4
		mark.		1											
			0	0	0	0	0	8	0	0	4	0	0	0	
77	5	5	0	0	0	0	0	8	0	0	4	0	0	0	Ex.5
		mark.						2							
			18	1	23	17	0	3	3	0	5	10	5	2	
78	2	17	18	1	23	17	0	3	3	0	5	10	5	2	Ex.6
		mark.					3			3					

Examples.

The marking of observation no. 6 in ex. 5 is applied to (3), but also in that case no correction has been undertaken due to showers. Finally, another situation (ex. 6) with showers has caused markings of two observations (no. 5 and no. 8). In both cases no rainfall are recorded, which also has been accepted by the manual control.

It must be pointed out, that nearly all cases where markings performed by the automatic quality control have not been accepted, are due to rainfall caused by convective cells.

Such cases will be difficult to exclude from a marking when using an automatic quality control system.

Concerning the last problem dealing with observations marked by the manual control but unmarked by the automatic control, it seems to be owing to incorrect rainfall measurements as for example, when the observer has noted rainfall to the wrong day for the whole month or part of month, and in case of accumulations and incorrect time of observation.

Such conditions make it difficult to find a reasonable algorithm.

On a long view, we direct our effords to a semi-automatic quality control system, where selection of erroneaus data and, in case of large-scale precipitation systems, also correction and interpolation of data can be undertaken automatically. In this system manual checking should be devoted to data deriving from convective cells or other small-scale systems.

References.

Abbott, P.F., 1986: Guidelines on the quality control of surface climatological data, WMO/TD-No. 111.

Bryant, G.W., 1979: Archiving and quality control of climatological data, Met. Mag., 108.

Dahlström, B., et al., 1980: A basic system for quality control of observations.

Dahlström, B., et al., 1988: Kvalitetskontroll av nederbördsdata vid Nordiska meteorologiska institut. ViN, No. 3.

Madsen, H., 1989: Quality control of Precipitation Measurements in Denmark. Fourth International Meeting on Statistical Climatology. Rotura, New Zealand, 1989.

Analysis of Homogeneity in Precipitation Measurements

Peter Allerup

The Danish Meteorological Institute

Abstract.

In series of precipitation measurements non-homogeneous observations are most frequently detected using a set of neighboring observations creating a reference value and calculated statistics -eg. ratios - are then submitted to statistical analysis. Different statistical 'tools' are used in this process but very seldom it is clear from the analysis how non-homogeneity is formally defined in contrast to a hypothesis of homogeneity. The aim of this paper is to suggest one formalization of homogeneity in terms of a comprehensive statistical model for precipitation data and to demonstrate how this model can be tested on real data.

Introduction.

When analyzing precipitation series in order to detect non-homogeneous observations, frequently the method is to select one 'suspicious' station and then make comparisons with other neighboring stations. The type of inhomogeneity in the selected station could be some point of discontinuity in the series or a developing trend of deviation across time as typical examples. In this paper no pre-knowledge concerning neither a single series to be 'suspicious' nor the specific type of inhomogeneity is assumed. The hypothesis of homogeneity concerns two different aspects (1): The distribution of expected precipitation in time and space (temporal and site specific characteristics) and (2): The structure of spatial correlation among the series.

Data.

The data used are measurements of monthly precipitation values in the period 1931-60 from 10 stations in the northern part of Denmark distributed evenly across an area of approx. 150 by 150 km. The data was first corrected according to usual quality check procedures.

Homogeneity.

Let x_{ti} denote $\log(y_{ti})$ where y_{ti} is the observed precipitation at station no. i and month no. t. We will say that the 10 (k) stations are homogeneous if the observations fit the following statistical model:

$$(1)\quad E(x_{ti}) = \alpha_i + \beta_t \qquad x_{ti} \sim \text{Normal}(0,\sigma_i^2)$$

$$(2)\quad \text{cov}(x_{ti},x_{tj}) = \gamma(t)[D(i,j)]^{\tau}$$

The formal specification (1) of homogeneity states that the expected (logarithmic) precipitation can be calculated in a two-way parametric frame with $\alpha_1,\ldots,\alpha_k$ as **site specific** parameters and $\beta_1,\ldots,\beta_n$ as **temporal specific** parameters (characterizing the general level for each month). The specification (2) specifies the spatial correlation function to be dependent if inter distance D(i,j) with structural parameters γ and τ to be estimated from data. This model has earlier proved to be successful when analyzing series of precipitation measurements (Allerup et al, 1982,1983). In this case we adopt the content of the mod-

el: **independence** of anything else than the **general** influence of the site and the month together with the **isotropic property** of the correlation function as the very basis for the definition of homogeneity.

Test statistics.
It is a consequence of the model (1), i.e under the hypothesis of homogeneity, that the residuals

$$(3) \quad r_{ti} = x_{ti} - x_{.i} - x_{t.} + x_{..} \approx 0$$

are randomly (normal) distributed around zero. From the r_{ti}'s we form the average cumulative studentized residuals

$$(4) \quad t_i(p) = 1/p \sum_{t \le p} r_{ti}/\sigma(r_{ti})$$

$$\text{with} \qquad \sigma(r_{ti}) = \sqrt{\text{var}(r_{ti})}$$

The $t_i(p)$'s will converge to zero if the series are homogeneous according to $\pm\epsilon$ confidence bounds, $\epsilon \downarrow 0$ for $p \to \infty$. For practical reasons we analyze the following $T_i(p)$'s which are Normal(0,1) if the series are homogeneous.

$$(5) \quad T_i(p) = t_i(p) \bullet \sqrt{p}$$

A plot of $T_i(p)$ for p=1,..n, n= no. of months in the series, i=1,...,k are the test plots for **1.order homogeneity** (i.e. test of hypothesis (1)).
In order to build a test for **2.order homogeneity** (i.e. test of hypothesis (2)) we rearrange homogeneity condition (2):

$$(6) \quad \log(1-\rho_{ij}) = \log(\gamma(t)) + \tau \log(D(i,j))$$

where ρ_{ij} is the correlation between the series nos.i and j. With estimated values of the structural correlation function parameters γ and τ and the correlations ρ_{ij}'s the series must fit into the linear relationship given by (6) in order for being **2.order homogeneous**. The test plot of $\log(1-\rho_{ij})$ versus $\log(D(i,j))$ (D=distance) may reveal a number of k-1 points out of all possible combinations of k correlations which don't conform with the linear structure (6). In such a case the deviating points may come from one series only indicating that this series correlates with the other series different from the isotropic spatial correlations defined by the other series.

Justification for test statistics.
Since we assume that no pre-knowledge concerning a suspicious series or the type of inhomogeneity is available me might have adopted a simple doublemass technique for the investigation of inhomogeneity. With k series at hand in stead of the usual two-by-two comparisons this will lead to the calculation of the cumulative precipitation at each site $c_{pi} = \sum_{t \le p} x_{ti}$:

	site			
month	1 .... i ...		k	total
1	c_{11}	c_{1i}	c_{1k}	$c_{1.}$
2	c_{21}	c_{2i}	c_{2k}	$c_{2.}$
.	.	.	.	.
.	.	.	.	.
p	c_{p1}	c_{pi}	c_{pk}	$c_{p.}$
.	.	.	.	.
n (total)	$c_{.1}$	$c_{.i}$	$c_{.k}$	$c_{..}$

At line p=month no. p we observe the total amount of gained precipitation c_{p1} ..c_{pk} at each of the k sites together with the total areal value $c_{p.}$. The principle of doublemass requires the fraction collected at a specific site in relation to the total areal value to be constant, i.e. $c_{pi}/c_{p.} \approx \text{const(i)}$. We take the total amount of precipitation collected across all months $c_{.1},...,c_{.k}$ to act as 'golden standards' for estimating the final values for these constants:

$$(7) \quad c_{.i}/c_{..} \approx \text{const(i)}$$

It is like collecting the precipitation across all months without emptying the gauges before end point. Then the requirements of the double mass technique can be ex-

pressed as a demand for for successive calculated $c_{pi}/c_{p.}$'s to be equal to the final ratios $c_{.i}/c_{..}$. This in turn means that

(8) $M_i(p)=(c_{pi}/c_{p.})/(c_{.i}/c_{..})\approx 1$

or

(9) $\log(c_{pi})-\log(c_{p.})-\log(c_{.i})+\log(c_{..})\approx 0$

The structure of (9) is evidently consistent with the tested structure of the residuals of $x_{ti}=\log(y_{ti})$ in (3). The statistical important difference, however is the fact that cutting the series into monthly segments x_{ti} provides us with equal stochastic 'unities' having estimable variances while an analysis of the growing cumulative sequence $c_{p1},\ldots,c_{pk}$ calls for detailled knowledge of precipitation totals summed over all kinds of time periods, eg three- four- seven month totals.

Results.

From data we estimate the α's, $\alpha_i\approx x_{.i}-x_{..}$

ser. no	1	2	3	4	5	6
α_i:	-0.13	-0.13	-0.02	0.07	0.10	0.13

ser. no	7	8	9	10
α_i:	-0.26	0.13	0.02	0.09

ranking series no. 7 with the lowest average precipitation and series no. 8 with the highest.

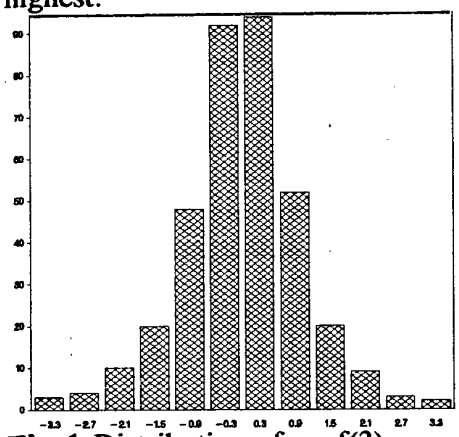


Fig.1 Distribution of r_{ti} cf(3).

Limited space allows us in the sequel to present examples only. Fig 1 shows the distribution of the residuals r_{ti} ,cf (3), indicating that the normal distribution in (1) seems adequate.

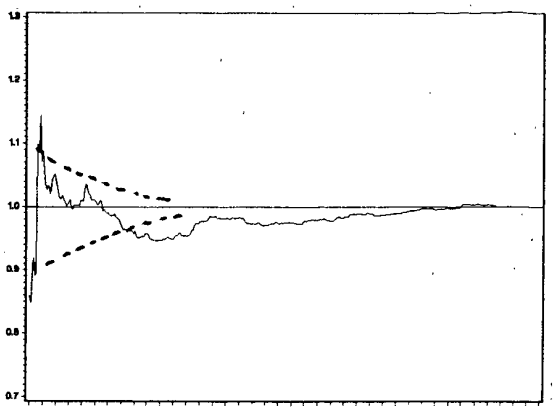


Fig.2. $M_i(p)$ for series no.3.

In fig 2 and 3 the double mass $M_i(p)$'s from (8) are displayed for the two series nos 3 and 10. Tentative confidence bounds are entered into the figures by hand - but it should be emphasized that **no** exact knowledge byond one-year totals is available.

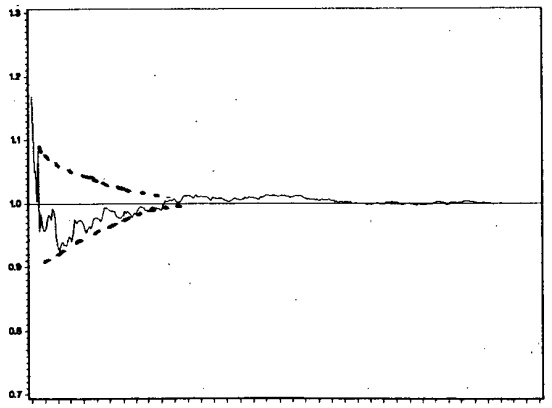


Fig.3 $M_i(p)$ for series no.10.

Fig 4 and 5 displays the result of applying the 1.order check of homogeneity on the series nos 3 and 10, Plot of $T_i(p)$ vs. month no p. The similarity between the structures found in fig nos. 2 and 3 with fig nos 4 and 5 confirms the agreement

between the test statistics $T_i(p)$ and the double mass $M_i(p)$'s . The confidence bounds ± 1.96 in fig nos 4 and 5 refer to the statistical property of the $T_i(p)$'s: for any given value of p $T_i(p)$~normal(0,1).

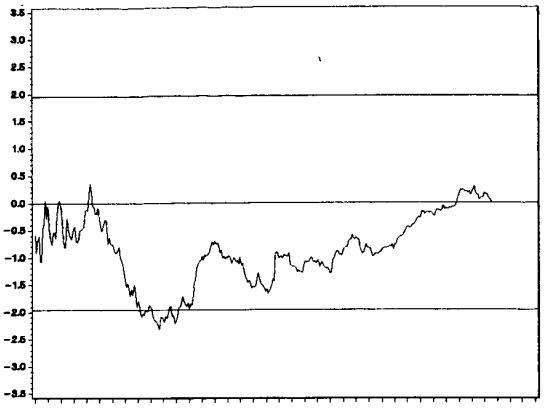


Fig.4 $T_i(p)$ for series no 3.

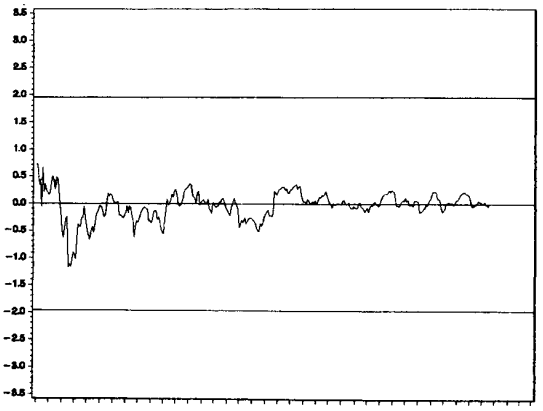


Fig.5 $T_i(p)$ for series no 10.

The test for 2.order homogeneity is displayed in fig 6. The figure contains the 45 possible combinations from the 10 series and a straight line is fitted to the points. The 9 combinations arising from series no 1 has been pointed out by dots indicating a slight bias of this series in comparison with the others.

Discussion.

From the 1.order analysis of homogeneity fig 3 indicates that series no 3 from some .point around month no 45 starts moving too 'slowly' relative to it's own final result - and relative to the movement of the other series. Notice that all $T_i(p)$'s end at zero by definition! A turning point where series no 3 moves upwards again can be read off to be approx. month no 100.What happened to the gauge between p=45 and 100?

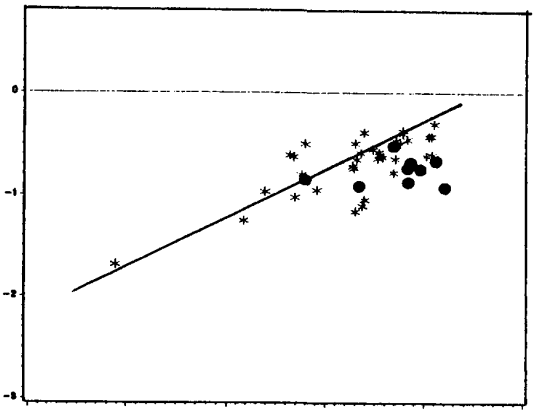


Fig 6 .Test for 2.order homogeneity: $\log(1-\rho_{ij})$ (y axis) vs. log Dist(i,j) (x axis).

From fig 6 we may conclude that series no. 1 on average has too low correlations the other 9 series. The $T_i(p)$ statistics for this series are not extreme but long distances to most of it's neighbors might be the reason for the violation of the isotropic condition for the spatial correlations?

The use of the test statistics $T_i(p)$ has always to be accompanied by other sources of information depending on the type of inhomogeneity, it be a point of discontinuity or a 'slowly moving trend' in the series. However existing automatic algorithms, Alexandersson(1986) for selecting a maximum 'breaking point' could be applied .

References.

Alexandersson,H 1986: A homogeneity test applied to precipitation data. Journ.-Met.Climat.,6.

Allerup,P Madsen,H 1982: Methods of calculating areal precipitation, Nord.Hydrol.,13.

Allerup,P Madsen, H 1983: A statistical model for precipitation data- with application to interpolation problems,Proc. IMSC Lisbon 1983.

STATISTICS FOR PRECIPITATION GAUGE SITE EXPOSURE CHANGES

BORIS SEVRUK AND LUBICA ZAHLAVOVA

Swiss Federal Institute of Technology; ETH Zentrum, Zurich, Switzerland
Slovak Technical University, STU, Bratislava, CSFR

ABSTRACT

The analysis of history of 398 precipitation gauge sites shows that there are on average almost three changes of gauge site exposure during the observation period. Only a small fraction of the total number of gauge sites shows no change and can be used for the climate change studies.

The fact that observed data of meteorological variables can be affected by measuring instrument and site makes the correct detection of climate change using time-series a complex matter, particularly for precipitation, P. Due to the systematic wind-induced error of P measurement, even small changes in construction parameters of P gauge and gauge site exposure, GSE, or gauge installation height can cause serious inhomogeneities of time-series which can be mistaken for P climate change.

Even at present, gauge type and site exchange occurs much more frequently than generally expected, as shown by Sevruk and Klemm (1989). Considering that a change of 5% over a period of a few years is to be expected in the climate change scenarios and that changes of P amount due to inhomogeneities arising from the change in a GSE can match this amount 2 - 3 times, it is clear that the gauge site history should be analyzed to detect possible sources of inhomogeneities. However, the history of gauge sites is as yet not generally available and in some cases not available at all. Such records, if not having been lost or destroyed, are usually stored in the archives of national meteorological services and not accessible from the data bank for direct use. The selection of important events from the archives ought to be made solely by the interested party. This can be very time-consuming. Thus the current practice consists either in taking all time-series available with no regard for inhomogeneities, or sometimes in using statistical tests to check inhomogeneities but without analyzing the station history. Hence the quantity of data is given preference over the quality, with the simple idea that the more time-series included in the analysis, the better will be the balancing of possible errors. It is not necessary to point out that any of these procedures is suitable for detecting climate changes. In addition, even when applying statistical methods to adjust time-series for inhomogeneities, errors affecting the trends can be introduced into the time-series. Moreover, the usefulness of some statistical tests may be in question in cases where the national type of gauge or its installation height has been changed at the same time over the entire territory.

Consequently, only time-series from gauge sites indicating no changes of gauge type and GSE during the period of observation can be used for the climate change analysis. Best suited are protected sites. Here, the effect of wind-induced error and its seasonal variation on the variability of measured P amounts is small.

One of the aims of the presented paper is to show (i) how much the measured P amount can change owing to the change of GSE when the same type of P gauge is used and (ii) how frequent GSE change can occur in a common network, as, for instance, the Swiss one. The ultimate goal, however is to develop a methodology for the selection of suitable and representative time-series for studies of P variability in time and space which would be based instead of on statistics alone, on the physics of phenomenon, that is, on the quantitative evaluation of the effect of GSE changes on wind-induced error. The merits of such a methodology are evident. Firstly, it would enable a more critical view of statistical methods used to detect time-series inhomogeneity, and secondly, it would improve the search for further possible sources of inhomogeneities, such as observational errors or instrument malfunction. A similar approach to deal with inhomogeneities due to exchange of different types of P gauges was presented at the last meeting of statistical climatology by the one of the authors (Sevruk, 1989a).

Basic reference used further is Sevruk (1986a). Additional references can be found in the proceedings of two workshops on precipitation measurement and correction (Sevruk, 1986b, 1989b). For applications see Sevruk and Kirchhofer (1992).

Methods

Essential to the following approach is the assessment of GSE and its changes as based on the analysis of the history records consisting of photos, sketches and written reports of gauge sites. Four classes of GSE have been used. They are characterized by the average vertical angle, α of obstacles around the gauge, as measured in eight directions of wind rose: Class 1 - open site ($\alpha = 0 - 5^o$); class 2 - partly open ($\alpha = 6 - 12^o$); class 3 - partly protected ($\alpha = 13 - 20^o$) and class 4 - protected site ($\alpha = 21 - 27^o$). Classes of 1.5, 2.5 and 3.5 have been used regularly. The respective values of α are 6, 13 and 18°. Figure 1 shows how α is defined and Sevruk (1986a) discusses its applications to estimate the wind-induced loss.

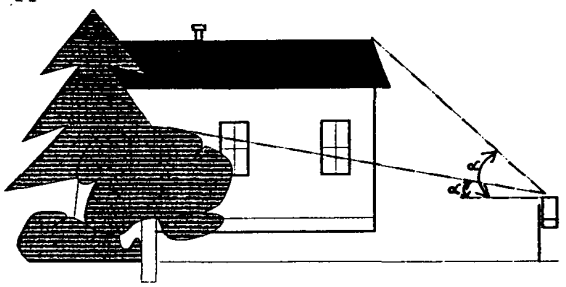


Fig. 1 Vertical angle ,α, of obstacles.

From the comparison of GSE classes as based on the measured values of α in the field and as assessed, the average error of GSE class estimates was computed to be considerably smaller than one class. Consequently, the chronology of GSE class changes can be established with fairly good accuracy. This was done by an experienced person during four months work for 398 gauge sites, in some cases as far back as the first half of the last century. The gauge sites are uniformly distributed over the territory of Switzerland in an altitude range from 198 to 2540 m a.s.l. (mean altitude is 760 m). The average record length is 80 years, spanning 10 to 173 years.

The frequency of changes in GSE classes were analyzed for different elevation zones and two time periods, the 90-year period, 1901-1990, and 50-year period, 1941-1990, and after 30 and 60 years. The aim was to see if there is any substantial and systematic change in GSE classes due to different time periods, altitude and urbanization effects.

To assess the effect of a change in GSE classes on the magnitude of wind-induced error, the latter was estimated in these region types: Less windy and windy lowlands and mountains. The former are situated in altitudes of 600 m a.s.l. and the latter up to 2000 m a.s.l. The portion of snow in total P in the lowlands is null for the summer season from April to September and 30% for the winter season from October to March. The respective figures for the mountains are 25 and 80%. Less windy GSE indicates wind speed at the level of 10 m above ground of 2 m s^{-1} for lowlands and 3 m s^{-1} for the mountains. The respective figures are 4 and 6 m s^{-1} for windy GSE.

The estimation of wind-induced error was made for the Hellmann gauge without wind shielding device (wind shield), according to Sevruk (1986a). This gauge type is used in Switzerland as the national standard.

Table 1 Percentage [1] of wind-induced losses for four exposure classes in less windy and windy lowland and mountainous locations, for the summer (April-September) and winter (October-March) seasons, according to Sevruk (1986a) .

Region / Season	Exposure class 1*		2*		3*		4*	
	less windy	windy	less windy	windy	less windy	windy	less windy	windy
Lowlands								
Summer	3	5	2	4	1	2	1	2
Winter	8	18	6	14	5	10	3	7
Mountains								
Summer	11	25	8	19	6	15	5	10
Winter	27	59	22	50	16	36	12	26

1) relates to measured precipitation amount *α = 2, 9, 16 and 24° respectively

Note: The values do not include wetting and evaporation losses and are smaller than the systematic error of precipitation measurement which includes all error components (Sevruk, 1986a).

Results

Table 1 contains wind-induced losses for all GSE classes in percent of measured P: It shows that with the exception of both lowlands in the summer season and partly classes 4, the losses are greater than 5%.

If a change in precipitation amount of ± 3% is taken as a limit, then with the exception of the summer season in the lowlands, a change in GSE by two classes can result, in a significant variability of measured P amounts, up to 10% in the lowlands and between 3 and 25% in the mountains, depending on the altitude, wind and P conditions and the season. If changes between less windy and windy GSE are also taken into consideration then a change in GSE of one class, or even within the same class, can produce a considerable increase or decrease of measured P, up to 30%. (For an example from the Swiss prealps see Sevruk, 1987).

On average, there are slightly less than three GSE class changes per gauge site, that is one change each 26-27 years, and they are more frequent in the altitudes below 800 m a.s.l. Almost 60 sites showed at least four changes. Only two sites showed 10 and more changes. As shown in Figure 2, the number of changes increases with the observation period length. The increase accelerates for periods longer than 100 years.

The most frequent exposure change was one class and more (130 out of 398). For 90 sites resulted a change of one-half of a class. A minimum two-class jump appeared for 23 sites.

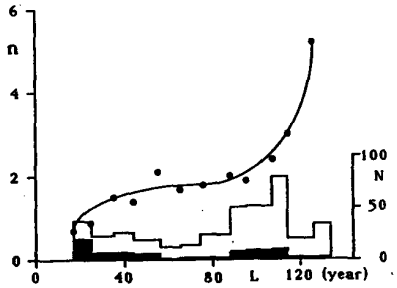


Fig. 2 Dependency of the number ,n, of changes in exposure class per gauge site on the length of observation period (points). Average values. N on the right scale indicates the total number of gauge sites (void columns) and the number of gauge sites without change (black columns).

The distribution of exposure classes for different altitude zones is shown in Figure 3. Below 2000 m a.s.l. the most frequent classes are 2 and 2.5 but the portion of classes 1 and 1.5 increases rapidly with increasing altitude, so that above 2000 m a.s.l. almost all gauge sites are in this class.

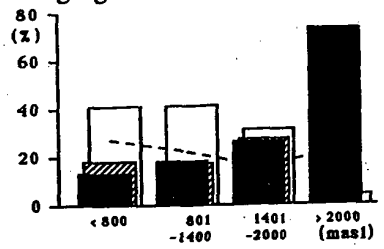


Fig. 3 Frequency ,in (%), of exposure classes for four altitude zones. The black column indicates classes 1, 1.5, the void one 2, 2.5 and the hatched one 3 - 4. The dashed line indicates no documented gauge sites.

The changes in exposure classes after 30 and 60 years are shown in Figure 4. Fortunately, the number of gauge sites with missing information on GSE classes has been decreasing rapidly during the past years, indicating a more concerned attitude of authorities to the problem of GSE documentation. The continual increase of GSE classes of 1 and 1.5 throughout this century is probably due to the fact that some of the new gauge sites were situated out of vilages and in higher altitudes.

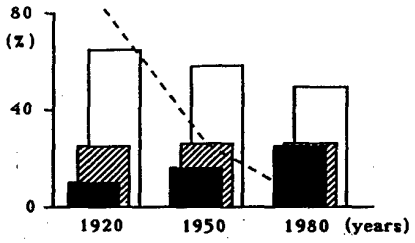


Fig. 4 Change in frequency ,in (%), of exposure classes after 30 and 60 years in percent of total number of documented gauge sites. The black column indicates classes 1, 1.5, the hatched one classes 3 - 4, and the void one classes 2, 2.5. The dashed line shows no documented gauge sites.

The total number of gauge sites where no change appeared is 69 (out of 398). Their distribution according to the length of observation period is shown in Figure 2. However, 19 gauge sites have to be eliminated owing to unreliable data during a certain time period (gauges too near to obstacles or fixed on a balustrade, missing information on exposure class, insufficient work of observers, etc.). If the 90-year period, 1901-1990, is considered, the number of gauge sites without change decreases to 19 and for the 50-year period, 1941-1990, is 23 gauge sites. Yet only five to six

gauge sites have the class of 3 - 4 as shown in Figure 5. Of course, not all changes mean a change in GSE, for instance, when the gauge is moved to a new site which has the same exposure class or when the GSE change is a half- class only. Consequently, the number of sites where no change or only a small change of exposure class is highly probable increases to 48 for the 90-year period and to 57 for the 50-year period.

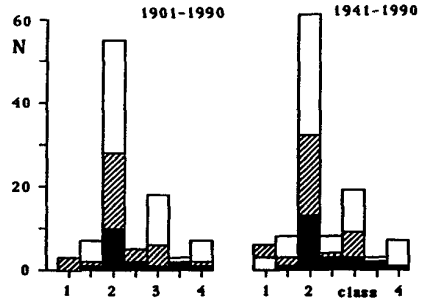


Fig. 5 Number ‚N, of gauge sites subdivided according to the exposure class for two time periods 1901 - 1990 and 1941 - 1990. Black columns indicate gauge sites without change of exposure class. Void columns indicate both gauge sites without change and with small changes in exposure classes and hatched columns are gauge sites with unreliable data.

Discussion

It seems that the number of suitable gauge sites is rather small for the assessment of the distribution of P in time and space over the territory of Switzerland. In any case, further checking of selected time-series is necessary, particularly for the effect of the change of observers and of installation height of P gauges above ground on P measurements. The regular height in Switzerland is 1.5 m but it varies from 0.9 to 2.7 m and it changes frequently as well. Here, the statistical tests can be applied with success. In addition to the points mentioned above, two serious changes of instrumentation took place in the Swiss P measurement network during the last hundred years. At the end of the last century, the old Swiss gauge was replaced by a new one, that is, the German Hellmann gauge which was used until recently as the Swiss national standard gauge. Ten years ago, the Hellmann gauge was replaced by the automatic, heating, tipping-bucket gauge at 70 gauge sites. On this occasion, the new gauges were moved to exposed sites and show a considerable deficit (Sevruk 1989a).

Conclusions

The analysis of precipitation gauge site history is an important part of studies of climate change. It contains not only the necessary information on which the studies of inhomogeneities and variability in time and trends of precipitation are based but it is the fundamental assumption for the efficient use of statistics in precipitation studies.

Acknowledgement

Thanks are due to the Swiss Meteorological Institute for the permission to use rooms and archive materials.

References

Sevruk, B., 1986a :Correction of precipitation measurements: Swiss experience. In: B. Sevruk (Ed.) Correction of precipitation measurements, Swiss Federal Institute of Technology, E.T.H. Zürcher geographische Schriften, ETHZ, No. 23, 187-196.

Sevruk, B. (Ed.), 1986b: Correction of precipitation measurement. Swiss Federal Institute of Technology, ETHZ Zürcher geographische Schriften, No. 23, 288pp.

Sevruk, B., 1987: Point precipitation measurements: why are they not corrected. In: J.C. Rodda and N.C.Matalas (Eds). Water for the Future: Hydrology in Perspective. Internat. Assoc. Hydrol. Scie. publ. No. 164, 477-486.

Sevruk, B., 1989a: Inhomogeneities in precipitation time-series. In: J. Sanson (Ed.). Proc. 4th International Meeting on Statistical Climatology, Wellington, New Zealand, 24-27.

Sevruk, B., (Ed.) 1989: Precipitation measurement. Proc. International Workshop on Precipitation Measurement, St. Moritz. Swiss Federal Institute of Technology, ETH Zurich, 584 pp.

Sevruk, B. and S. Klemm, 1989: Catalogue of national standard precipitation gauges. WMO Instruments and Observing Methods Rep., No. 39, WMO/TD-No. 313, 52 p.

Sevruk, B. and W. Kirchhofer, 1992: Mean annual corrections of measured precipitation amounts 1951-1980. In: M. Spreafico and R. Weingartner (Eds.) Hydrologischer Atlas der Schweiz, Kartenblatt 2.3, Landestopographie, Bern.

THE DEVELOPMENT OF A LONG-TERM MONTHLY WIND SPEED RECORD AT SABLE ISLAND

J.D. Smith[1], R.D. Brown[2], V.R. Swail[2] and K. Pietrzak-Smith[1]

1. Introduction

Several studies over the past few years have attempted to characterize the climate variability over the oceans in an effort to understand global climate change (e.g. Cardone, et al., 1990; Franzen, 1991; Palutikof, et al., 1986; Ramage, 1987; and Wright, 1986). Characterization of climatic variability is important for understanding the various mechanisms influencing a climatic variable, and for assessing the adequacy of stochastic parameters such as means and return period design values. The problem with long-term characterization of wind speed variability is that it is greatly complicated by frequent changes in instrumentation and observing practices; furthermore wind speed statistics are sensitive to site changes. The purpose of this brief note is to outline, by example, some of the problems that were encountered while developing a long-term wind speed dataset from the original climate station reports.

The example of Sable Island, N.S. was chosen because of its truly offshore nature (see Fig. 1) and because of its proximity to the main hydrocarbon and fishery resources on the Scotian Shelf. This report is part of a larger study conducted by the Atmospheric Environment Service to characterize the historical over-water winds of the Scotian Shelf.

2. Sable Island Wind Observing Program

a. Recording Practice

The first meteorological observations were probably taken on Sable Island from the East Point site in 1871. Sable Island first appears in the Annual Reports of the Meteorological Service of Canada in 1883. In 1890, the West or Main Station was opened. In 1895, both stations were indicated as being Class II with three observations per day at fixed local times - 8:00 am, 2:00 pm, and 8:00 pm.

In this early period, East Point appeared to be the more important site (e.g. in the 1895 Annual Report, the Summary Tables are presented for East Point). However, early in 1906, the Main Station was upgraded to class IT which meant that observations taken at 8:00 am and 8:00 pm 75th meridian time were telegraphed to Toronto immediately after the reading. The three observations per day format (8:00 am, 2:00 pm and 8:00 pm 75th meridian time) was re-introduced in 1909. In the first "Monthly Record" published in 1916, East Point was shown as a class II station which was defined as "a station where the equipment consists of a maximum and a minimum thermometer and a rain gauge, ordinarily, although at a few the equipment is more extensive". It appears that during this period, the early reliance on East Point changed when the telegraph was installed at the Main Station; the East Point site then became less important, and ceased to operate in 1930.

In the first Monthly Record of Meteorological Observations published for 1916, the three local times shown were 9:00 am, 2:00 pm and 9:00 pm Between the years 1922 and 1928 the mean wind speeds recorded at the Main Station were reported as *either* wind speed in miles per hour (mph) *or* Beaufort Scale Force with no apparent pattern; units varied even from month to month. In 1929 the Main Station observations were taken only at 8:00 am and 8:00 pm and in 1935 these observations began to be recorded as monthly wind run rather than as mean wind speed.

b. Instrumentation

In the 1895 Annual Climate Report, it was stated that the anemometer in general use was a modification of that by Green of New York. This anemometer was connected to either a mechanical or electrical anemograph which recorded hourly wind run and direction. In 1919, John Patterson (1920), physicist to the Meteorological Service of Canada, presented a paper to the Royal Society of Canada on the correction factor for the Canadian standard anemometer in which he describes the Canadian standard anemometer as being "a [4 cup] Robinson anemometer with cups 4 inches in diameter on arms 6.72 inches long".

In a comparison with a Dines' Pressure Tube Anemometer, Patterson found that the 4 cup anemometer over-estimated wind speed, particularly for higher velocities. Patterson proposed the following

[1] The MEP Company, Markham, Canada
[2] Canadian Climate Centre, Toronto, Canada

correction based on experimental data obtained in a wind tunnel with wind speeds up to about 35 mph:

$$V_c = (1.24 - 0.251 \log V_r)\, V_r \qquad (1)$$

where

V_c = "true" velocity
V_r = recorded velocity

Patterson indicated that the correction should only be applied for $V_r > 10$ mph, and that the correction was an extrapolation for wind speed exceeding 35 mph. There is no evidence that this correction has ever been applied to the observed wind data.

The historical evidence for wind instrumentation on Sable Island is summarized in Table 1. It is not known when this anemometer was first installed; however, Patterson indicated that the Robinson anemometer was designed in 1846, and was in extensive use. Subsequent experiments (beginning in 1875) to calibrate the system were performed in England, Canada, U.S.A., and Russia which demonstrates the wide-spread use of the instrument.

In a 1926 paper to the Royal Society of Canada, Patterson presented results of an investigation into anemometer design and concluded that the three cup anemometer is decidedly superior to the four cup. This then became the 45B anemometer which came into regular service during the mid-1930's. A 45B replaced the official anemometer at Sable Island sometime in 1935-37. However, photographs taken in 1931 clearly show a 45B mounted on a 10 ft. wooden mast near the beach, in addition to the main anemometer mounted on a wooden tower on top of the superintendant's house. (Perhaps this was for field comparison purposes).

Table 1. Summary of Wind Instrumentation

- 1899: Photo taken of the superintendant's house in 1899 shows an anemometer on top of wooden mast attached to roof. The anemometer height is estimated to be ~24 ft. above ground level. The superintendant's house was apparently built in 1899 and lasted until 1950.
- 1917: A photo of the main station from Canada Archives shows an anemometer mounted on slightly taller mast at the same location on the superintendant's house. Height is estimated to be ~26 ft. above ground level. A third floor loft had been added to the north side of the building which would cause some disturbance to winds from the northeast quadrant.
- 1931-July 9, 1937: A photo in the inspection report file taken in 1931 showed a wooden tower has replaced the mast on the superintendant's house. The anemometer height is shown on the photo as 39 ft. The anemometer was likely to have been a 4 cup Robinson. Another photo taken in 1931 shows a 45B mounted on a smaller 10 ft. wooden mast near the beach.
- July 9, 1937: A new steel tower was erected. Anemometer height indicated at 37 ft. above ground. The station history file assumes a 45B is also installed at this time. This is unclear from the inspection reports and may well have been earlier.
- August 14-25, 1959: The inspection report indicates no wind measuring equipment in use (no reason given). The inspection report of August 26-29, 1959 indicates the system is operational.
- June 4, 1968, 1200Z: The anemometer tower is relocated on new 10 m mast.
- May 9, 1969: U2A replaces the 45B.

c. Observers

Between the years 1909 and 1942 there have been no fewer than 8 different observers stationed on Sable Island. Each change of observers introduces a systematic change in observing practice which in many cases can be quite substantial.

The value of research into changes in observers and observing practices cannot be over-emphasized. For instance, the meteorological inspection report for Sable Island dated September 5, 1936 stated that:

"The anemograph does not work too well especially after heavy all day rain as anemovane cover is rusted and leaks and short circuits interfere with operation".

3. Wind Record

Prior to this study, the wind data for Sable Island, existed only as paper climate charts at the Canadian Climate Centre for the years 1895-1952. After 1952, these data were recorded as hourly observations in the Canadian Digital Archive of the Canadian Climate Centre.

A digital database of raw observed values of wind speed and direction (as well as pressure) has been constructed in this study from the monthly climate reports for the years 1895-1952. This digital dataset (1895-1989) includes: monthly mean wind speed, monthly frequency of wind direction, extreme wind speed and direction, the number of gales per month, and the mean sea-level pressure at Sable

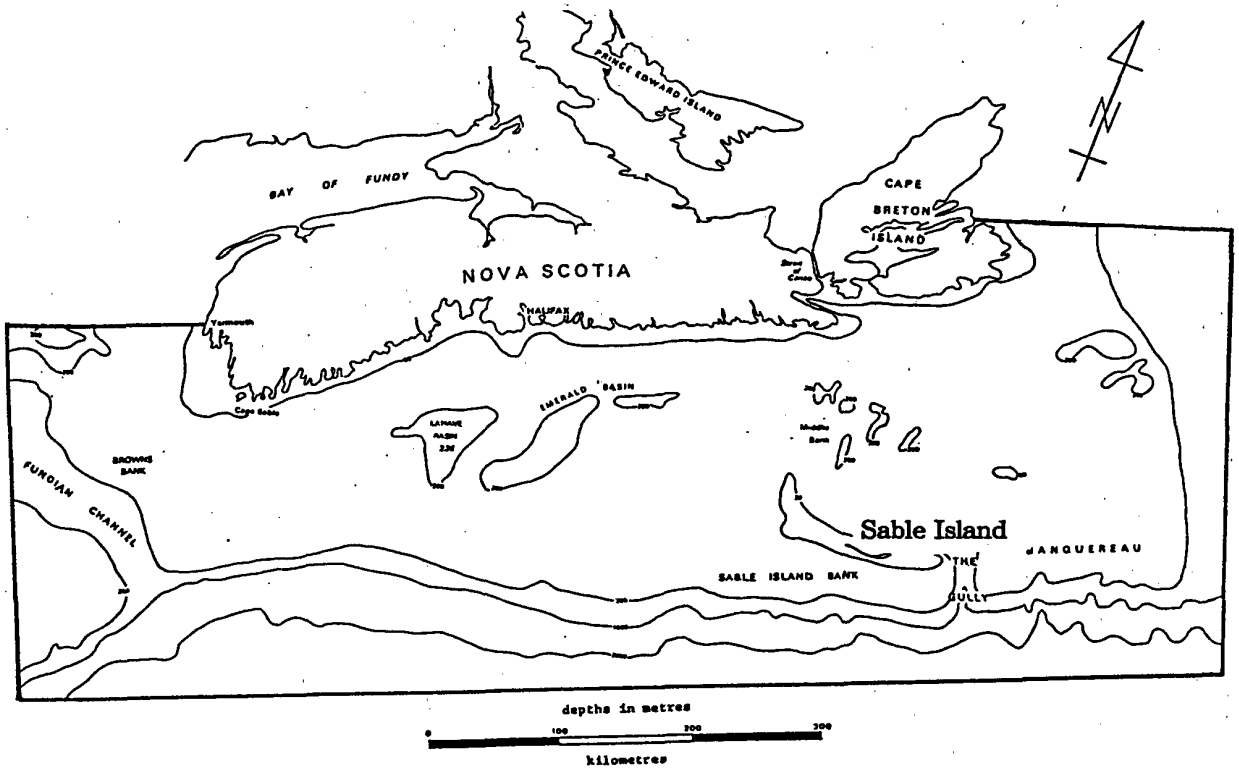


Fig. 1 Location of Sable Island

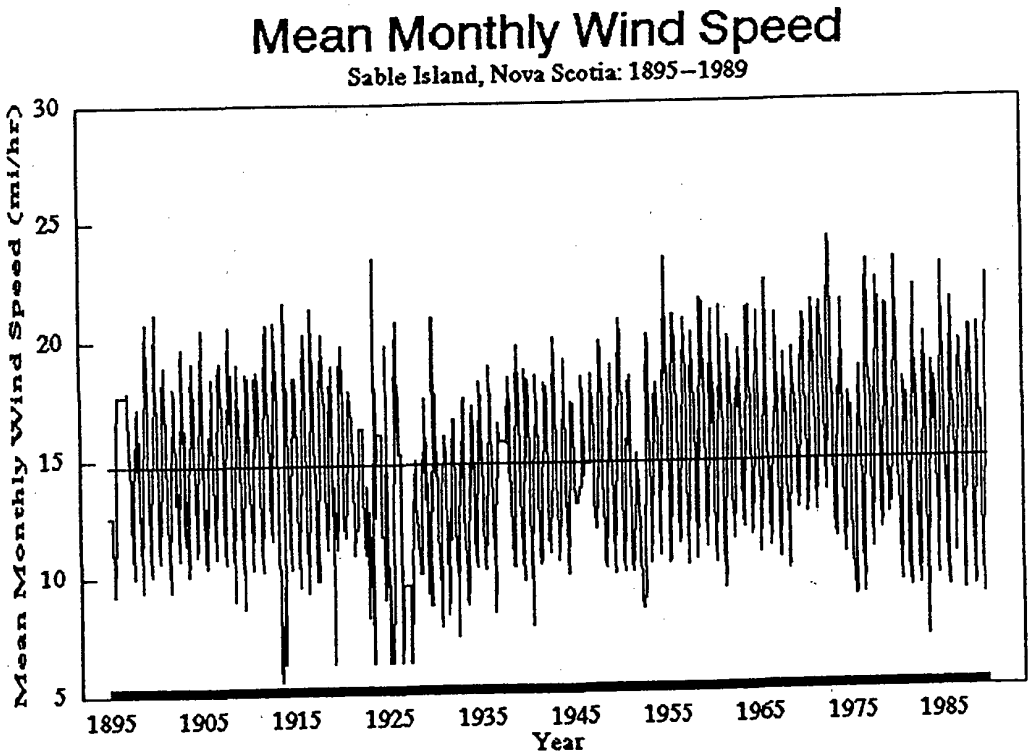


Fig. 2 Mean Monthly Wind Speeds at Sable Island 1895-1989

Island, Sydney and Yarmouth. An example of the information which can be assembled from this longterm dataset is presented in Fig. 2. It appears from Fig. 2 that there is a marked change in the wind regime during the 1920's. Unfortunately, this period coincides with the greatest frequency of change in observing practice. Without further analysis (such as correlation with pressure and temperature data) it is not known whether we are observing a climatic phenomenon or not.

The mean wind speed values (Fig. 2) have been corrected for changes in instrumentation and for record units: (a) equation (1) was used in an effort to make the anemometer data prior to May 9, 1969, compatible with the U2A anemometer data after May 9, 1969; (b) the wind speeds recorded using the Beaufort Scale were converted to mph using the method of Cardone (1969). It should be noted that there are significant problems associated with estimating wind speed for the Beaufort Scale (e.g. Frost, 1966; Isemer and Hasse, 1991; Kinsman, 1990 and 1991; and Peterson and Hasse, 1982). This study will also examine whether the changes in the observing practice and in instrumentation introduced any significant systematic trends in the longterm dataset.

The dataset will then be compared with ship observations within 100 km of Sable Island taken during the same period. This will determine whether the two datasets are homogeneous and whether over-water winds for the Scotian Shelf can be estimated.

4. Conclusion

It has been demonstrated that in the development of long-term wind records using historical data, there are many considerations that have to be addressed in order to minimize systematic errors in the datasets. These include changes in: instrumentation, site location and height, observation practice, units of observation and observers.

5. Acknowledgements

We would like to thank Dr. D. McGillivray and Mr. Gord McKay for their many helpful comments on this report. This research was funded by the Atmospheric Environment Service.

References

Cardone, V.J., J.G. Greenwood and M.A. Cane, 1990: On Trends in Historical Marine Wind Data. *Journ. of Clim.*, 3,113-127.

Cardone, V.J., 1969: Specification of the wind distribution in the marine boundary layer for wave forecasting. Report TR69-1, New York University, New York, N.Y., 131 pp. [NTISAD 702-490].

Franzen, L.G., 1991: The Changing Frequency of Gales on the Swedish West Coast and its Possible Relation to the Increased Damage on Coniferous Forests of Southern Sweden. *Inter. Journal of Climatology,* 11, 769-793.

Frost, R., 1966: The Relation Between Beaufort Force Wind Speed and Wave Height. British Meteorological Office, Scientific Paper No. 25.

Isemer, H.-J., and L. Hasse, 1991: The Scientific Beaufort Equivalent Scale: Effects on Wind Statistics and Climatological Air-Sea Flux Estimates in the North Atlantic Ocean. *Journal of Climate,* 4, 819-836.

Kinsman, B., 1990: Who Put the Wind Speeds in Admiral Beaufort's Force Scale? Part I-The Original Scale. *Mariner's Weather Log,* 34, 2-8.

----, 1991: Who Put the Wind Speeds in Admiral Beaufort's Force Scale? Part II-The New Scales. *Mariner's Weather Log,* 35, 12-18.

Palutikof, P., P. Kelly and J. Davies, 1986: Windspeed Variations and Climate Change. *Wind Eng.*, 10, 182-189.

Patterson, J., 1920: The Correction Factor of the Canadian Standard Anemometer and the most Probable Maximum Velocity and the Possible Extreme Velocity in Gusts. *Proceedings and Transactions of the Royal Society of Canada,* Series III, 18, 81-96.

----, 1926: The Cup Anemometer. *Proceedings and Transactions of the Royal Society of Canada,* Series III, 20, 1-54.

Peterson, E.W. and L. Hasse, 1982: Did the Beaufort Scale or the Wind Climate Change? *Journal of Physical Oceanography,* 17, 1071-1074.

Ramage, C.S., 1987: Secular Change in Reported Surface Wind Speeds Over the Ocean. *Journal of Climate and Applied Meteorology,* 26, 525-528.

Wright, P.B., 1986: Problems in the Use of Ship Observations for the Study of Interdecadal Climate Change. *Monthly Weather Review,* 114, 1028-1034.

A STATISTICAL ANALYSIS OF INVERCARGILL UV DATA

X. Zheng and R. Basher
New Zealand Meteorological Service, Wellington.

Introduction

High rates of skin cancer and premature materials failure in New Zealand [Basher (1981)] have led to popular speculation that the ultraviolet (UV) climate in New Zealand is relatively harsh. For monitoring UV climate, the New Zealand Meteorological Service installed a Robertson-Berger (RB) sunburn meter [Robertson, (1969) and Berger, (1976)] in 1981 at Invercargill. The meter failed in 1986 and was sent to the US manufacturer for repair. On return at Invercargill in 1988 it was found to read too high by about 60% relative to pre-1986 conditions and it was electronically adjusted to correct this change [Kenny and Basher (1990)]. It has not been possible to arrange a calibration against the international standard instrument. In its raw form the data set cannot be used for the detection of trends in UV radiation. In this paper, we show how we produce a homogeneous UV data set. The key approach is to compare the measured UV data with the UV estimates from a well established physical UV transmission model [Green et al. (1974) modified by Basher (1987)]. We then show that there is a long term upwards trend in UV at Invercargill and that it is probably due to the measured long term ozone depletion of about 5% per decade for this location.

Physical and statistical UV transmission model

The physical UV transmission model is a non-linear spectral model which needs ozone data input. Using satellite derived ozone data, we are able to calculate the theoretical daily UV level. We refer to as calculated UV and denote it as UVc. An alternative statistics model [Barton and Robertson (1970)] is to assume that the fractional change in UV can be linearly related to the fractional change in ozone

$$(1) \qquad \frac{UV - UV_f}{UV_f} = a + b\frac{Oz - Oz_f}{Oz_f} + \varepsilon$$

where the UV values are clear sky daily totals, UV_f and Oz_f are climatological references for the annual variation of daily UV and ozone respectively and ε is a family of i.i.d. random errors with mean 0. It is important to choose the most suitable references here. We obtain our ozone reference by locally robust regression of the clear-sky satellite data using function "lowess" in the statistics analysis software "S-plus". Since UVc is a function of ozone, the best reference for UVc is from the physical UV transmission model using the ozone reference Oz_f as input. We refer to it as UVc_f. In the case of the measured UV we do not know exactly how it relates to ozone, so we choose its reference UV_f as $UVc_f*(UVr_f/UVcr_f)$ where UVr_f and $UVcr_f$ are references for measured UV and calculated UV by robust regression respectively. $UVr_f/UVcr_f$ represents the ratio of measured UV seasonal variation to calculated UV seasonal variation. The correlations between UV fractional changes and ozone fractional changes for several data sets and references are listed in Table 1. We could choose UVr_f as a reference for measured UV as Barton and Robertson (1970) did, but we demonstrate UV_f is a better choice because it makes better physical sense and shows a stronger linear relationship in Table 1. Table 1 also shows that the calculated UV fits the statistics model very well. This demonstrates the close relationship between the two models.

Table 1. Correlations between UV fractional changes and ozone fractional changes

	1981-86	1988-90	1981-90
Calculated UV & $UVcr_f$	-0.910	-0.920	-0.918
Calculated UV & $UVcr_f$	-0.988	-0.976	-0.982
Measured UV & UVr_f	-0.725	-0.917	-0.785
Measured UV & UV_f	-0.835	-0.912	-0.842

Homogenising the measured UV data set

The changed 1988-90 RB meter sensitivity results in a constant factor change for the 1988-90 measured UV, say k_m. It also results in a constant factor change in the measured UV reference UV_f, say k_r, i.e. $UV1_f = k_r\, UV_f$ where $UV1_f$ comes from the same procedure as UV_f under the assumption k_m=1. Suppose the regression equation for 1981-86 data is

$$(2) \qquad \frac{UV - UV1_f}{UV1_f} = a_1 + b_1 \frac{Oz - Oz_f}{Oz_f} + \varepsilon$$

This is equivalent to the regression equation

$$(3) \qquad \frac{UV - UV_f}{UV_f} = k_r\, a_1 - 1 + k_r + k_r\, b_1 \frac{Oz - Oz_f}{Oz_f} + \varepsilon$$

where $(k_r\, a_1 - 1 + k_r)$ and $k_r\, b_1$ can be estimated. Then the adjusted consistent 1988-90 UV data, k_m*UV, should satisfy the regression equation

$$(4) \qquad \frac{k_m{*}UV - UV_f}{UV_f} = k_r\, a_1 - 1 + k_r + k_r\, b_1 \frac{Oz - Oz_f}{Oz_f} + \varepsilon$$

By the least square estimation procedure, we get

$$(5) \qquad k_m = \frac{\sum_1^{n2} ((k_r\, a_1 - 1 + k_r) + 1 + k_r\, b_1\, Oz / Oz_f)\, UV / UV_f}{\sum_1^{n2} (UV/UV_f)^2}$$

$$= \frac{\sum_1^{n2} (a_1 + 1 + b_1\, Oz / Oz_f)\, UV / UV1_f}{\sum_1^{n2} (UV / UV1_f)^2} = 0.945$$

where n_2 is the number of 1988-90 clear sky days. The crucial point here is that this k_m is independent of k_r theoretically. If it were dependent on k_r, then we would have achieved nothing because k_r depends on the 1988-90 RB meter sensitivity. We call this procedure the single sensitivity adjustment.

Figure 1 shows some interesting seasonal variation in the relation between calculated and measured UV. These are probably due to systematic errors in the instrument and the Green model, e.g. with solar angle. We can calculate seasonal adjustments further by taking k_s and k_w as the adjustment factors for the summer half year and the winter half year 1988-90 measured UV records respectively. We refer to n_{s2} and n_{w2} as the number of clear sky days

for the 1988-90 summer half year and winter half year respectively and similarly refer to n_{s1} and n_{w1} for the 1981-86 period. Then we refer to m_2, m_{s2} and m_{w2} as the ratio of the single adjusted UV against the calculated UV for the 1988-90 full year, summer half year and winter half year respectively, and similarly refer to m_1, m_{s1} and m_{w1} for the 1981-86 period. To keep the same overall change as the single sensitivity adjustment, we must impose the condition

(6) $$(1 - k_w m_{w2} / m_2) n_{w2} = (k_s m_{s2} / m_2 - 1) n_{s2}$$

Then we estimate the k_s and k_w by minimising

(7) $$(k_s m_{s2} / m_2 - m_{s1} / m_1)^2 + (k_w m_{w2} / m_2 - m_{w1} / m_1)^2$$

under the condition (6). This gives k_s=1.033 and k_w=0.966.

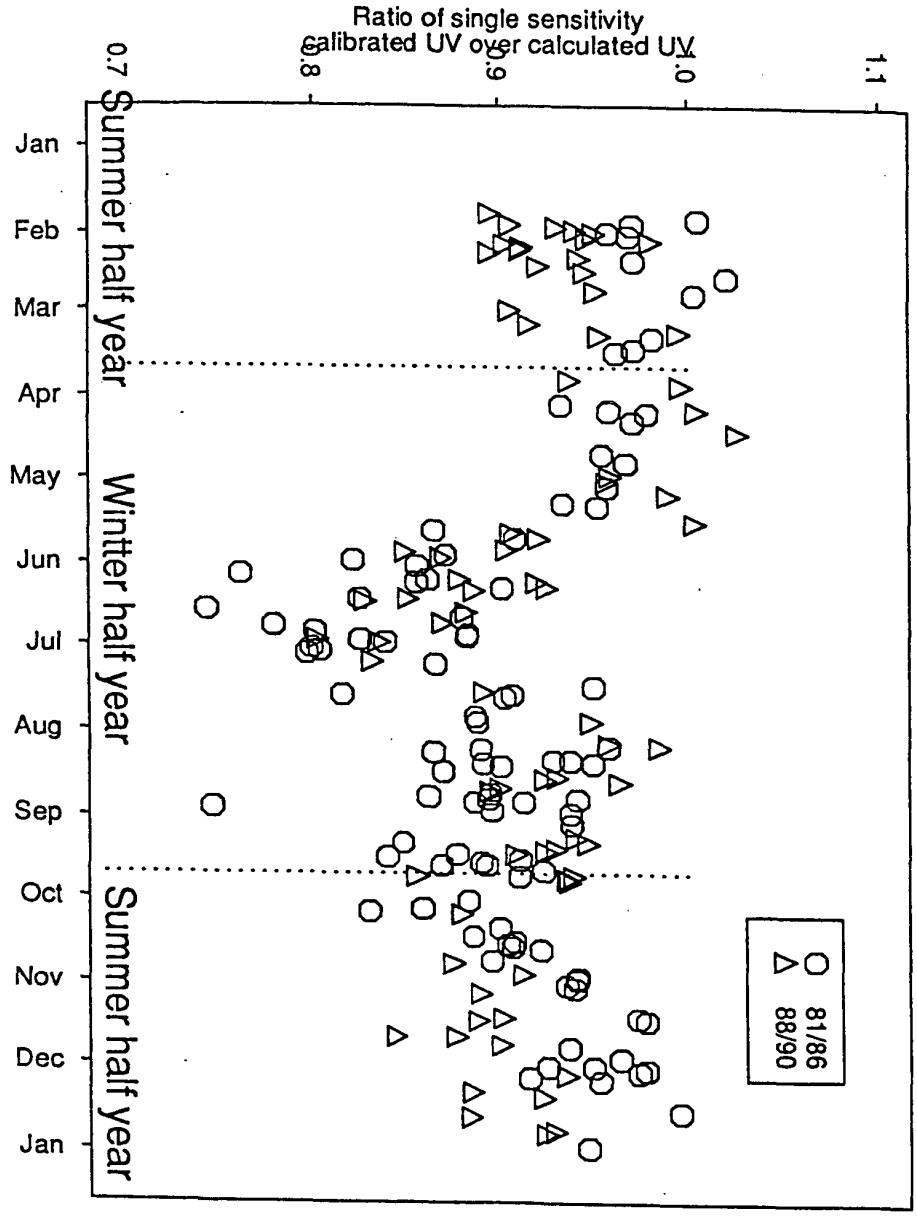
The difficulty with the two-season adjustment is that we do not have a physical basis for treating the two data periods differently, and it will introduce artifacts into the dataset at the seasonal change. In our final adjustment we have chosen the weighted regression procedure giving weight function $1/UV2_f^2$ to regression equation 3, because days recorded higher UV levels are more important. This adjustment factor is 0.99. Residual analysis shows that the seasonal adjustment gives the best fit. Although we did not use it, it still reflects some natural feature of the original UV data.

Trend analysis

We detect the trends for adjusted UV data, ozone data and the residuals of equation 3 separately by weighted regression giving weight function $1/UV_f^2$. It shows that the UV radiation has a upwards trend of about 6% /decade and ozone has a downwards trend of about 5% /decade at the 95% level. The trend for the residuals is not significant. In effect, we have used the physical UV model together with the statistics model of the UV/ozone relation derived from day to day variations, to test the consistency of the UV/ozone long term trend relationship. This strongly suggests, as expected, that UV levels are increasing in direct relation to the ozone decline.

References

Barton, I, and D Robertson (1979). Measurements of erythemally effective ultraviolet radiation and total ozone content. Nature 258, 6, 68-69.
Basher, R. (Ed.) (1981). Proceedings of the seminar on solar ultraviolet radiation, Wellington. New Zealand Meteorological Service, Wellington.
Basher, R. (1987). The calculation of solar ultraviolet radiation and its sunburning effect under clear skies. Scientific report 26, New Zealand Meteorological Service.
Berger, D. (1976). The sunburning ultraviolet meter: design and performance. Photochem. Photobiol., 35, 187-192.
Green, A, T Sawada, and E Shettle (1974). The middle ultraviolet reaching the ground. Photochem. Photobiol., 19, 251-259.
Kenny, G. and R. Basher (1990). Report on Invercargill UV data progress to 30 July 1990. Scientific Report, New Zealand Meteorological Service.
Nichol, S. and R. Basher (1987). Sunburning ultraviolet radiation at Invercargill, New Zealand. Weather and Climate (Journal of the Meteorological Society of NZ), 7, 21-25.
Robertson, D. (1969). Long-term field measurements of erythemally effective natural ultraviolet radiation. Biologic Effects of UV Radiation, (Edited by Urbach F.), 433-436, Pergamon Press, Oxford.


Ratio of single sensitivity
calibrated UV over calculated UV
0.7
0.8
0.9
1.0
1.1
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Jan
Summer half year
Winter half year
Summer half year
81/86
88/90

CONFRONTED BY CLICOM

K.J.A. Revfeim
Meteorological Service, Nadi Airport, Fiji

Via a 1983 circular letter to member countries WMO solicited interest in the development of a pc based climate data storage and processing system. A statement of primary objectives viz:-

* to promote the development .....
* to compile and disseminate .....
* to co-ordinate the exchange .....
* to co-ordinate education .....,

was followed by the most important objectives i.e. "The project is primarily directed at developing countries ....." and "Emphasis is placed on easy to use software ....". Specifications dealt with 'the system' from the entry point of 'data assimilation' and did not express any need to be particularly concerned about the nature of data. From that point on everything in the specifications assumed an acceptable if not ideal data set for subsequent processing; almost too good to be true!

It would seem from elements allocated numbers 1,2,3,4,5.... that daily observations were considered first and doubtless the identified target market was non-computerised climate centres with a modest network of stations. However it was too tempting not to generalise to all types of data, from synoptic records to specialised research interests, and put everything in the one 'package'. The objectives were written in that disconcerting 'international project' language wheras one suspects that if its prime users had been asked they might have stated

1. simple to learn the system basics
2. rapid start to key-entry of data
3. accommodates the flaws of real data
4. easy extraction of useful products

Anyway this would surely have been addressed by the committee set up to oversee the Clicom project.

Did any of the system designers spend some time in a 'target' country observing how records were kept and how manual processing was done for routine monthly climate returns? Or was a lot left to the memories of experts and downsizing of the procedures in major climate data processing centres?

Clicom as delivered seems to have acquired more ambitious objectives that go well beyond the simple system that might have satisfied more than 90% of the target market:-

1. * Learning the basics involves tracing through additional options to monthly/daily observations for 3 hourly (synoptic), hourly, 15 minute, and upper air (pressure level increment) data;
   * About 300 element codes that appear to have no structure within (and little between) the time/level scales above;
   * Nearly 1000 files to store programmes, instructions and data is a daunting introduction rather than promotional 'plus'.
2. * Before key-entry of any data, information must be loaded for each station on GEOGRAPHY, OBSERVATION (frequency), ELEMENTS (observed), and EXTREMES (for each element and month) i.e. 1 station with daily observations and 6 observed elements (dry & wet bulb, max & min temperatures, rainfall and windrun) requires 80 records (screenfuls) entered prior to any daily data. 50 climate stations with 5-12 elements and 70 (single element) rainfall stations requires 5000+ preloaded screenfuls; scarcely a rapid start to key-entry of data! There is a copy procedure that some frustrated early user inserted into the 'FULL REPORTS' which is not in the Clicom manual. This hastens the initiation and avoids being 'beeped' by the key-entry system, but copying country-wide or global monthly extremes to every station dilutes the much vaunted quality control specially built in to enhance the system.
   * Key-entry of data sets is assumed to be most natural along rows of different element data for the same day i.e. the 'traditional' non-intelligent 'pool of punch operators' concept. However a few minutes in a manual climate data centre would have revealed that data is/was check-added down columns for each element. Except for the qc relationships in the dry/wet/max/min temperatures (and possibly global/diffuse radiation) this familiar procedure is easily transferrable to single column data sets.
   * Other global limits for each element and relativity restrictions (min,wet-bulb<dry-bulb<max & diffuse<global radiation) must also be setup in a quality control file.

3. * The properties of real data are assumed to be sufficiently qualified by a single character 'flag'. This would be a parsimonious description for climate data anywhere let alone developing countries. Certainly how these flags might be used in analysing data seem to descend to the soft option 'if in doubt, leave it out'! Apart from the resolution curio 'T'=trace (i.e. < 0.05 mm) for rainfall, the suggested flags are 'I'=incomplete, 'M'=missing, 'G'=generated (from other data, humidity?). There is no recognition, at least mention, of the distinction between 'missing = data lost or not recoverable' and 'missing = observation not made (observer left early morning to go fishing!). In the latter case rainfall, windrun and evaporation would be accumulated to the subsequent observation and max/min temperatures extended to cover a longer period. One presumes that 'missing' will also be invoked if the arbitrary threshold set by Clicom of "at least 20 days in the month" is not met. Again whether 'incomplete' means exclusion from analysis based on the WMO '3/5 rule' (a statistically pre-maximum-likelihood relic) or the less strict Clicom '20 days rule' is uncertain.

4. * Standard 'properties of the numbers collected' are provided in the menu of Clicom products, plus elementary graphical presentations; some of more use than others. According to the recent release of V 3.0 most development over the last 4 years has gone into improvements to the graphics e.g. better wind rose presentations. Evaporation estimation by the Penman or Priestley-Taylor formulae is not considered a 'standard product'. Anything else that might be called a climate product as distinct from a data product has yet to be included as a Clicom routine.

With my brief exposure to Clicom I have been so bold as to implement some changes and would suggest others that I wish had been included in the original development.

1. * A basic 'stripped down' system for daily data leading to decad, monthly, annual summaries- to do these simple things has been made very complicated for the uninitiated.
   * A simple structure of element codes with functional abbreviations that can be used for deciding whether to summarise an element as a total or mean (or not at all).
   * Using a blank field and flag 'M' for missing data- nothing looks less like being missing than -99999! However the chosen database system does not distinguish between a blank (no data entered) and a keyed-in zero for within record computations on numeric fields. If they were defined as text fields the distinction would be made.

2. * An option for the present and obvious future with key-entry of data as 'one day-many stations' in addition to the provided 'one station-many days' for the traditional monthly returns of the past. This would be the most natural form of data entry for subsequent processing with the facilities of the chosen database software which can process functions such as total/mean/se/max/min across but not within records. One day-many stations was the implied form of data entry with on-line telemetry in the original specifications but not as a manual key-entry option.
   * Shortcuts for keying in rainfall data recognising that in many target countries there is often no rain in a month- 60 key depressions is tedious.

3. * Flags by themselves are a weak qualifier unless supported by an associated measure such as number of days observations included in the total or average (this is derived in a separate 'inventory' record in Clicom), or width of interval in which the true data value lies etc. Accumulations over, or extremes extended through, more than one day will naturally lead to decad/monthly totals/averages that are confined within an interval. Common recording failures with autographic recorders such as pens drying up or traces lost in a large ink blot, blockages by blown leaves, twigs or bird stools etc, give lower bounds. Fixed clock based observations provide interval bounds for arbitrary time base data, and dominant events can give maxima interpretable as upper bounds. Hence in the real world we must allow data as point values, lower bounds, upper bounds, or within some interval. For any chosen statistical model the estimation of parameters from such data is not much more complicated than for the ideal data set entirely of point values. The information content of this qualified data is relatively high in statistical measure terms and simply to avoid wasting information, or to obtain a sufficiently large data set for analysis, it should be used.

Examples from the author reference below demonstrate the information content of sparse data sets.

REFERENCES

Revfeim, K.J.A., 1990: Daily observations; necessity, ritual or imposition. J. Climatology, 10, 105-110.
Clicom Reference Manual, V2.1 1988, V3.0 1992: NOAA/NWS, International Activities Division.

5th International Meeting on Statistical Climatology

Comparative Energetics of FGGE Reanalyses using the Normal Mode Expansion

by
Hiroshi L. Tanaka
Institute of Geoscience
University of Tsukuba, Japan

and
Qiang Ji
Geophysical Institute
University of Alaska Fairbanks, U.S.A.

1 Introduction

One of the main purposes of FGGE (First GARP Global Experiment) was to improve the medium-range forecasting. The FGGE III-b datasets were produced by two central institutions: the European Center for Medium Range Weather Forecasting (ECMWF) and the Geophysical Fluid Dynamics Laboratory (GFDL). After the release of the FGGE III-b datasets, numerous diagnostic studies have been conducted to evaluate and compare the data assimilation systems of different institutions (e.g., Kung and Tanaka 1983; Lau 1984; Paegle et al. 1986). With the results from those comparative studies, understanding of the distinct characteristics of different FGGE datasets has increased.

It is noted that the GFDL analysis of the FGGE data is noisy containing significant small-scale structure, and the mass and momentum fields are unbalanced. The spurious excitation of the small-scale disturbance may be resulted from the continuous data insertion. On the other hand, the ECMWF analysis is pointed out to have major deficiencies in the tropics. The vertical velocity is weaker in the ECMWF than the GFDL (Kung and Tanaka 1983), and the divergent wind has been damped by a diabatic version of the nonlinear normal mode initialization (Paegle et al., 1986). The imposed 95% constraint toward the geostrophic balance seems to be too restrictive in the tropics.

Since the production of the FGGE original analyses, there has been a continuous process of modifications and improvements. The day-to-day operational experience has provided invaluable information on the performance of the system. With these accumulated knowledge of the deficiencies in the former data analysis methods, both of the ECMWF and GFDL conducted the FGGE III-b *reanalyses* based on the improved assimilation techniques. Refer to Uppala (1986) for the detail of the ECMWF reanalysis and Stern and Ploshay (1992) for the GFDL.

The purpose of this study is to compare the energetic characteristics of the FGGE reanalyses with those of the FGGE original analyses. The comparison is made of both the ECMWF and GFDL versions of the original and reanalyses.

One of the useful diagnostic tools for that purpose is the normal mode energetics developed by Tanaka and Kung (1988). In the normal mode energetics, observed atmospheric data are expanded in the 3-D normal modes for a resting atmosphere which are constructed by a product of the vertical structure functions and Hough harmonics. The eigenfrequency of a 3-D normal mode may be regarded as a 3-D scale index for each of the gravity (first kind) and rotational (second kind) modes due to the intrinsic dispersion relation. The atmospheric energy spectrum is represented as a function of the 3-D scale index, and the nonlinear energy interactions among the different scales of motion are examined.

This diagnostic analysis is especially useful in comparing the amount of high-frequency gravity modes which are closely related to the divergent field and are most sensitive to the assimilation system. Therefore, we have paied a special attention to the energy spectra of the gravity modes over the 3-D scale index. The difference in the high-frequency gravity mode energy between the ECMWF and GFDL analyses is presented. Finally, an assess-

ment is made of how the difference in the original analyses is modified by the new assimilation systems.

2 Data and analysis scheme

The FGGE reanalyses by the ECMWF and GFDL for the Special Observing Period 1 (SOP-1) are obtained from the National Center for Atmospheric Research (NCAR). Twice daily (0000 and 1200 UT) meteorological variables of horizontal wind vector $V = (u, v)$, vertical p-velocity $\omega(= dp/dt)$, temperature T, geopotential height Z, and relative humidity R are defined on 1.875° by 1.875° horizontal grids at 12 mandatory vertical levels of 1000, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50 mb. For the GFDL dataset, the top two vertical levels are given at 50 and 30 mb rather than the aforementioned levels. These FGGE data are interpolated onto 5° longitudinal grids and 60 Gaussian latitudinal grids.

The analysis scheme of the normal mode energetics has been detailed in Tanaka and Kung (1988). The brief description of the scheme is summarized here. First, applied to a sequence of vertical, Fourier, and Hough transforms, the primitive equations become a system of ordinary, dimensionless differential equations in the 3-D spectral domain:

$$\frac{d}{d\tilde{t}} w_{klm} + i\sigma_{klm} w_{klm} = n_{klm} + d_{klm}, \qquad (1)$$

where the complex variables of w_{klm}, n_{klm}, and d_{klm} represent respectively the expansion coefficients for a vector (u, v, Z), nonlinear term vector for the momentum and mass fields, and the diabatic processes including friction. The symbol i represents the imaginary unit, time t is normalized as $\tilde{t} = 2\Omega t$ with the angular speed of the earth's rotation Ω, and σ_{klm} is a dimensionless eigenfrequency obtained as a solution of Laplace's tidal equation with a basic state at rest. The subscripts k, l, m are for the zonal wavenumber, meridional index, and vertical index, respectively. The vertical indices m=0 and $m \neq$0 represent barotropic (external) and baroclinic (internal) modes, respectively. We have used a total of 50 meridional modes, including 26 Rossby modes and 12 eastward and 12 westward gravity modes.

An element of total energy is simply defined by an energy norm of the system:

$$E_{klm} = \frac{1}{2} p_s h_m |w_{klm}|^2, \qquad (2)$$

where the dimensional factors p_s and h_m are respectively a constant pressure near the surface and an equivalent height of mth vertical index. The energy spectrum expressed as a function of σ_{klm} represents the 3-D spectral energy distribution. By summing the energy terms with respect to all indices but k, the result describes the zonal energy spectrum, while the summations over k and l describe the vertical energy spectrum.

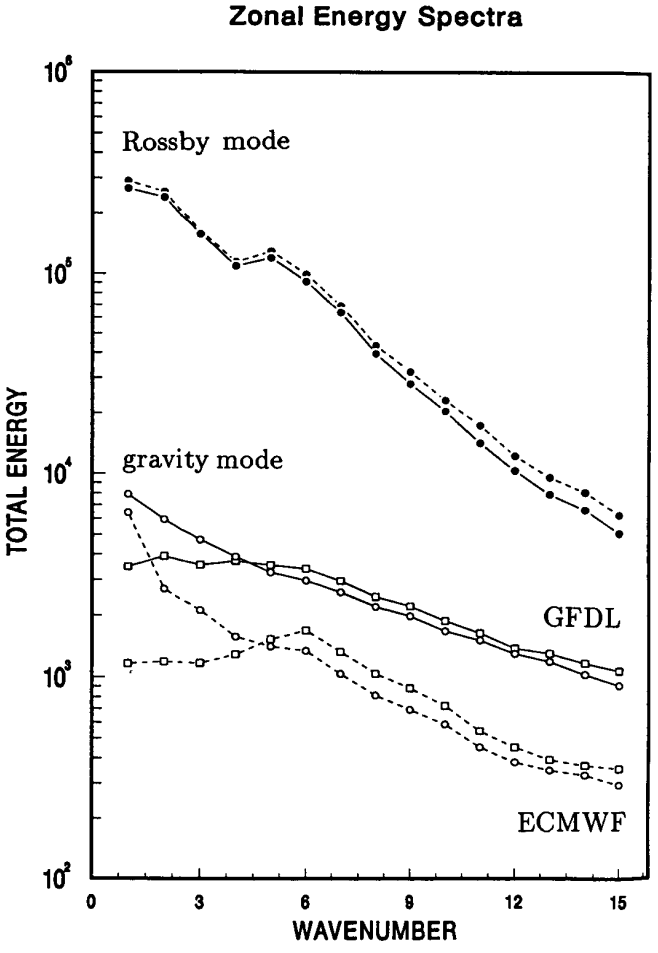


Fig. 1. Zonal energy spectra for the ECMWF (dashed) and GFDL (solid) reanalyses for k=1-15. Dots, squares, and circles denote Rossby modes and westward and eastward gravity modes.

3 Results of the analysis

Figure 1 illustrates zonal energy spectra for k=1-15 as obtained by summing all meridional and vertical indices for Rossby and gravity modes. The Rossby mode energy levels indicate minor difference between the ECMWF (dashed) and GFDL (solid). The gravity mode energy level is, however, substantially different between these two reanalyses. The GFDL analysis is about 4 time larger than the ECMWF analysis at k=15. For planetary waves of $k \leq 4$, the eastward gravity modes are greater than the westward gravity modes due to the dominant Kelvin waves.

Figure 2 shows the energy spectra of the ECMWF and GFDL reanalyses expanded in the 3-D NMFs. The energy elements of E_{klm} in Eq. (2) are plotted as functions of the scale index σ_{klm} for 50 meridional indices l, zonal wavenumbers k=1-6, and vertical indices m=0 (Fig. 2a) and m=2 (Fig, 2b). The difference in the modal energy levels are visually presented by deviation bars. The energy levels for the GFDL reanalysis are plotted with symbols, and the deviation bars are extended toward the the ECMWF reanalysis without symbols. The comparison of the external modes of m=0 for the two reanalyses shows similar energy spectra for the Rossby modes. However, the discrepancy is evident in the gravity-mode energy spectrum. The gravity mode energy levels for the ECMWF are lumped together near 10 Jm^{-2}, and a power law is not recognized. The low-frequency gravity modes are significantly damped in the ECMWF; the energy level is one order of magnitude lower than the results of the GFDL reanalysis. Figure 2b compares the same energy spectra, but for an internal mode of m=2. Again, the Rossby-mode energy spectra are similar for these two reanalyses. It is shown that planetary-scale Kelvin modes dominate the other gravity modes, and the energy levels reach up to 1000 Jm^{-2} for these two reanalyses. A discrepancy is found in the high-frequency (i.e., small-scale) gravity modes; the energy level of the GFDL is about one order of magnitude higher than that of the ECMWF for the smallest resolvable-scale modes.

4 Concluding remarks

This study conducted the comparative diagnosis of the FGGE III-b reanalyses produced recently by the ECMWF and GFDL. The energy levels of the atmospheric gravity modes are examined by means of the normal mode energetics scheme.

As a result, we confirmed that the Rossby mode energy levels are virtually same for these FGGE datasets. The Rossby modes appear to be less sensitive to the analysis technique. However, a notable discrepancy is found in the gravity-mode energy spectrum between the ECMWF and GFDL analyses. For the external modes, the GFDL analyses exhibit a power law in the gravity mode energy spectrum with respect to the 3-D scale index σ_{klm}. In contrast, the gravity mode energy levels for the ECMWF are lumped together near 10 Jm^{-2} level, and the power law is not recognized. In the ECMWF reanalysis, the low-frequency (i.e., largest-scale) external gravity modes are one order of magnitude lower than the results of the GFDL reanalysis. The internal modes are also damped by an order of magnitude at the high-frequency (i.e., small-scale) component, compared with the corresponding GFDL reanalysis.

It is worth noting that the gravity mode energy has increased by 50% for the GFDL reanalysis compared with its original analysis, whereas it has decreased by 10% for the ECMWF reanalysis. Hence, the discrepancy in the gravity modes between the ECMWF and GFDL has grown rather than reduced by the reanalyses. This result was unexpected from the fact of smoother fields in the GFDL reanalysis and intensified Hadley circulation in the ECMWF reanalysis. It is speculated that the GFDL reanalysis contains larger amount of ageostrophic wind as found from the poor fit of the geopotential field to the observed field (see Ploshay et al. 1992).

Acknowledgements: This research is supported by the National Science Foundation under Grant No. ATM-8923064.

References

Kung, E. C. and H. Tanaka, 1983: Energetics analysis of the global circulation during the special observation periods of FGGE. *J. Atmos. Sci.*, **40**, 2575-2592.

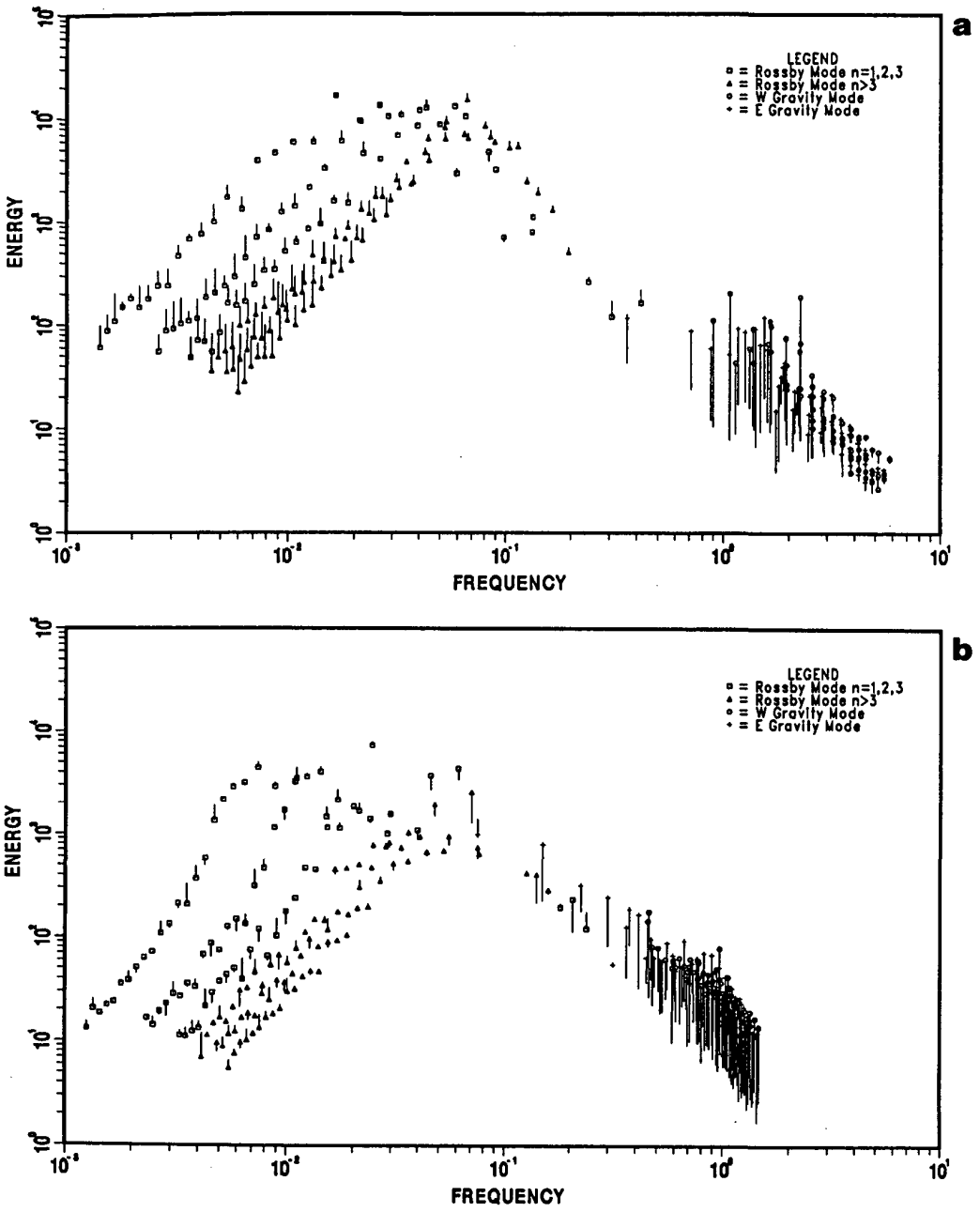


Fig. 2. Energy spectra of the GFDL (with symbols) and ECMWF (deviation bar) reanalyses for vertical indices m=0 (Fig. 2a) and m=2 (Fig, 2b) as functions of the scale index σ_{klm}.

Lau, N.-C.,1984: A comparison of circulation circulation statistics based on FGGE level III-b analyses produced by GFDL and ECMWF for the special observing periods. *NOAA Data Report ERL GFDL-6,* 237pp.

Paegle, J, W. E. Baker, and J. N. Paegle, 1986: The analysis sensitivity to tropical winds from the Global Weather Experiment. *Mon. Wea. Rev.,* **144**, 991-1007.

Ploshay, J.J., W.F. Stern, and K. Miyakoda, 1992: FGGE re-analysis at GFDL. *(to be submitted.)*

Stern, W. F., and J. J. Ploshay, 1992: A scheme for continuous data assimilation. *(to be submitted.)*

Tanaka, H. L., and E. C. Kung, 1988: Normal mode energetics of the general circulation during the FGGE year. *J. Atmos. Sci.,* **45**, 3723-3736.

Uppala, S., 1986: The assimilation of the final level IIb data set at ECMWF, Part 1. *Prep. National Conference on the Scientific Results of the FGGE, Miami, Florida,* 24-29.

The North Atlantic Climatological Dataset (NACD). - A necessary tool for monitoring climatic change in the Nordic Region.

John Cappelen
Danish Meteorological Institute
Lyngbyvej 100,DK-2100 Copenhagen,Denmark

Abstract:
Nine meteorological institutions from northwestern Europe are involved in NACD. The dataset comprises monthly data from more than 70 stations in the North Atlantic Region (see enclosed map). Our common purpose is to homogenize existing time series of not only temperature and precipitation but all the observed elements of climate. Each country involved has been observing various climate-elements for more than 100 years, and now the time has come to gather this information into a common computerized dataset which can be used for analyzing climatic trends and variability on a larger scale.

A number of analysis can be performed with a homogenous dataset, which comprises more than 20 climatic elements. Basically we aim at 4 kinds of analysis of the homogenized dataset.

1) Spatial and temporal variations in single elements which will include both trend and variability analysis.

2) Spatial and temporal coveriance between several elements, like f.ex. cloud cover, number of clear and overcast days, hours of bright sunshine and precipitation. Another such analysis could be done by comparing pressure gredients with mean wind speed and wind direction, and number of days per month with storm.

3) Correlation with relatively short time series like cloud cover or radiation measured from satellites, upper-air soundings of pressure, wind or temperature or ground-based observations of soil temperature. Through such analysis it is possible to extrapolate monthly means of these parametres back to the previous century. Such time series can be used for validating the results of regional climate modelling.

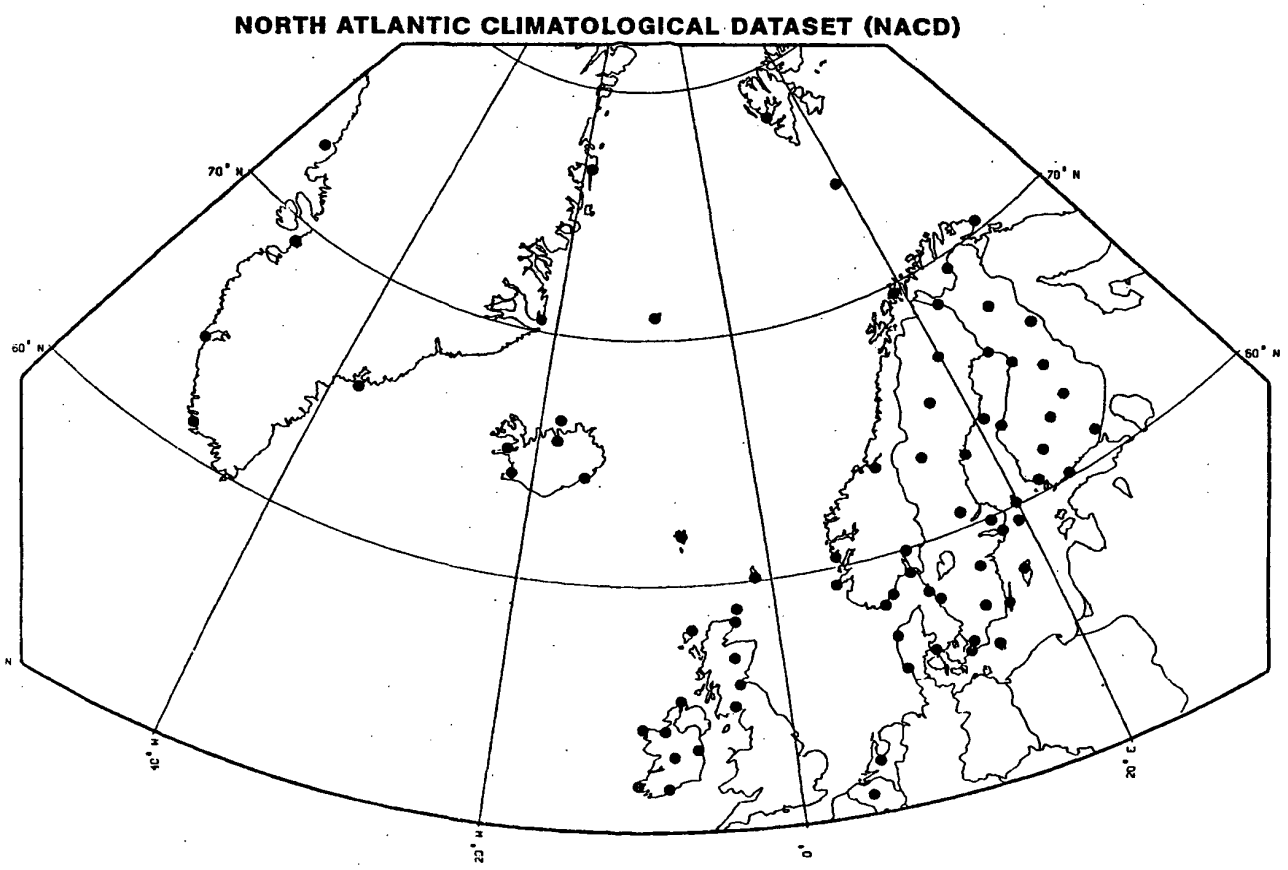
4) Correlation with relatively long time series of proxy data like f.ex. ice core data of stable oxygen isotope variations, tree ring data or deep sea sediments, and thereby supply valuable information concerning a more detailed climate history of the North Atlantic region prior to the instrumental period.

The NACD - dataset could also be used for analysis of changes in the atmospheric circulation in the North Atlantic region and for validation of general atmospheric-circulation modelling.

The Danish contribution to NACD consists of data from 4 stations in Denmark, 1 station on the Faroe Islands and 7 stations in Greenland. More than 20 climate-elements, including cloud cover, hours of bright sunshine, air pressure, dominant wind direction, storminess and frequency of various weather phenomena have been digitized and some preliminary results will be shown. The poster will also describe some of the more severe inhomogeneities in climatic time series during the past 120 years.

Recent results from coupled ocean-atmosphere models show that the most prominent warming should be expected in the Arctic during the NH winter. Most Greenland stations show a warming trend during winter, and this is accompanied by an increasing trend in cloudiness. Whether this ongoing change should be ascribed to the enhanced Greenhouse effect still remains an open question.

The problem of monitoring climatic change would be much easier to solve if a comprehensive and homogeneous dataset could be established. In this respect NACD could become a valuable tool for decision-makers in the near-future.



OBJECTIVE ANALYSIS MODEL FOR THE TIME - SPACE VARIATION OF THE WIND VECTOR

Aristiţa Busuioc
National Institute of Meteorology
and Hydrology
Bucharest-Romania, Sos. Bucureşti-Ploieşti 97

ABSTRACT

The paper presents an objective analysis method of the daily wind data by using the method of principal components analysis applied to the vector fields of data. The extension of this method has been tried for the analysis of the monthly data set for a long time period. The essential space features existing in the initial observations data are given by the primary spatial eigenvectors and the time variations are reflected by the principal component associated to the primary eigenvector. This analysis model has been applied to the study of the wind field for Romanian littoral.

1. INTRODUCTION

The statistical method of principal components analysis, frequently referred to as empirical eigenvector analysis, and the closely related procedures of factor analysis have been used to study many different types of observational data. The most of these studies were limited to the analysis of the scalar fields of data. Hardy and Walton (1978) presented a generalization of the principal components analysis to the vector values.

Vector parameters are very important in many physical problems. Treatment of vector observations is achieved by an appropriate use of complex rather than real numbers in the analysis.

All the applications which refer to the analysis of the principal components begin with a measurement set of a physical parameter that can be grouped into a rectangular matrix. For this matrix the eigenvectors and the eigenvalues are calculated.

The eigenvector corresponding to the largest eigenvalues has the highest significance in representing the data and is called the primary eigenvector. The most significant eigenvectors can be identified with physically important patterns in the original data. Temporal expansions of the data in terms of the most significant eigenvectors are known as the principal components.

This paper presents the results obtained by applying the principal component method to the analysis of wind vector for the Romanian littoral. It has been tried the extension of this method to the treatment of the monthly data set for long time period.

2. DESCRIPTION OF THE MODEL

The vectorial observations to be analyzed are the wind vectors $\vec{V}$ measured in horizontal plane, simultaneously in a number N of geographical locations. These data are represented in the analysis as complex numbers, each individual observation $\vec{V}$ having a complex number associated (Hardy D., Walton, J., 1978),

$$\sigma = s e^{i\theta} \qquad (1)$$

where s - the wind observed velocity
θ - the corresponding direction.

Within this analysis, the wind direction is defined countre-clockwise (from the positive axis x) in accordance with the usual mathematical convention. The positive axis x coincides with the east direction and the positive axis y with the

north direction.

The complete set of observations is represented as a matrix S of N x M size, of the form:

$$S = (\sigma_{km})_{\substack{k=\overline{1,N} \\ m=\overline{1,M}}} \quad (2)$$

where k indicated the geographic location and m - the time moment.

Matrix H is defined by the relationship:

$$H = \frac{1}{M} S \cdot S^{+} \quad (3)$$

where S^{+} is the complex conjugate transpose of S. It follows from its definition that H is Hermitian and of dimension N x N, so all its eigenvalues are real and the eigenvectors are orthogonal, the elements of which are generally complex. Note that if Θ_k is an arbitrary phase factor associated with the k^{th} eigenvector $\hat{E}_k$, the set of vectors defined by:

$$\hat{B}_k = e^{i\Theta_k} \cdot \hat{E}_k, \quad k=1,\ldots,N \quad (4)$$

are also eigenvectors of H and satisfy the orthonormality condition.

The component of order k, ε_{kj}, of each eigenvector $\hat{E}_j$ is associated with the meteorological station having the number k. ε_{kj}, being complex numbers, can be written in the form:

$$\varepsilon_{kj} = a_{kj} \cdot e^{i\alpha_{kj}} \quad (5)$$

Values a_{kj} and α_{kj} can be used in order to plot the N two-dimensional vectors at the locations corresponding to the meteorological stations, where the observations have been carried out. The eigenvectors are rotated with a common phase angle by using relationship (4) so that they may take over the characteristic direction of the data in the area.

The N eigenvectors form a complete orthonormal basis of the N dimensional space. Any of the vectors S_m of this space can be thus:

$$\hat{S}_m = \sum_{k=1}^{N} C_{km} \hat{E}_k \quad (6)$$

where,

$$C_{km} = \langle \hat{E}_k^{+}, \hat{S}_m \rangle \quad (7)$$

C_{km} coefficients are generally complex and the series of C_{km} coefficients ($m=\overline{1,M}$) is a time series expansion in terms of $\hat{E}_k$ and is known as the principal component of order k.

For each eigenvector $\hat{E}_k$, the relative magnitude is defined:

$$F_k = \lambda k / \sum \lambda j, \quad k=1,\ldots,N \quad (8)$$

representing the figure of merit of the eigenvector $\hat{E}_k$.

The closer to 1 F_k is, the more exact the data representation is, as the product between the coefficient C_{km} and the eigenvalue $\hat{E}_k$. Out of this reason the eigenvectors corresponding to the highest eigenvalues are called primary eigenvectors and emphasize the spatial characteristics of the wind in the area. The principal component corresponding to the primary eigenvector emphasize the wind time features in the area.

3. THE PRACTICAL RESULTS OBTAINED

The model presented in paragraph 2 was applied to the study of the wind field for the Romanian littoral, considering the 6 meteorological stations, arbitrarily numbered, as it follows: Sulina (1), Sf. Gheorghe (2), Constantza (3), Mangalia (4), Unirea (5), Corugea (6). The data selected for the analysis include measurements for wind direction and speed, every 3 hours, in horizontal plane, over a 12 month period (January-December 1984).

The dimension of matrix S depends on the number of days of the month that varies between 6x248 (31 days) and 6x232 (29 days). The eigenvalues the eigenvectors of matrix H as well as the figure of merit of the primary eigenvector (F_1) for each separate month are calculated.

Table 1. The figure of merit F_1, eigenvalue (λ_1) and eigenvector elements for the primary eigenvector derived for May 1984

F_1	1	Eigenvector elements 1	2	3	4	5	6
0.6	85.2	1.0	0.9	1.6	1.9	1.1	0.6
		310° SE	309° SE	307° SE	309° SE	300° SSE	309° SE

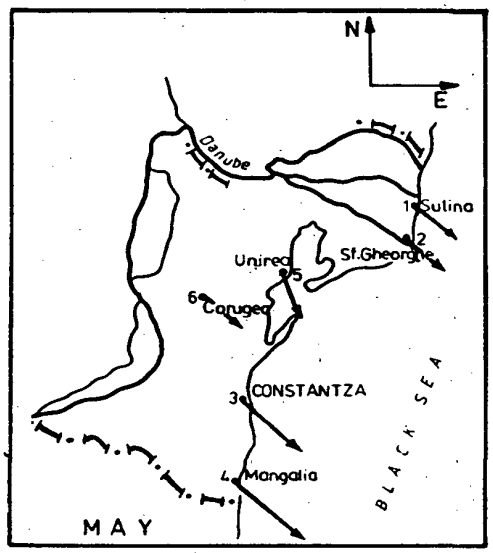


Fig.1 Geometric representations of the primary spatial eigenvector for May 1984

From the analysis of F_1 for each month, generally, it has been noticed that the eigenvectors for the warm months had the highest significance (the maximum value was recorded in May) and those for the cold months, the lowest (the minimum value was recorded in December). This difference is due to the persistency of a certain type of circulation during the warm months as against the cold months showing a higher meteorological variability. The eigenvectors resulted from the calculation have rotated by a θ_1 angle allowing, each two-dimensional vector of the primary eigenvector geometric representation to achieve a meaningful orientation of the data for each meteorological station.

From the analysis of the primary eigenvector argument, the predominant direction and sectors were emphasized for each separate month. This it was noticed that for January, April, August, November and December, generally, the winds from northern sector are predominant and in May, June, September and October the winds from the S-SE sector are predominants.

These results are very close to those obtained from the study of the real data fact which shows that the elements of the primary eigenvectors objectively emphasize the spatial charactristics of the wind in the area. As an exemple table 1 shows the primary eigenvalue λ_1, the elements of the primary eigenvector E_1, the corresponding figure of merit F_1 for May. For each element of the primary eigenvector its magnitude expressing the wind velocity is given (in relative units) as well as the argument expressing the wind direction (in degree).

Fig.1 shows the geometric representation (using the data of table 1) of the primary eigenvectors for May (predominant direction S-SE).

The coefficient C_{1m} associated to each primary eigenvector shows, as previously mentioned, the modifications depending on time, of the wind vector in the studied area. Fig. 2 presents, as an example, the magnitude and the argument of coefficient C_{1m} for the first 6 days of May.

In case of the magnitude of C_{1m} the existence of a periodicity can be noticed for the daily cycle. For each day the wind velocities minimum during the night and during the first hours of the morning and maximum towards noon and in the afternoon.

The argument of C_{1m} (rad.) shows unlike the magnitude, a less regular variation. The argument values represent a common rotation of the wind velocity at the 6 stations.

It was noticed that the maximum

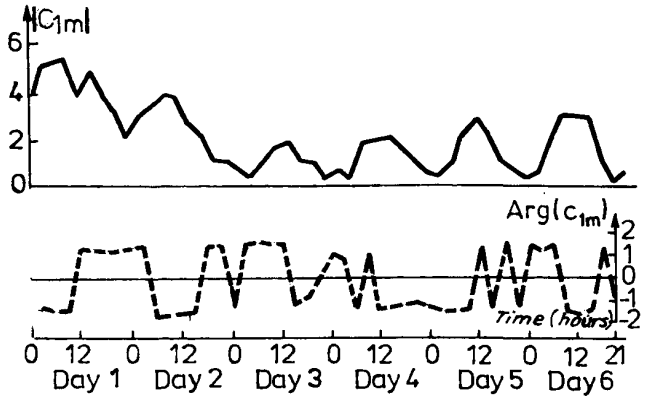


Fig.2 The magnitude and argument variation of the principal component C_{1m} for the first 6 days of May 1984

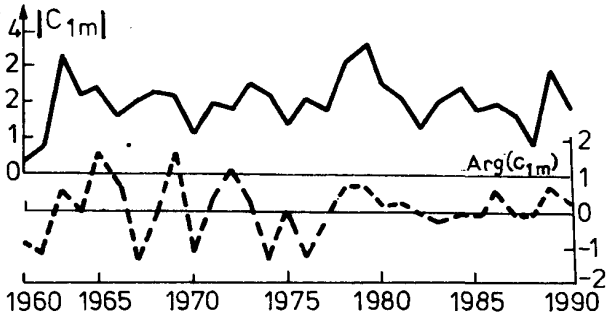


Fig.3 The magnitude and argument variation of the principal component C_{1m} for May (1961-1990)

values of the argument, over the 6 day period, are approximately equal to 1.5 rad., fact which shows a deviation of the wind direction westwards (in case of the negative values of the argument) or eastwards (in case of the positive values of the argument).These changes of the wind direction can also be noticed in the real data being determined by NE or SW circulations added to effects such as "breeze" generated by the temperature obstacle of the Black Sea(Drăghici,1988).

It has been also tried to apply the model presented above to the analysis of monthly wind data for a 30 years period (1961-1990). For this as initial data was considered the wind speed corresponding to the most frequent direction for every month. The matrix S dimension for each separate month is 6x30.

From the analysis of the primary eigenvectors of matrix H for 4 representative months (January, May, Julay, October) it has been noticed that it is kept the same space features as in case of the daily values analysis.

As an example, fig.3 shows the time variation of the magnitude and the argument of the coefficient C_{1m} for 1961-1990 period. The magnitude shows,unlike the magnitude of daily data, a less regular variation.After 1975 the monthly predominant direction at all 6 stations maintains, generally, the Southern Sector.

4. CONCLUSIONS

The presented results show that the model allows to objectively obtain,from the initial observations, the essential spatial and temporal features of the wind vector in the analyzed area, over the considered time period.

The spatial wind features in the area for each separate month area given by the primary eigenvectors emphasizing the dominant direction for each station.Thus,generally, certain difference can be noticed between the cold period of the year (when the predominant direction of the wind is from the northern sector) and the warm period (when the predominant wind direction is from the southern sector).

The temporal features are reflected by the principal component associated to the primary eigenvector emphasizing the existence of a periodicity within the daily cycle as regards the magnitude of the wind velocity.

The extension of the respective method for the analysis of the monthly predominant wind direction for a 30 years period has emphasized the same space features as in case of daily data.

REFERENCES

1. DRAGHICI, I.,(1988) Dinamica atmosferei,Editura Tehnică, Bucureşti, 398-471.
2. HARDY,M.D.,WALTON,J.J.,(1978) Principal components analysis of vector wind measurements.J.Appl. Meteor. 8, 1153-1162.

APPROACHING CLIMATIC REPRESENTATIVITY OF SHORT SAMPLES THROUGH SUBSAMPLES WITH MACROSYNOPTIC SIMILARITY TO CLIMATE

J. Mika, T. Szentimrey, P. Domonkos, G. Baranka, T. Szabó,
Hungarian Meteorological Service, H-1525 Budapest, P. O. Box 38,
C. Károssy, Berzsenyi Coll. H-9701 Szombathely, Eötvös 1, Hungary

A method based on local macrosynoptic types is introduced to compile 'climatic samples' within any shorter sample, for which the actual relative frequency distribution of the macrotypes is not equal to their 'a priori probability' (ie. climatic average). Therefore 'climatic' subsamples belonging to the different macrosynoptic types are formed by decreasing the subsamples' sizes to obtain relative frequencies of macrotypes, similar to the a priori probabilities. Optimum approach for making the samples' relative frequencies similar to the a priori ones is delivered by minimizing a quantity, which is an upper estimate of the squared difference between climate and the average of the climatic sample. An algorithm for finding this optimum frequency distribution is established as a theorem. After specifying the necessary frequencies of macrotypes, a randomization method is applied to chose the concrete days involved. The method can be applied for any climatic or environmental elements depending on actual weather, not integrating the effects of previous days.

1. INTRODUCTION

Climatic representativity presumes at least a few decades of measurements under constant conditions of observation, which is not the case for numerous climatic and air-quality observations or expeditional measurement programs. Several weeks or months of observations under the same circumstances can be rather valuable and difficult to repeat. However these short samples are hardly representative for climate of the given region, because they bear specifics of the given period that might be quite anomalous. Therefore 'climatic subsamples' will be defined.

2. SELECTION OF CLIMATIC SUBSAMPLES

The statistical problem is the following. Let's assume that:

the number of macrosynoptic types is K;
the a priori probabilities of these macrotypes are:

$$p_i \qquad i=1,\ldots\ldots,K$$

and the conditional distribution functions belonging to the given macrotypes are:

$$F_i(x) \qquad i=1,\ldots\ldots,K$$

In this case the distribution function of 'climate' can be formulated like that

$$F(x) = \sum_{i=1}^{K} F_i(x)\, p_i$$

Let's assume that there are given K conditional subsamples belonging to the corresponding macrotypes i.e.:

$$\delta_{i,j} \qquad j=1,\ldots\ldots\ldots,N_i;\ i=1,\ldots\ldots,K$$

where N_i is the size of the i. subsample.

In this case the size of the joint sample is

$$N = \sum_{i=1}^{K} N_i$$

and the distribution function of this joint sample

$$F_N(x) = \sum_{i=1}^{K} F_i(x)\,(N_i/N)$$

As can be seen the similarity of the distribution functions $F(x)$, $F_N(x)$ depends on the differences of the a priori probabilities p_i and the relative frequencies N_i/N $(i=1,\ldots,K)$.
The aim is the minimization of the supreme

$$\sup_x \left(F_N(x) - F(x) \right)^2$$

by decreasing the sizes of the conditional subsamples to obtain a climatic sample, and the conditional distribution functions $F_i(x)$ $(i=1,\ldots,K)$ are not known in general.
But taking into consideration the following inequality

$$\sup_x (F_N(x) - F(x))^2 \le \sum_{i=1}^{K} (N_i/N - p_i)^2/p_i$$

we can turn to solving of the minimum problem mentioned below to specify the sizes of the conditional subsamples.

Minimum problem:

$$\min \sum_{i=1}^{K} (n_i/n - p_i)^2/p_i \qquad (1)$$

where $0 \le n_i \le N_i$ and $n = \sum_{i=1}^{K} n_i$.

This minimum problem can be solved by the following <u>algorithm</u>.

<u>i</u>, the 0. step:

$$N_i^{(0)} = N_i \qquad i=1,\ldots,K$$

$$H^{(0)} = \sum_{i=1}^{K} (N_i^{(0)}/N - p_i)^2/p_i$$

<u>ii</u>, the M. step $(M=1,\ldots,N-1)$:

$$N_i^{(M)} = N_i^{(M-1)}, \quad N_j^{(M)} = N_j^{(M-1)} - 1$$

for $j \neq i$, <u>if and only if</u>

$$(2N_i^{(M-1)} - 1)/p_i = \min\left((2N_j^{(M-1)} - 1)/p_j\right)$$

and

$$H^{(M)} = \sum_{i=1}^{K} (N_i^{(M)}/(N-M) - p_i)^2/p_i$$

<u>Theorem:</u>
The solution of the minimum problem (1) is equal with the minimum $H^{(M)}$ $(M=0,\ldots,N-1)$.

The presented algorithm makes possible to solve the minimum problem (1) effectively, and after specifying the necessary subsamples' sizes, a randomization method should be applied to generate the 'climatic sample'.

3. THE APPLIED AND IMPROVED MACROSYNOPTIC CLASSIFICATION

<u>Péczely</u> (1957) has defined the macrosynoptic types around Hungary on the basis of sea-level pressure maps (Fig 1). He also published (<u>Péczely</u>, 1983) transition-matrices and several other statistical characteristics of the code-series for 1881-1980. (Since his death, - 1984 - one of the authors - <u>C. K.</u> - determines the actual codes.)

This typization has been widely used, although it has two shortcomings, narrowing the possibilities of its applications. <u>First</u>, standard deviation of climate elements remain high within several macrotypes.

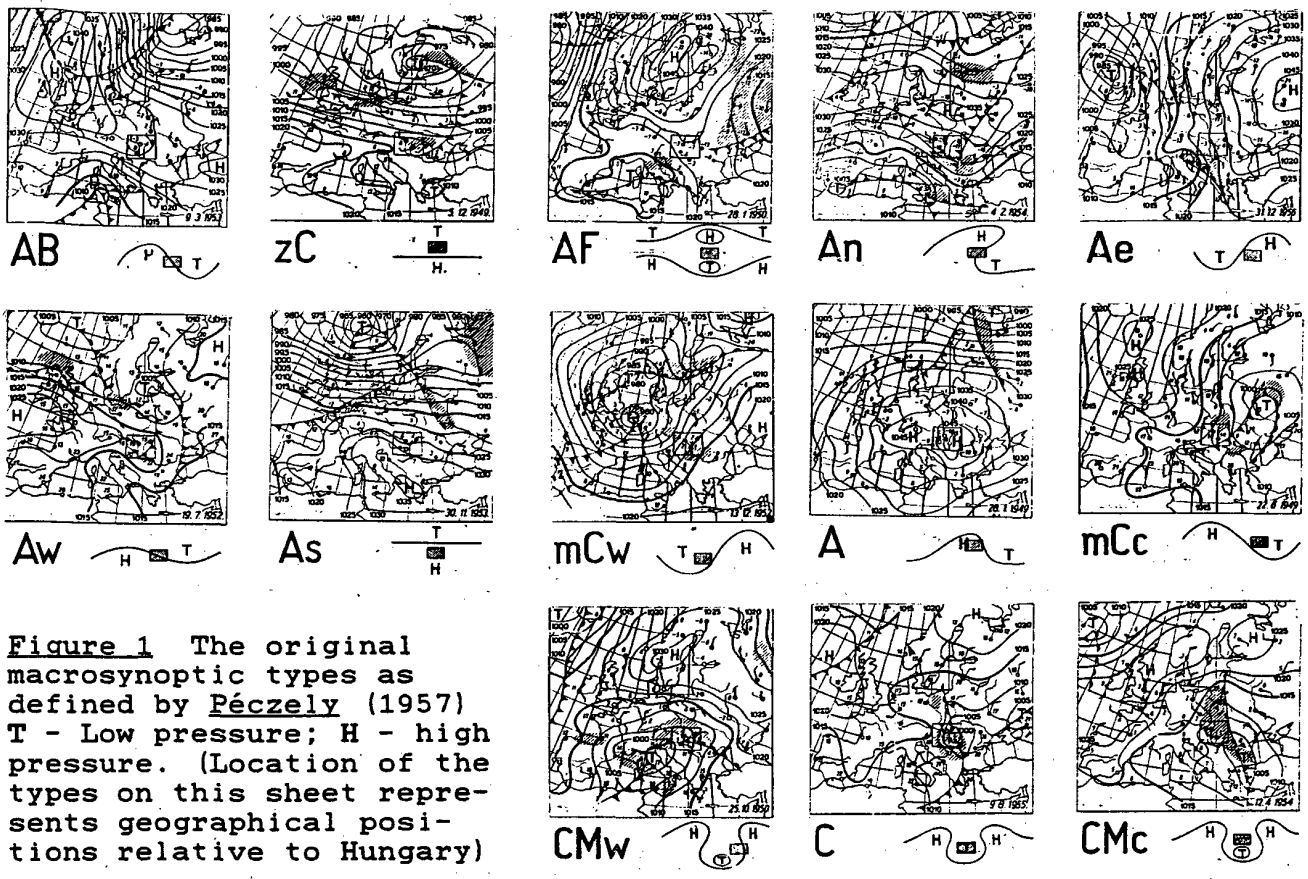


Figure 1 The original macrosynoptic types as defined by Péczely (1957) T - Low pressure; H - high pressure. (Location of the types on this sheet represents geographical positions relative to Hungary)

Table I. Split of 7 macrotypes as determined by the previous ones: I - zonal (high index) II - meridional (low index) realization of the original types. (Results of the first iteration step are set in bold) **Algorithm of sorting the original macrotypes into I or II:** First, it is investigated, wether the types which preceed the actual ones with a significant frequency, can appear one following the other with a significant frequency, or not. In the latter case two groups (I and II) are defined as two different macrotypes of the improved system. In the first case the macrotype remains unchanged. In this first step of sorting only 2-5 preceeding macrotypes could be sorted to groups I or II. For the remainder it has been determined that members of which newly formed group can preceed or follow them with a significant frequency. If this step was still not resultative, a new iteration has been done on the same basis. For types having significant connections to both groups, the decision has been made on the majority principle: The macrotype was attributed to the group, fro which these connections climatically repeated themselves with a higher frequency.

Onto: \From:	mCc	AB	CMc	mCw	Ae	CMw	zC	Aw	As	An	AF	A	C
AB	I	--	I	I	II	I	I	I	I	II	II	II	I
mCw	I	I	I	--	II	II	I	I	I	II	II	I	I
Ae	I	II	I	I	--	I	I	II	I	II	II	I	I
zC	I	I	II	I	I	I	--	I	I	II	II	I	I
Aw	I	II	I	I	II	I	I	--	I	II	II	II	I
An	I	II	I	I	II	I	I	I	I	--	II	II	I
A	I	II	I	II	II	I	I	I	I	II	II	--	I

Second, macrotypes defined in a window around Hungary can frequently be formed by different continental-scale processes.This can be proved applying the objective typization, (Ambrózy et al., 1984), which is a generalization of European macrotypes by Hess and Brezowsky, (1969). Applying calculations of Péczely (1983), the similarity in averaged 500 hPa fields of the European macro-types, appearing parallel to any fixed Hungarian macrotype with a significant frequency (at 90% level) is frequently weak or medium. Using scoring values, scaled from 1 to 9 (Ambrózy et al., 1984), 54 % of comparisons yield values 3-5.

Facing these shortcomings, it was decided to registrate macrotypes, following each other with a significant frequency (according to chi-square test at 80% level), as compared to the unconditional probabilities. Transition from one type into an other was established in a season, if i. the frequency of transition was significantly higher than the unconditional one in at least two months of the season, and ii. the average of monthly relative frequencies were higher than the threshold.

Sorting the transitions into groups means answering the question, that wether the flow of significant transitions between the macro-types form one or more groups (see Table I for details)

Splitting of the existing types, introduced completely by a mathematical algorithm, can be interpreted in terms of large-scale circulational patterns. Transitions forming group I correspond to zonal (high-index) processes, while those from group II represent blocking type (low index) patterns at larger scales

4. APPLICATION OF THE METHOD

Skill of the approximation depends on the statistical structure of the simulated element. The task of verification will be attempted by comparing approximated and 'true' averages for several climatic elements with different statistical structure.

The method will also be demonstrated on time series of the surface ozone content, observed at a regional background station (K-Puszta, Hungary) for 20 months, and also temperature data of the active surface above a well-watered sunflower canopy, which is only 35 days long.

The method can be used for a first estimation of the effects of any external forcing or global climate change too, in the way, that instead of the many-years average of the macro-types we use disturbed or hipo-tetically calculated frequencies

REFERENCES

Ambrózy P., Bartholy J., Gulyás O., 1984: A system of seasonal macrocirculation patterns for the Atlantic-European region. Idôjárás 88, 121-133

Hess, P., Brezowsky, H., 1969: Katalog der Grosswetterlagen Europas. Berichte des Deutschen Wetterdienstes, Bd. 15, N. 113

Péczely G., 1957: Grosswetterlagen in Ungarn. Kleinere Veröffentlichungen der Zentralanstalt für Meteorologie, No. 30, Budapest

Péczely G., 1983: Catalogue of the macrosynoptic types for Hungary (1881-1983). Publications of the Hungarian Meteorological Service vol. 53, Budapest

STOCHASTIC CLIMATE MODELS: THEORY AND APPLICATIONS

BY GERALD R. NORTH, TEXAS A&M UNIVERSITY

U.S.A.

Simple energy balance models forced by broad band noise in both space and time lead to fluctuating solutions which in many cases can be studied analytically or with simple algorithms. These models have now been compared and fitted to both real earth fluctuations and to GCM simulations of an idealized planet. In both cases the agreement for the geographical distribution of spatial correlation and variance statistics is excellent over intra- and inter- annual frequency bands with only the noise strength as an adjustable parameter. Applications will be demonstrated in studies of predictability of future states and in questions of the adequacy of the observing system to estimate trends in the surface temperature field.

Fluctuations of regional rainfall in climate models: Implications for climate change studies.

N.Nicholls, M.Speer, and W.Drosdowsky
Bureau of Meteorology Research Centre,
Melbourne, Australia.

INTRODUCTION

General circulation model (GCM) regional-scale predictions of rainfall changes associated with an enhanced greenhouse effect are considered unreliable. For example, the 1992 Supplement to the Intergovernmental Panel on Climate Change (IPCC 1992) notes that "Confidence in regional climate patterns based directly on GCM output remains low", and, "The prediction of climate change at regional level precision is subject to great uncertainty. Precipitation changes are particularly uncertain, although these changes are of great practical significance. Progress in the development of GCM models is urgently needed, particularly in terms of improving their capability for regional predictions". These statements reflect a general concern through the atmospheric science community regarding the credibility of regional-scale rainfall predictions from GCMs.

The cause of the lack of credibility of regional rainfall predictions is not, however, obvious. One possible cause is the low resolution of GCMs. This low-resolution means that the model orography does not truly represent the detail of the real world. Rainfall is often caused by interactions between the orography and the large-scale flow, so the lack of detail may invalidate the direct estimates of rainfall from a GCM. At the simplest level, the lack of detail will mean that the detail of regional rainfall is lacking in a low resolution model. This appears to be the reason for the lack of credibility of the regional rainfall predictions accepted by the authors of IPCC 1992 who assert that "GCM results can be interpolated to smaller scales using statistical methods (correlating regional climate with the large-scale flow) or a nested approach (high-resolution, regional climate models driven by large-scale GCM results). Both methods show promise but an insufficient number of studies have yet been completed to give an improved global picture of regional climate change due to increases in greenhouse gases". This focus on interpolation implies that the large-scale flow predictions are more credible than are the direct regional rainfall predictions.

The IPCC 1992 Supplement does note, however, that "both interpolation methods depend critically on the quality of the large-scale flow in the GCM". So it may be that the regional rainfall predictions are not credible because the large-scale flow predictions are themselves not credible. If this was the case then use of an interpolation method would not lead to more credible regional rainfall predictions.

It would seem important to isolate the source of the lack of credibility of regional scale rainfall predictions. There seems to have been little attempt, however, to determine exactly why regional rainfall estimates are unreliable, i.e., to what extent this is simply due to resolution or, alternatively, to inappropriate physical parameterisations. If the inadequate regional rainfall is due to the latter, this may invalidate the use of observed relationships between large-scale flow and regional rainfall in assessing regional rainfall changes.

One approach to this question of the causes of inadequate regional rainfall estimates is to examine the ability of GCMs to simulate the observed relationships between regional rainfall anomalies and large-scale flow anomalies. If a model reproduces the observed relationships between regional rainfall and large-scale flow anomalies, this would suggest that the model's direct estimates of regional rainfall are **not** less credible than its broad scale flow predictions. If so, the use of interpolation methods would not seem to be necessary, except if very detailed (sub-regional scale) predictions were required. If, on the other hand, the model rainfall-flow relationships were extremely dissimilar to those observed, this may reflect inadequate large-scale flow simulations, perhaps due to inadequate parameterisations. In an extreme case, for example, a region which receives most of its rainfall in frontal situations may, in a poor simulation, receive more rainfall from convective events. In such a case use of an interpolation method would be unlikely to improve the credibility of regional rainfall predictions.
The use of interpolation methods may be justified where the model and observed rainfall-flow relationships are similar in character but, perhaps, the model relationships are rather weaker or reveal less detail. In such a case it would seem reasonable to assume that the differences reflected a rather simple influence of the low resolution of the model on the regional rainfall, i.e., it simply meant that the detail required in the regional-scale

rainfall predictions was lacking in the direct model predictions.

Here the ability of two models, one from the Max Planck Institute (MPI) for Meteorology, and the other the climate model of the Australian Bureau of Meteorology Research Centre (BMRC), to simulate large scale circulation anomalies associated with interannual rainfall variability over the Australian continent is assessed. The effect of doubling CO_2 on these relationships is also examined.

The spatial structure of Australian rainfall anomalies has been determined by rotated principal component analysis (Drosdowsky, 1992) from which composites of wet and dry years (for the regions investigated) of seasonal Southern Hemisphere mean sea level pressure (MSLP) anomalies were derived for comparison with the model results. The Southern Hemisphere MSLP anomalies associated with these spatial patterns are examined using 19 years (winter 1972 to Autumn 1991; Southern Hemisphere seasons) of gridded analyses from the National Meteorological Centre, Melbourne. The output from the MPI model (Cubasch et al.,1991), was produced from two 100 year runs (1 X CO_2 and 2 X CO_2). The BMRC model (Colman et al., 1992) produced two 15 year runs similarly of 1 X CO_2 and 2 X CO_2. Seasonal MSLP anomalies for the two GCMs were calculated and standardized using their own means and variances based on the last 80 years in the case of the two MPI runs and all 15 years of the two BMRC runs, in order to calculate t-tests for individual gridpoint significance.

RESULTS

Drosdowsky (1992) demonstrated that Australia can be divided into eight regions within which rainfall variations exhibit similar temporal behaviour. We now examine large scale seasonal MSLP anomalies associated with rainfall fluctuations from composites of model wet and dry years based on some areas from this analysis. Three areas were examined over the Australian continent where large scale seasonal circulation anomalies are known to be related to fluctuations in seasonal rainfall. The three areas are shown in Figure 1.

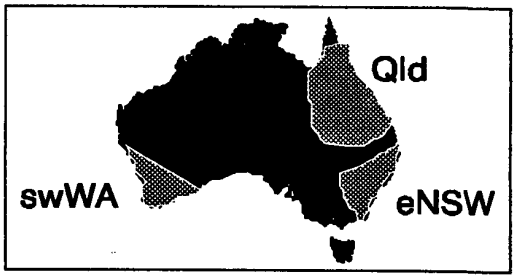


Figure 1. Locations of three areas considered in this study.

Southwestern Western Australia (swWA)

Empirical studies (Wright 1974a,b) have shown large scale circulation anomalies to be related to winter regional rainfall in the swWA. A composite of MSLP anomalies for the three driest observed years shows an area of positive anomalies centred south of Western Australia indicating a lack of baroclinicity and hence minimal frontal activity. For the composite of the three wettest years, the positive anomalies are replaced by negative anomalies, centred southwest of Western Australia indicating rain produced from midlatitude northwesterly winds associated with fronts moving southeast out of the main baroclinic area.

In the BMRC 1 x CO_2 run, MSLP anomaly composites of the three driest and three wettest out of the 15 years for one grid point in the swWA area, match closely the observed MSLP anomaly composites. The composites for the 2 x CO_2 run show similar patterns to the 1 x CO_2 case for both dry and wet composites although the anomaly centres are weaker and the zero anomaly line has shifted slightly poleward.

On the other hand, in the MPI 1 x CO_2 run for a gridpoint in swWA, MSLP anomaly composites of the eight driest and eight wettest years, based on the last eighty years of the run, reveal no similarity with the observed composites in the Australian region.

In the 2 x CO_2 case the MPI composite of the eight driest years produces a different pattern again to the control. As was the case with the control run, the composites in the MPI 2 x CO_2 run are very different to the observed patterns. Similarly, there are inconsistencies for the wet composite of eight years; a weak area of positive anomalies centred in the Great Australian Bight suggesting rain from a weak, presumably moist, east to northeast flow.

Eastern New South Wales (eNSW)

The dry observed composite for eNSW shows the typical spring El Nino pattern for southeastern Australia with a strong negative anomaly centre south of New Zealand and positive anomalies over the continent indicating frontal activity in the baroclinic area is only affecting Tasmania and perhaps southern Victoria. The wet composites

show positive anomalies south of New Zealand in contrast to the negative anomalies associated with the dry composite. There is also a broad band of weak negative anomalies centred over eastern Australia in the wet composite.

For the BMRC model, the composites based on a gridpoint in eastern NSW, exhibit little similarity with the observed. In the dry composite there are positive anomalies in the South Pacific (c.f., negative anomalies in the observed composites) and a weak negative anomaly over eastern Australia replaces the observed positive anomaly. Similarly the wet periods appear to result from weak baroclinicity in the westerlies rather than from mesoscale lows and troughs in the easterlies. The 2 x CO_2 dry composite resembles the observed pattern more closely with negative anomalies southeast of New Zealand and positive anomalies over the continent. However, the wet composite is very poorly represented.

For a gridpoint in eastern NSW from the MPI model the dry MSLP anomaly composite looks more like the observed composite in that negative anomalies are located southeast of New Zealand with positive anomalies over the continent although the anomalies are weak. The wet 1 x CO_2 composite is more credible than the corresponding BMRC model composite with Tasman Sea moisture interacting with midlatitude baroclinicity extending from the negative anomalies further south. The 2 x CO_2 dry composite shows a very weak pattern. There is some similarity with the control in that negative anomalies are south east of New Zealand. The wet composite is similar to the control but with more easterly influence and the zero anomaly line is located further south.

Queensland (Qld)

The observed MSLP anomaly composites for Qld reveal the patterns of the El Niño - Southern Oscillation over the Eastern Pacific and Australian region (Drosdowsky and Williams, 1991). Many studies have documented relationships between Australian rainfall and the Southern Oscillation (e.g. McBride and Nicholls, 1983).

In the dry composite, positive anomalies are centred over northern Australia and negative anomalies centred southeast of New Zealand and stretching into the Eastern Pacific, though much weaker there. Conversely, for the wet composite, negative anomalies are centred over northern Australia indicating an active monsoon trough and positive anomalies through the Eastern Pacific.

For the BMRC model, based on a rainfall gridpoint in northeast Queensland, both the 1 x CO_2 dry and wet composites show weak patterns with almost no resemblance to the observed patterns. The control wet composite shows weak positive anomalies over northern Australia instead of the observed negative anomalies. Similarly, the 2 x CO_2 MSLP anomalies in the dry composite are negative and positive in the wet composite.

The MPI composites based on a rainfall gridpoint in northeast Queensland shows that the dry 1 x CO_2 MSLP anomaly composite is somewhat similar to the observed but the wet composite is not at all like the observed. The 2 x CO_2 composites reveal no similarity to the observed patterns.

DISCUSSION

This examination of the abilities of the BMRC and MPI models in simulating some observed large scale circulation anomalies related to Australian regional rainfall anomalies shows that:

1. The relative skill of the two models in reproducing the observed rainfall-flow relationships varies with area and season. In some areas and seasons the models cannot reproduce even the gross character of the observed relationships; elsewhere they may reproduce them in considerable detail.
2. The model rainfall-flow relationships can change dramatically under enhanced "greenhouse effect" conditions.

There are several implications of this study for the use of GCM's in regional climate prediction. Firstly, the similarity between the observed relationships and those in the BMRC model for the swWA area suggests that model resolution and physical parameterisations do not cause problems in predicting regional climate change in this area. It may not, therefore, be necessary to use interpolation methods to produce regional rainfall change estimates for this area and season, unless much finer detail (sub-regional scale) was required than is feasible using the low-resolution model predictions. However the MPI model was less able to reproduce the observed relationships in this area so, for this model, resolution or physical parameterisations may be preventing the control from simulating the observed relationships. If so, this would cast doubt on the use of the model in either seasonal prediction or climate change studies for this area. The large change in the model relationships in the MPI 2 x CO_2 run compared to the observed may cast further doubt on the use of this model in regional rainfall change prediction for this area in an enhanced "greenhouse effect"

scenario. The differences between the two models' composites cast further doubt on the use of models in rainfall simulation. These differences imply that regional rainfall simulations are not robust and may alter dramatically with changes in the models.

The results from the other two areas in eastern Australia and northeast Queensland were disappointing. The inability of the models to reproduce the relationships which result in rainfall fluctuations in these areas indicates that the direct model rainfall predictions are unreliable. Inability of a model to simulate these rainfall-flow relationships, even crudely, must lead to doubts on the credibility of the direct regional-scale rainfall predictions. On the other hand, the inability of the models to reproduce the observed relationships in the tropics and subtropics would appear to preclude the use of the interpolation approach in predicting regional climate change over Australia. As was noted earlier, it seems reasonable to expect that the observed relationships should be reproduced at least crudely, if interpolation methods are to be used to produce more detailed regional-scale predictions.

In none of the area-season combinations examined here did it appear that there was substantial evidence for accepting that the use of interpolation methods to predict regional rainfall is likely to result in more realistic predictions than the use of direct GCM rainfall predictions. If the model cannot reproduce, at least crudely, the observed relationships, there seems little justification for interpolating the model results to provide finer detail. On the other hand, the inability of the models, with the possible exception of the BMRC model for southwest Western Australia rainfall, to reproduce the observed flow-rainfall relationships in the real world also casts doubt on the use of the direct model predictions of rainfall. If these relationships, which account for much of the rainfall variations, cannot be simulated, then the model does not provide a true representation of the processes producing rainfall at these regional scales. Overall, this study provides little reason to accept that GCMs, either by themselves or in association with interpolation tools, can produce credible simulations of regional rainfall variations or changes. At the very least, more extensive examination of the ability of models to reproduce the relationships between regional rainfall variations and the large-scale flow anomalies is necessary before reliance is placed on regional-scale climate change predicted with these models.

Acknowledgements

This project was suggested by Hans von Storch, who also arranged access to the MPI model results. The BMRC model results were provided by the BMRC Greenhouse Group.

References

Colman, R.A., B.J. McAvaney, J.R. Fraser and R.R. Dahni, 1992: *Slab Ocean and Thermodynamic Sea Ice Models in the BMRC GCM.* BMRC Res. Rep. 30 (pp 35).

Drosdowsky, W. 1992. An analysis of Australian Seasonal Rainfall Anomalies: 1950 - 1990. Part I: Spatial Patterns. *Int. J. Climatol.*, (in press).

Drosdowsky,W., and Williams,M., 1991. The Southern Oscillation in the Australian region. Part I: Anomalies at the extremes of the Oscillation. *J. Climate*, **4**, 619-638.

Cubash, U., K. Hasselmann, H. Hock, E. Maier-Reimer, U. Mikolajewicz, B.D. Santer and R. Sausen, 1991: Transient greenhouse warming computations with a coupled ocean-atmosphere model. *Nature*, (submitted).

McBride, J.L. and N. Nicholls., 1983: Seasonal relationships between Australian rainfall and the Southern Oscillation. *Mon. Wea. Rev.*, **111**, 1998-2004.

Wright,P.B., 1974a: Seasonal rainfall in Southwestern Australia and the general circulation. *Mon. Wea. Rev.*, **102**, 219-232.

Wright,P.B., 1974b: Temporal variations in seasonal rainfalls in Southwestern Australia. *Mon. Wea. Rev.*, **102**, 233-243.

ON STATISTICAL PROBLEMS CONCERNING CLIMATE CHANGE DETECTION

BY G.V. GRUZA

Institute for Global Climate and Ecology, U.S.S.R.

Probabilistic hypotheses in which case the climate change detection problem can be interpreted as a statistical one, are discussed.

The simplest hypothesis of the same kind is a presence of random errors in empirical mean data, such as air temperature, averaged over some regions or time intervals. The random error estimates are presented for the annual air temperatures over the Northern and Southern Hemispheres and the Globe.

Another statistical problem is to detect a climate change **signal** due to *forcing factors* as distinct *from* the natural climate variability. An instability in estimating of relationships between the climate variables and forcing factors (such as greenhouse gases emission, atmospheric transparency, SOI, ENSO and some other indices) are shown.

The necessity of qualification of the empirical inferences uncertainty is substantiated. That is the same as well for empirical climate change forecasts as for the statistical connections.

PROBABILISTIC FORECAST OF EXPECTED CLIMATE CHANGE AS A RESULT OF THE GREENHOUSE GASES ANTOPOGENIC EMISSION

By G.V. Gruza, and E.Ya. Ran'Kova
Institute for Global Climate and Ecology, U.S.S.R.

A new simple statistical procedure to predict a probability of expected climate changes relative to some forcing factors is presented. The assumption is that perfect knowledge is available for the forcing factor behaviour in the future.

The climate changes forecasts are verified for such factors as carbon dioxide and other greenhouse gas concentrations, Southern Oscillation Index and atmospheric transparency as depending on the statistical estimating period.

There is presently only the greenhouse gases changes forecast for the coming decades. So, the climate changes probabilistic forecast was made for the global air temperature (over the Globe, the Norther and Southern Hemispheres) by using the real observed data relative only to the greenhouse emissions and only for the coming 10-20 years.

At present similar forecasts are being constructed for the individual seasons and for different regions.

5TH INTERNATIONAL MEETING ON STATISTICAL CLIMATOLOGY

Sampling Errors in Estimating the Global Temperature

BY

Samuel S. Shen, Department of Mathematics, University of Alberta, Edmonton, Canada T6G 2G1

AND

Gerald R. North and James W. Hardin, Climate System Research Program, College of Geosciences, Texas A&M University, College Station, Texas 77843, U.S.A.

Abstract

Several groups around the world are attempting to estimate the trend in the apparently increasing global average temperature (cf., Hansen and Lebedeff, 1987; Jones et al., 1986; Houghton et al., 1990) since it is now recognized that the increasing concentrations of certain trace gases have the potential to drive the temperature to even higher levels (Houghton et al., 1990). There are three main thrusts of physical science research concerning the so-called greenhouse effect:

1. Estimation of the trends in the earth's climate based upon observations over the last century;

2. Estimation of the trends in the greenhouse gases according to past observation data and future levels based upon assumed greenhouse gas emission scenarios;

3. Formulation and study of climate models which simulate and forecast the trends.

In this paper we concentrate on a very small portion of the first of these problems: for an idealized fluctuation field on the sphere (earth, represented by $\oplus$) with a given network of isolated unbiased point gages, what is the mean squared error in the estimation of the global average value of the temperature field? The problem lies in the fact that the gages are separated by finite spatial distances and that these spatial gaps lead to an inevitable 'sampling error'. Complicating the problem is the fact that the field has correlations from one point on the sphere to another and that these correlations depend on the length of temporal averaging employed. We do not intend to present the final analysis of this problem here, but rather explore some simplified models of the procedure, both the field and the measurement design, in order to better understand the estimation issue.

Our attention is focused on an idealized earth whose surface temperature fluctuation statistics are rotationally invariant on the sphere (the analog of stationarity in time series analysis); we also discuss the limits in which such an analog is faithful. For a given temporal smoothing filter, only one function is necessary to specify such a field: the spatial autocorrelation function or equivalently, the spherical harmonic variance spectrum. We will formulate the mean squared error (MSE) for the problem as a sum over the spherical harmonic indices of a factor dependent on the sampling design and the spherical harmonic variance spectrum. A simple stochastic model (the linear noised forced energy balance climate model) of the temperature field will be developed which is useful as a guide in the problem. Based upon the power spectrum density function derived from this simple model, the ratio of the sampling error to the internal fluctuation of the stochastic temperature field is given by a both physically and mathematically transparent formula. Finally, a few examples will be presented which give a feeling for the magnitude of error to be expected in some typical designs. The problem will then be cast into the perspective of the real problem of taking the earth's average surface temperature within a specified tolerance.

Statistically Optimal Averaging for the Determination of Global Mean Temperatures

Rudolf Weber, Paul Scherrer Institute, CH-5232 Villigen, Switzerland

One of the serious problems in determining time series of global mean surface temperature extending back to the last century is the poor spatial coverage of the observation network (IPCC, 1990, Chap. 7). One method of obtaining estimates of spatial averages of poorly sampled quantities is statistically optimal averaging as proposed by Vinnikov et al. (1990). The main idea of this statistically optimal averaging is to get an estimate of the spatial average which is unbiased and whose mean square difference (MSD) from the true spatial average is minimal. These goals are achieved by assigning a suitable weight to each station in the observation network.

The purpose in this work was twofold. First the method of statistically optimal averaging was tested with some climate model data to see possible limitations and errors of the method. Second the method was applied to real data by taking for each year the actual available set of stations and not restricting the number of stations as it was done in Vinnikov et al. (1990) where only that subset of 566 stations was taken into account, which had measurements in a specific year.

The first dataset analyzed consists of monthly means of 991 hpa temperature from 120 perpetual January realizations (Branstator, 1990; Madden et al., 1991) of the National Center for Atmospheric Research Community Climate Model available on a R15 grid. For the years 1860 through 1979 the subsets of gridpoints with actual observations were taken from Madden et al. (1991) where data from land stations and the Comprehensive Ocean-Atmosphere Data Set were used to determine which gridpoints have observations in a specific year. As the model data are available at all gridpoints for all realizations it is possible to compute the true spatial average for all ensemble members.

Statistically optimal averaging is closely related to optimal interpolation (Gandin, 1965) and is based on the following principles. Suppose a certain ensemble of data, typically for a number of years, is available on a grid on the sphere. Let $X_i(\tau)$ denote the value at gridpoint i for the ensemble member τ ($i = 1,\ldots,M; \tau = 1,\ldots,T$). The true spatial mean $\mu(\tau)$ is then given by

$$\mu(\tau) = \sum_{i=1}^{M} g_i X_i(\tau) \,, \qquad (1)$$

where the g_i's are normalized geometrical weights on the sphere; here the Gaussian weights of the R15 grid are used. If data is only available on a subset (of size N) of the full globe-covering grid a poor mean $p(\tau)$ can be defined by

$$p(\tau) = \sum_{j=1}^{N} g_{i_j} x_{i_j}(\tau)/G \,, \qquad (2)$$

with $G = \sum_{j=1}^{N} g_{i_j}$ a normalization constant. As here renormalized geometrical weights are used, we term this mean also geometrical mean.

A statistically optimal average $o(\tau)$ is introduced by

$$o(\tau) = \sum_{j=1}^{N} w_j(\tau) X_{i_j} \,, \qquad (3)$$

where the weights w_j are determined by the requirement that $o(\tau)$ is an unbiased estimate of the true mean, $\bar{o} = \sum_{\tau=1}^{T} o(\tau)/T = \sum_{\tau=1}^{T} \mu(\tau)/T = \bar{\mu}$, and that it has a minimal MSD with respect to the true mean

$$D^2 = \sum_{\tau=1}^{T} [\mu(\tau) - o(\tau)]^2 = \text{minimal} \,. \qquad (4)$$

The weights w_j are determined by the variational equations $\partial D^2/\partial w_j = 0$ which leads to a linear system of equations

$$\sum_{j=1}^{N} w_j C_{jk} = C_{\mu k} \,, \; k = 1,\ldots,N \,, \qquad (5)$$

where $C_{jk} = \sum_{\tau=1}^{T}(X_{i_j}(\tau) - \bar{X}_{i_j})(X_{i_k}(\tau) - \bar{X}_{i_k})/T$ is the covariance between gridpoints i_j and i_k and $C_{\mu k} = \sum_{\tau=1}^{T}(\mu(\tau) - \bar{\mu})(X_{i_k}(\tau) - \bar{X}_{i_k})/T$ is the mutual covariance of the values at gridpoint k and the true mean. Here, the weights were additionally normalized to one, $\sum_{j=1}^{N} w_j = 1$, in order to be able to handle data with unknown norm (Gandin, 1965). The additional condition can be incorporated into the linear system of equations by means of a Lagrange multiplier.

If data on the full globe-covering grid is available for some ensemble, let's think of the last two or three decades where spatial coverage of the observation network is high on the globe, one may use the sample covariance matrix and mutual covariance vector of this ensemble to determine the optimal weights. Statistically optimal averaging is then equivalent to multiple linear regression. By using the model data and different ensemble sizes T the optimal weights were determined for gridpoints of random subsets of varying size N. The MSD of the optimal mean and the true mean was computed for the whole series of 120 perpetual model Januaries and not only for the smaller ensemble of length T which determined the weights. As long as $N < T$ the MSD remains small which means that the weights can be reasonably well used for other ensembles than the one which determined the weights. For $N > T$ the MSD increases dramatically, hence no extension to other ensembles, or time periods, is possible.

As is known from statistical theory (Kendall, Stuart, Ord, 1987), a population covariance matrix is nonnegative definite, hence any meaningful sample estimate should have the same property. Under some assumptions about the random variables it was shown (Muirhead, 1982 and references therein) that the sample covariance matrix is nonnegative definite if $T > N$, which is equivalent to the statement that it has no negative eigenvalues. In our case the assumptions are not fulfilled but the computation of the eigenvalues of the sample covariance matrix confirms that negative eigenvalues occur for $N > T$ only. In this case the method may fail and can give meaningless weights and averages. Statistically optimal averaging does however not necessary break down if the sample covariance matrix is not positive semidefinite. If all grid points were used the covariance matrix was indefinite but the optimal weights still become just equal to the geometrical ones and the MSD between optimal and true mean was within numerical accuracy.

In general the true mean is not known for any ensemble and the data is not continuously available on all points of a given subset. The covariance matrix and the mutual covariance vector can then be estimated by means of the spatial correlation function (Gandin, 1965), one of the essential tricks of optimal interpolation. Here the same procedure was used as in Vinnikov et al. (1990) where the covariance matrix and vector are approximated by means of analytical spatial correlation functions $r(\rho_{ik})$, which are assumed to be homogeneous and isotropic within latitude belts of 30°, and by zonally averaged ensemble standard deviations $\bar{\sigma}_i$. The covariance between gridpoints i and k is given by $C_{ik} = r(\rho_{ik})\bar{\sigma}_i\bar{\sigma}_k$. With this procedure it is not obvious which is the appropriate ensemble size T. Therefore, the definiteness of the covariance matrix, formed with the above mentioned analytical functions, was checked for the different latitude belts and the subgrids of the years 1860 to 1979 by computing their eigenvalues. Negative eigenvalues were found in the latitude belt $60-90°$ south, thus indicating that the

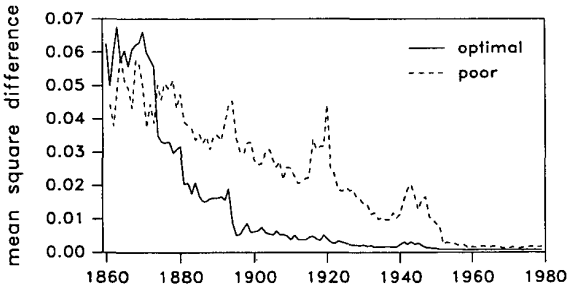


Figure 1 Mean square difference of the statistically optimal mean (solid line) and the poor mean (dashed line) with respect to the true mean for an ensemble of 120 climate model January averages. For each year only the gridpoints with actual measurements are used for the calculation of poor and optimal mean.

method can fail in this region.

For each latitude belt the MSD can be computed by use of (4,3,1) which gives the MSD as a function of the optimal weights w_j (Vinnikov et al., 1990, Eq. (3)) and it is therefore inherent to the optimal averaging method. As the model data allow the computation of the true global mean, an independent comparison of true and optimal mean can be carried out, too. With exception of the latitude belt $60-90°$ south the two estimations of MSD are in good agreement in each latitude belt. In the latitude belt $60-90°$ south the inherent MSD obtained from (4,3,1) remains reasonably small, but gives negative values for some arrays of gridpoints, thus indicating problems with optimal averaging in that latitude belt; the independent MSD gives large values as the averaging procedure fails.

Restricting to latitudes north of 60° south, the MSD between optimal and true mean can be computed for the whole globe (actually the optimal mean was computed for latitudes north of 60° south and the true mean for the whole globe). In Fig. 1 the MSD of the optimal mean is shown as solid line as a function of the year whose subset of gridpoints of available data was used for the optimal averaging procedure. For comparison the MSD of the poor mean (2) is shown as dashed line. Except for the observation networks before 1880 the optimal mean has a fairly lower MSD than the poor mean.

If a linear least square fit is done with the values at a single gridpoint only, the gridpoint with highest correlation with the true mean was chosen (105° W, 64.4° N), the corresponding MSD results to be ≈ 0.018. Different ensemble sizes as small as $T = 10$ were used to determine the weight and showed no significant differences in the resulting MSD.

The method of statistically optimal averaging was applied to the combined land air and sea surface dataset (Parker and Jones, 1990) available from the National Center for Atmospheric Research archive. In this database monthly mean values are available on a 5° by 5° grid for the period from 1854 to 1989. For each month analytical spatial correlation functions were determined by taking all pairs of grid points into account with at least 30 years of common data. Within each of the six latitude belts defined as above and for each year the covariance matrix of the subset of gridpoints with actual data was computed and diagonalized. No negative eigenvalues were found, i.e. all sample covariance matrices are positive semidefinite.

For each month the hemispheric and global means were determined and these values were then averaged to the time series of annual means

	1870-1989	1890-1989
North	0.49 (0.39)	0.57 (0.47)
South	0.44 (0.43)	0.59 (0.53)
Globe	0.47 (0.41)	0.58 (0.50)

Table 1 Linear trends (in K/100 years) of annual mean temperature over different periods for Northern and Southern hemisphere and the globe, in brackets the corresponding values of the IPCC report (1990).

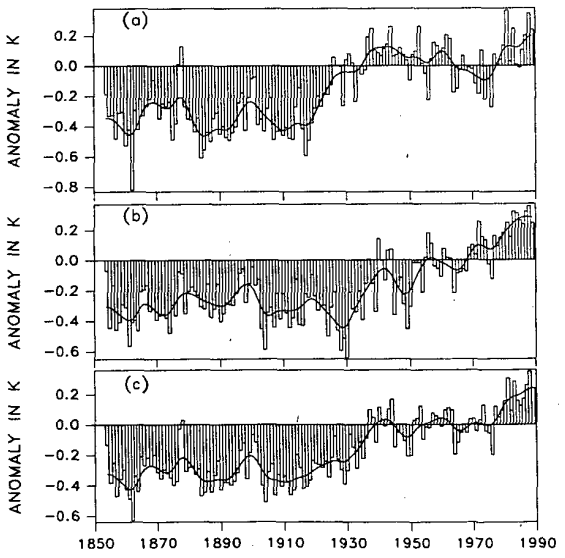


Figure 2 Annual means of the statistically optimal averages for the northern hemisphere (a), the southern hemisphere (b) and the whole globe (c). The smoothed curves are obtained by a low pass binomial filter with 21 terms.

which are shown in Fig. 2, where the global and hemispheric means are plotted for the entire range from 1854 to 1989. A comparison with the corresponding figures of the IPCC report (Fig.7.10) shows a good agreement of the overall behaviour of the time series. In detail, however, there are substantial differences, e.g in the southern hemisphere around 1930 where our values are about 0.2 K lower. By taking the period from 1960-1989 as ensemble with known true mean the MSD of the optimal means are calculated for each month and each year. The mean square error of the annual mean is approximated by the average of the twelve monthly MSD's. In Fig. 3 the root mean square error of the annual mean is shown.

Fitting a monotonic linear trend to the time series of hemispheric and global means the following slopes are obtained (Table 1). The different time periods have been chosen in order to facilitate a comparison with values given in the IPCC report (1990), which are given in brackets in Table 1. The linear trends computed in this paper are higher than the ones reported earlier. The largest differences occur for data in the northern hemisphere where the values of this paper are 0.1 K/100 years larger than the ones of the IPCC report.

In summary it has been shown that statistically optimal averaging can give a significantly better estimate than simple geometrical averaging. The method can fail, however, and it should be tested in all applications whether the sample covariance matrix is positive semidefinite. The application of the method to a gridded dataset of observed values results in global and hemispheric means which are at least for some periods quite different than the ones obtained with other spatially averaging procedures. The linear trends are slightly larger than previously published ones. Which values are closer to the true ones can of course not be answered. The conclusion to be drawn is that there is still a large amount of uncertainty in the determination of global means from such poorly sampled data.

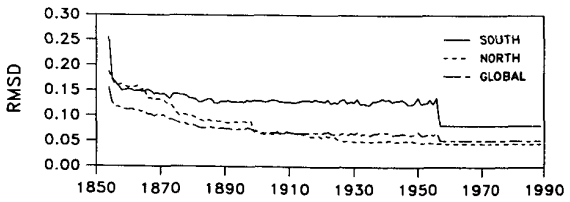


Figure 3 Estimated root mean square error of the global mean (in Kelvin) as obtained by statistically optimal averaging.

Acknowledgement. Most of this work was carried out during a visit at the National Center for Atmospheric Research in Boulder, sponsored by the National Science Foundation.

References

IPCC, 1990: *Scientific Assessment of Climate Change.* IPCC WGI, WMO, UNEP. J.T. Houghton, G.J. Jenkins and J.J. Ephraums, Eds., U. of Cambridge Press. 365 pp.

M.G. Kendall, A. Stuart, K. Ord, 1987: *Kendall's Theory of advanced Statistics, Vol. 1, Distribution Theory.* Oxford University Press, New York. 604 pp.

R.A. Madden, D.J. Shea, G.W. Branstator and J.J. Tribbia, 1991: The Effects of Imperfect Spatial and Temporal Sampling on Estimates of the Global Mean Surface Air Temperature, *in preparation.*

R.J. Muirhead, 1982: *Aspects of Multivariate Statistical Theory.* John Wiley and Sons, New York. 673 pp.

D.E. Parker and P.D. Jones, 1991: Global Warmth in 1990, *Weather*, 302-311.

Vinnikov, K., P. Ya. Groisman and K.M. Lugina, 1990: Empirical Data on Contemporary Global Climate Changes (Temperatures and Precipitation), *J. Climate*, 3, 662-677.

The likelihood of climate change: A methodology to assess the risk and the appropriate defence.

Harvey Stern
Special Services Unit
Bureau of Meteorology
Box 1289K, Melbourne, 3001
Australia.

1. Introduction.

"The analogy between meteorology and astronomy is often made...There is a closer resemblance, to my mind, between meteorology and economics. Both deal fundamentally with the problem of energy transformations and distribution - in economics, the transformation of labour into goods and their subsequent exchange and distribution; in meteorology, transformation and distribution of the energy received from the sun. Both systems are subject to extremely capricious external influences..." (Sutton, 1951).

"Economists are sometimes challenged to cite a discovery from economics that is both true and surprising. For many years, the principle of comparative advantage was the best example. An equally good reply, thanks to Messrs Black and Scholes (Black and Scholes, 1973) is the theory of option pricing" (The Economist, 1991).

The beautiful connection between these two quotes will become apparent during the course of the present paper as one realises that

. the basis for costing (economics) of climate change (meteorology) is options pricing theory (economics); and,

. that theory originates, in part, from heat transfer theory (meteorology) - see Kreith and Bohm (1986).

The Australian government recently established an Industry Commission (Industry Commission, 1991) to study the costs and benefits of reducing greenhouse gas emissions.

The Industry Commission concluded that "current knowledge does not definitively establish that an enhanced greenhouse effect with adverse implications exists ...(and that)...there are many major uncertainties, and it will be another decade before it is possible to say whether or not observed recent climate change has been driven by anthropogenic influences".

And, the costs of premature action may be great. In the News Release which accompanied the release of its report, the Industry Commission warns that, for Australia, "...achieving a 20 per cent reduction in emissions from 1988 levels (the Toronto target) would reduce national output by 1.5 per cent".

The issue of global climate change has aroused considerable debate in the scientific community. Views on trends in global climate have ranged over the years from those who forecast a substantial (economically significant) temperature rise, to those who see the present apparent upwards trend in mean temperature as a part of the natural fluctuations, and finally, to those who have predicted a substantial trend to lower temperatures.

The physical bases for these predictions are as varied as the explanations advanced for previous climate shifts. These explanations cite, among other causes, changes in the earth's axis of rotation and its orbit around the sun, aerosols emitted by increased volcanic activity, carbon dioxide and other greenhouse gas emissions, as being responsible for climate variations. More recently, global climate shifts have been attributed to variations in the output of the sun's energy, some correlation of the mean temperature to the period length of the sunspot cycle being noted.

That no numerical global climate model (GCM) referred to in the literature has ever been reported as correctly replicating the "warming-slight cooling-warming" scenario observed this century, underlines how fragile is our understanding of the complexity of the processes involved.

These issues make necessary the development of an appropriate risk-return model which would enable the calculation of the appropriate strategy to adopt both from a micro-economic and macro-economic standpoint.

2. Methodology.

Firstly, regard the global mean temperature

GMT data series)
. Interest rate = 0% (assuming that the only money which changes hands is that associated with variation margins).

Utilising the Black and Scholes (1973) call option formula, as modified for future style options (Gastineau, 1988), the calculation yields 0.15 - see also Black (1976).

So, for protection under the aforementioned assumptions, the full cost of protection is $15 for every $100 of industrial growth to be achieved over the next 100 years.

But if, for example, the effect of reducing industrial output by 1.5% (referred to earlier) is compounded annually, after 100 years (and if one assumes a 3% annual growth rate otherwise) we are

. looking at a level of annual industrial output 77% below that which otherwise would have been produced after 100 years; and,
. furthermore, the total industrial output sacrificed under these assumptions would be 47% of that otherwise achieved over the full period.

The important issue is not so much the validity (or otherwise) of the many assumptions that are made, but that they form the basis for opening up a quantitative approach to determine cost and appropriate action.

And, there is one further assumption that is implied by the methodology. This is the assumption that taking actions to reduce greenhouse gas emissions will, indeed, arrest the global warming, if it is, in fact, occurring at all.

Example 2:
Protecting against the risk of decreasing value of a company likely to be adversely affected by global warming (e.g. a manufacturer of ski equipment).

Assumptions -

(1) That long term GMT is 288K (approximately 15°C) and departures are regarded as being from that figure (e.g. -0.21 represents 287.79K)
(2) That the insurance is calculated for protection against a decrease in the value of a company, which occurs as GMT rises.
(3) That the value is unaffected by global warming as the GMT rises, until the temperature reaches a departure from the long term of 1.34°C. (1.00°C above the CDIAC 1988 value). A temperature increase from this point is assumed to adversely affect value, causing it to decline in a linear manner as GMT rises further to +2.34°C, at which point the value is reduced to zero. Continued rise in GMT from this point has no further effect upon the company's value as it cannot decline in value below zero.
(4) Protection is required for a period of 100 years.

Calculation -

This is equivalent to calculating the difference between the cost of two American call option contracts on the value of the Futures GMT contract with the following characteristics:

First contract -

. Spot = 288.34K.
. Strike = 289.34K.
. Volatility = 0.0069% (based on the CDIAC GMT data series)
. Interest rate = 0%

Utilising Black and Scholes call option formula, as modified by Black for futures contracts, the calculation yields 0.15.

Second contract -

. Spot = 288.34A.
. Strike = 290.34A.
. Volatility = 0.0069% (based on GMT data series)
. Interest rate = 0%

Utilising the Black and Scholes (1973) call option formula, as modified for future style options (Gastineau, 1988), the calculation yields 0.03 - see also Black, 1976.

So, the cost of insurance is the cost of the first contract minus the cost of the second contract, namely 0.12, or 12% of the future value of the company. Note again that no money changes hands initially, and it is possible that only at the end of the options' life will settlement occur.

(GMT) in the same manner as one would a financial commodity futures contract and value it accordingly (Black, 1976; Gastineau, 1988). The theoretical value of such a futures contract is a function of 'holding costs minus income generated by the underlying financial instrument'. To illustrate, the value of a share price index (SPI) futures contract is a function of 'interest rate minus dividend return from the underlying shares'.

The above process yields theoretical valuations (on this basis the theoretical value of a GMT futures contract will equal the current GMT).

The operation of financial markets then leads to departures from the theoretical value of futures contracts, which are a function of anticipated movements in the value of the underlying.

Secondly, assume that GMT futures contracts are available to be bought and sold and that associated put and call option contracts (Gastineau, 1988) are available to be written or taken, and so alter the risk-return characteristics associated with the GMT contract.

The strategy, therefore, is to establish the economic consequences of movements in the GMT, these economic consequences being applied across the complete range of scales; i.e. from the global economy down to the smallest company (e.g. a ski equipment manufacturer).

These economic consequences can be replicated in a combination of GMT futures contracts and an associated set of 'written' and 'taken' put and call option contracts at various strikes and expiry dates.

The use of the methodology presented derives from an assumption that the two time series (SPI and GMT) follow a "random walk" principle (Cheng and Deets, 1971 and Gordon, 1991), that is, the value of the next element in a series is independent of preceding values.

Alternatively, it may be argued that the market would evaluate the cost of insurance against global warming by taking into account the spread of predictions from GCMs. However, the running of GCMs is quite analogous to the efforts which are undertaken to forecast the movement of the SPI. Rarely do the contracts significantly vary in value from the theoretical.

3. Illustrative examples

Two examples, illustrating the methodology, are now presented.

Example 1:
Protecting against the risk of diminishing industrial output associated with global warming.

Assumptions -

(1) That long term GMT is 288K (approximately 15°C) and that the departures given by the Carbon Dioxide Information Analysis Center (CDIAC, 1990), are regarded as being from that figure (e.g. -0.21 represents 287.79K).
(2) That the insurance is calculated for protection against a diminishing in the globe's increase in its industrial output that would have occurred except for adverse effects of global warming.
(3) That industrial output is unaffected by global warming as the GMT rises, until the temperature reaches a departure from the long term of +1.34°C (or 1.00°C above the 1988 CDIAC value). A temperature increase from this point is assumed to adversely affect industrial output, causing it to decline in a linear manner as GMT rises further to +2.34°C, at which point the annual rate of increase in industrial output is reduced to zero. Continued rise in GMT from this point is assumed to lead to an adverse effect increasing at the same rate. So, by the time the GMT departure is at +3.34°C, the rate of decline in global industrial output is equivalent to the current rate of increase. Note that these scenarios of GMT change are consistent with GCM output (Bureau of Meteorology, 1992).
(4) Protection is required for a period of 100 years.

Calculation -

This is equivalent to calculating the cost of an American call option contract on the value of the Futures GMT contract with the following characteristics:

. Spot = 288.34K.
. Strike = 289.34K.
. Volatility = 0.0069% (based on the CDIAC

4. Summary and conclusions.

A methodology for calculating the appropriate cost of defending against the possible onset of global warming has been presented.

The calculation procedure is based on the premise that current knowledge does not definitively establish that an enhanced greenhouse effect with adverse implications exists.

Under this premise and utilising the results of other work on the time series which records the global mean temperature record, the assumption of "random walk" with regard to global mean temperature is regarded as valid.

One then borrows from financial markets, the behaviour of which are also regarded as "random walk", to value hedging and speculative instruments that might equally apply to climate fluctuations.

In valuing the climate fluctuation hedging and speculative instruments, wide-ranging assumptions are made. But, the methodology does provide a tool whereby the cost of the risk faced can be determined (whether it is in the case of determining that risk on a global scale, or on a company specific scale).

Use of these financial instruments would lead to those concerned being compensated provided they were on the correct side of the contract. Conversely, those on the wrong side of the contract would have to provide that compensation.

5. References.

Black, F. 1976. The pricing of commodity contracts. Journal of financial economics (Sept.).

Black, F. and Scholes, M. 1973. The pricing of options and corporate liabilities. Journal of political economy (Sept.).

Bureau of Meteorology, Australia. 1992. Weather and climate services for sustainable development. p.10.

CDIAC, 1990. Trends '90. A compendium of data on global change. The carbon dioxide information analysis center (US).

Cheng, P. and Deets, M. 1971. Portfolio returns and the random walk theory. Journal of finance (March).

Gastineau, G. L. 1988. The options manual. 3rd. edn. McGraw-Hill Book Co.

Gordon, A. H. 1991. Global warming as a manifestation of a random walk. Journal of climate (June).

Industry Commission, 1991. Costs and benefits of reducing greenhouse gas emissions. Report No. 15, Vol. 1, Aust. Gov. Pub. Service.

Kreith F. and Bohm, M. S. 1986. Principles of heat transfer. 4th edn. Harper and Row, NY.

Sutton, O. G. 1951. Mathematics and the future of meteorology. Weather (Oct.)

The Economist, 1991. Of butterflies and condors. The Economist (16 Feb.).

6. Acknowlegements.

The author thanks Dr. R. R. Dahni and Messrs J. R. Fraser and L. W. Logan for their helpful comments.

Detecting Climatic Change Along the Longest Belt of Largest Population Density

By

L.S. Hingane
Indian Institute of Tropical Meteorology, Pune-411008,INDIA

ABSTRACT

Socio-economic and biophysical impacts on the tropospheric level of ozone is assessed along the longest belt of largest population density of the globe. This belt which lies along the foot of the great Himalayas mountain in Asia constitute 1% of the World's land area and supports 10% of its population of both human and animal, tremendous amount of Ch_4, CO, CO_2 and NO_2 is emitted continuously into the atmosphere through their inadvertent activities. Keeping in view the importance of the subject, the anthropogenic production of these gases is estimated over the region of interest. Production of CH_4 by paddy field and enteric fermentation of herbivorous animals where as that of CO, CO_2 and NO_2 from biomass and fossil fuel burning is estimated.

The values obtained above are used to define the lower boundary condition in a 1-D photochemical dynamical model (fundamental equation of continuity with vertical transport due to eddy diffusion only) to see the effect of these gases on the tropospheric level of ozone. The flux conditions (f_z mol $cm^{-2}s^{-1}$) according to the emission rate (I_z) of Ch_4 and that of CO obtain as above and deposition velocity (w_d = 0.5 cms^{-1}) for O_3 are prescribed at the lower boundary at surface such as

$$f_z = -w_d.n_i + I_z$$

photochemical equilibrium condition is assumed for the upper boundary at 15 kms. Model consists 31 species of oxygen-hydrogen-nitrogen-carbon and chlorine atmosphere and incorporate about 150 chemical and photolytic reactions. The reflection of solar radiation and Rayleigh scattering are included. The numerical technique applied here

utilizes central differences in spatial variables and a leap-frog method in temporal variables with a modified forward differences scheme at the first half-time step. Model does not include heterogeneous reactions. The 1-D model is run at the interaction of 30°N and 80°E, for winter season as the residence time of air during this season is relatively more with surface flux condition and without surface flux condition. Inspite of several limitations of the model, results clearly show the anthropogenic effect on the tropospheric ozone. The level of tropospheric O_3 is certainly enhanced due to surface emission of CH_4, NO_x and CO.

The population density of the belt mentioned above remained to be substantially high right from the beginning of the century. The living habits of the population also had been not changed significantly. Therefore, keeping above facts in mind it is thought of interest to discuss the long term surface air temperature in the light of tremendous amount of anthropogenic emission of heat as well as greenhouse gases emission in the atmosphere with the assumption that some acknowledgment of the reciprocal interaction between them would be an appropriate additions. Surface air temperature data of fairly widespread stations over this belt for the period 1901-1990 have been used to examine the nature of variation. The long-term variation in the temperature have been evaluated by linear trend. The results clearly indicate a definite warming trend with its value of 0.56 C/100 yr (significant at 1% level) in the mean winter season (December to February) temperature series. As residence time of air over this region during winter is more, the above signal definitely give the sense of climatic change over one of the most anthropogenically perturbed.

Window technique for detection of CO_2 effect

Tamás Szentimrey
Hungarian Meteorological Service,
P.O. Box 38, H-1525 Budapest, Hungary

Abstract. Climatic characteristics are affected by various systematic and occasional impacts: besides the changes in the observing system (locations of the stations of meteorological network, instruments, observing procedures), the possible local-scale and global natural and antropogenic impacts on climatic conditions should be taken into account. Perhaps one of the most important question in connection with the antropogenic impacts there is the Greenhouse effect, and chiefly the effect of increasing of CO_2 concentration to the global annual mean temperature. The essence of the problem, that there is some significant relationship between the CO_2 concentration and the global annual mean temperature, but the relation of cause and effect can't be verified unambiguously.

In this paper a new statistical method - the afore named window technique - is presented and applied to exam whether the statistical relationship of CO_2 concentration and the global annual mean temperature is changing or not. The aim of the examination is to get as much significant information as possible about the characteristics of the mentioned statistical relationship in time. For the evaluation of the statistical relationship in time, a set of hypothesis about the validity of the relationship attributes should be jointly tested by means of statistics on a given significance level. The latter requirement is essentially equivalent to the total (mutual) independence of statistics, and in this respect, the application of the orthogonal methods is advantageous in the case of normality.

The window technique is based on both the linear regression (in our case it is a bivariate regression) and the orthogonal series expansion, and the specific purpose of this proposed method is to estimate the possible changes of the linear relationship of the examined elements within certain time intervals (windows) on a given significance level. This method makes possible the joint examination of $[n/2]-1$ windows, where n is the length of the observed time series, and $[n/2]$ denotes the integer part of $n/2$, i.e. we can get $[n/2]-1$ independent information about the linear relationship of the examined elements within these windows in the case of normality.

This method is applied for global and hemispheric data series of the annual mean surface temperature alike.

METHOD OF 'SLICES' TO ESTIMATE REGIONAL FEATURES OF THE GLOBAL WARMING AT EXTRATROPICAL LATITUDES

MIKA J., Hungarian Meteorological Service
H - 1525 Budapest, P.O. Box 38, Hungary

Method of 'slices' is developed to investigate connections between regional climate elements and two hemispherical characteristics, ie. the average temperature and air temperature contrast between continents and oceans. Hemispherical and regional time series are divided into subperiods of uniform length and multiregression analysis is fulfilled for the averages of these subperiods. A 'quasi-equilibrium' division from 13 years long sequences with neglible trends in temperatures above the continents and 5 'non-equilibrium' divisions of different length, irrespectively of trends are defined. The range of hemispherical mean temperature as covered by the 'slices' in 1881-1980 is 0.5 K.

The main feature of sea-level pressure in the Atlantic-European area is a weakened zonality in the winter half-year parallel to a hemispheric warming. Intensity of North-Atlantic pressure-dipole exhibits a positive connection with the continent-ocean contrast.

Temperature fields of the areas to the North from 30° N in semiannual resolution can be characterized by large positive relative sensitivity around Greenland, concerning the hemispherical mean temperature. In addition 3-4 areas of significant sensitivity are identified, that are nearly symmetrical around the Pole in the winter half-year, but not at all in the summer one. Significant regression coefficients are generally greater in quasi-equilibrium resolution than in non-equilibrium ones. Relative sensitivity fields of standard deviation for temperature, demonstrating the change in the variability are also presented.

1. THE METHOD

The aim of this study is to promote regional climate scenarios by developing a method, which is a generalization of that applied by Lough et al. (1983). In the referred study local climate averages for two appropriate 20 years long periods are compared.

We use each year from 1881-1980 so, that original regional and hemispherical series are divided into subperiods of equal lengths and regression analysis is fulfilled for time averages of the subperiods. A 'quasi-equilibrium' division (EQ) from periods with no trends in the continental temperature, and five 'non-equilibrium' ones (NON-EQ) with consequent 5, 9, 13, 17 and 21 years are defined.

Regional features of the global warming are estimated by a three variable linear regression in connection with the annual mean temperature of the Northern Hemisphere ($\langle T \rangle$) and the air temperature contrast above the continents vs. the oceans (ΔT).

$$Y = Y_0 + (\delta Y/\delta \langle T \rangle) \langle T \rangle + (\delta Y/\delta \Delta T)\ \Delta T$$

Elements of the samples are time averages for the above defined quasi-equilibrium or non-equilibrium periods. Regression coefficients are found by the method of least squares. As correlation coefficients are not realistic according to the non-parametric regression model, T-tests of the regression coefficients at 95 and 80 % levels are made.

Hemispherical mean temperature and continent-ocean contrast are derived from air temperatures above the continents (T_C: Jones et al., 1986) and oceans (T_O: Folland et al., 1984) by linear weighting according to the areas of the two domains. Correlation of the two introduced hemispherical variables is neglible for all applied divisions during the investigated period, which makes possible to avoid the problems of multicollinearity of the physically plausible variables.

The Russian data set (eg. Vinnikov, 1986) for pressure and temperature fields is used for regional series. In the latter case only 90 yrs. (1891-1980) were available in necessary completeness. Series after 1980 can be used for verification of the connections established from the basic period.

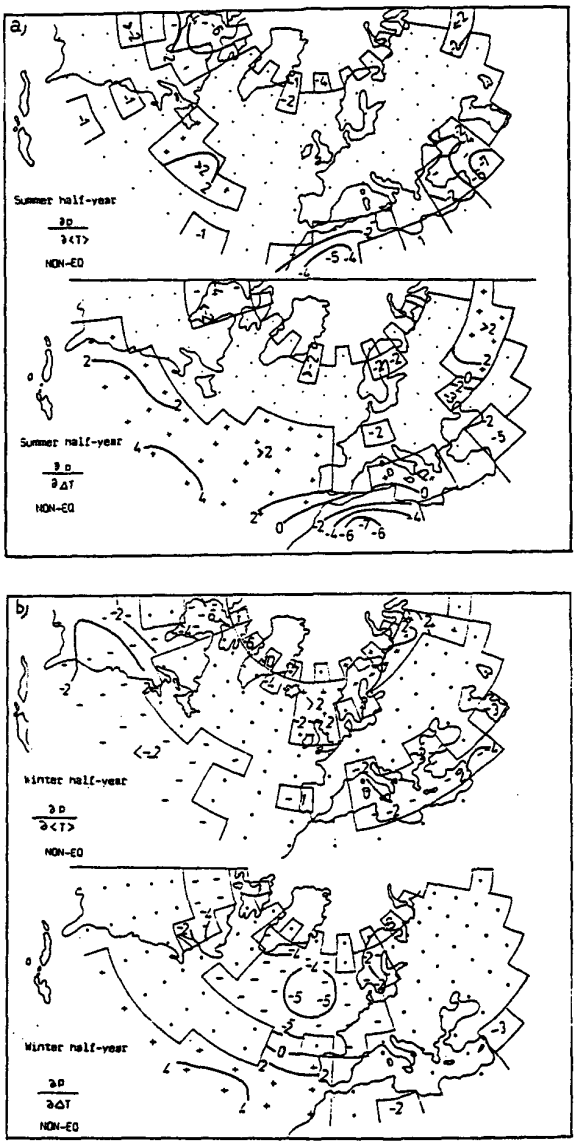


Figure 1 Regression coefficients of sea-level pressure comparing to the hemispherical mean and continent-ocean contrast in the summer (1 a) and winter (1 b) half-years

2. RESULTS FOR AIR PRESSURE FIELDS

In the winter half-year (Fig. 1a) meridional pressure gradient between southern and northern parts of the Atlantic-European area significantly decreases parallel to a hemispherical warming. At the same time hemispherical air temperature contrast between continents and oceans coincidences with a strengthened pressure-dipole in the Northern Atlantics. Fields of summer half-year coefficients are less characteristic (Fig. 1b).

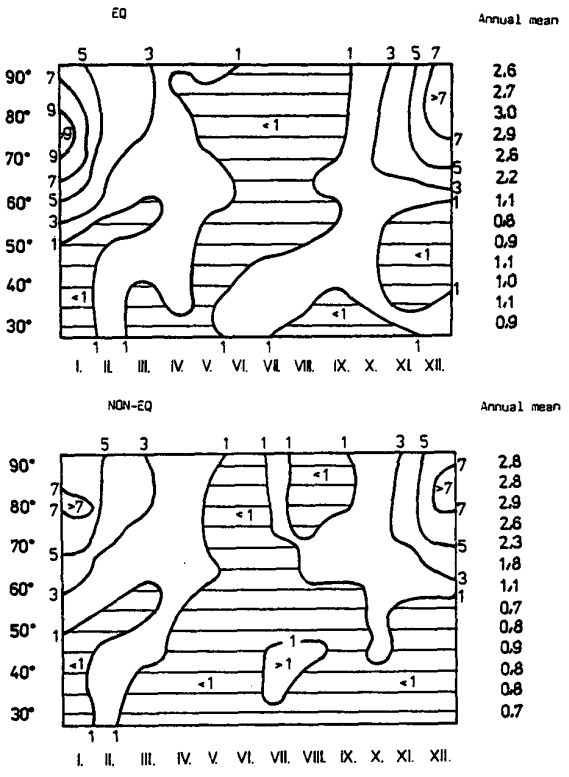


Figure 2. Regression of zonal mean temperatures, compared to hemisperical indicators

3. RESULTS FOR TEMPERATURE FIELDS

3.1 Sensitivity of the averages Regression coefficients, derived from the quasi-equilibrium division are usually more significant than those calculated from non-equlibrium ones (Table I.).

Table I. Percentage of regression coefficients, significant at the 80 % level, comparing to all investigated divisions and gridpoints

			90-65	60-45	40-30°N
		Division			
δT	summer half-yr	EQ	47	49	50
		NON-EQ	58	40	50
δ<T>	winter half-yr	EQ	97	61	49
		NON-EQ	92	38	49
δT	summer half-yr	EQ	36	24	31
		NON-EQ	38	19	8
δ(ΔT)	winter half-yr	EQ	80	58	35
		NON-EQ	35	30	20

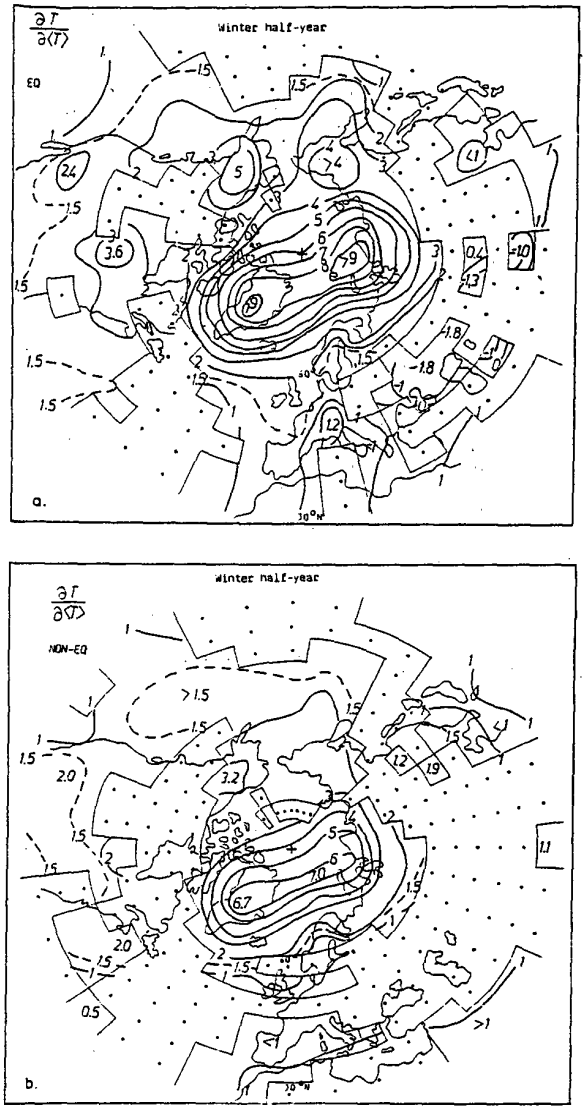


Figure 3 Coefficients of regression for the winter half-year from the quasi-equilibrium division (3a) and from the 5 non-equilibrium (3b) divisions in average. In gridpoints indicated by ⊓ just a minority of the coefficients is significant even at 80%

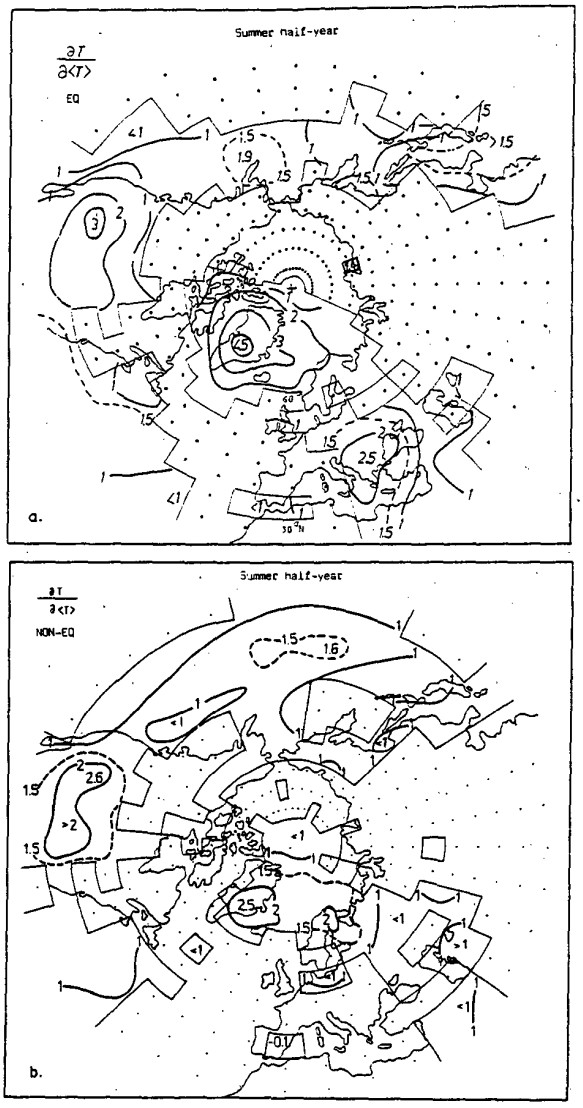


Figure 4 The same as Fig. 3 for the summer half-year

For the temperature field a maximum of the relative sensitivity as compared to the hemispherical mean temperature changes is developed over polar latitudes in the winter half-year and an isolated maximum over Greenland in the summer half-year (Figs. 4a,b). Regression coefficients are generally larger in the quasi-equilibrium resolution than in the non-equilibrium ones in both half-years, in accordance with Table I. The difference is bigger in the cold half-year.

Approximate similarity of the relative sensitivity fields in quasi-equilibrium and non-equilibrium divisions is a positive experience, as it contributes to apply the GCM-outputs of equilibrium experiments in regional scenarios. But due to differences in coefficients and their significancy, the application may be only qualitative.

3.2 Sensitivity of the standard deviations

Main feature of these fields is an irregular, frequently meridional run of isolines, which separate areas with different tendencies as compared to a hemispherical warming (Fig. 5).

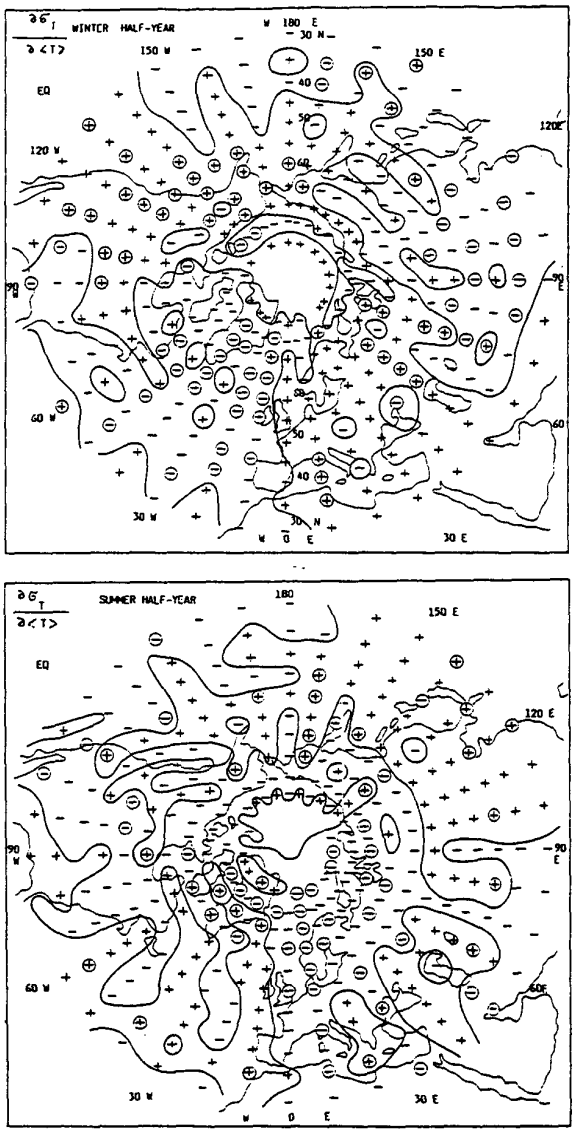


Figure 5 Signs of relative sensitivity for standard deviations of temperature. Grid-points, significant at 80 % are encircled

4. DISCUSSION

Papers adressed to the estimation of regional features of a global warming (see eg. Webb and Wigley, 1985; or Vinnikov, 1986 for references) are based on the hypothesis, that changes in regional climate elements are similar, no matter what the mechanisms of the global warming are. However there is an argument against the hypothesis (besides that real warming-up patterns may differ from the equilibrium ones, which exhibits similarity).

Table II. Linear trends of three hemispherical temperature characteristics for periods, defined by Jones and Kelly (1983)

10^{-2} Kyr^{-1}	$\frac{dT_C}{dt}$	$\frac{d\langle T\rangle}{dt}$	$\frac{d(\Delta T)}{dt}$
1917-39	2.3 ± 0.4	1.9 ± 0.3	0.6 ± 0.3
1940-64	-0.6 ± 0.4	-0.9 ± 0.3	-0.6 ± 0.3
1965-80	1.6 ± 0.8	1.6 ± 0.6	-0.1 ± 0.8

This doubt is based on the study of Jones and Kelly (1983) presenting different local linear trends in two periods, monotonously warming according to continental (T_C) mean temperatures. However, as it is demonstrated in Table II., the two warming-up periods 1917-39 and 1965-80, defined by Jones and Kelly exhibit a marked difference in the continent-ocean temperature contrast. So they can not be considered as completely similar ones. Therefore the difference in their regional temperature trends is not a solid evidence against the basic hypothesis.

REFERENCES

Folland,C.K., Parker,D.E., Kates,F.E., 1984: Worldwide marine temperature fluctuations 1856-1981. Nature 310, 670-673

Jones,P.D., Kelly,P.M., 1983: The spatial and temporal characteristics of Northern Hemisphere surface air temperature variations J. Climatology 3, 243-252

Jones,P.D., Raper,S.C.B., Bradley,R.S., Diaz,H.F., Kelly P.M., Wigley,T.M.L., 1986: Northern Hemisphere Surface Air Temperature Variations 1851-1984. J. Climate Appl. Meteor. 25, 161-179

Lough,J.M., Wigley,T.M.L., Palutikof,J.P., 1983: Climate and climate impact scenarios for Europe in a warmer world. J. Climate Appl. Meteor. 22, 1673-1684

Vinnikov, K.Ya., 1986: Sensitivity of climate. Leningrad, 219 p. (in Russian).

Webb T. III, Wigley T.M.L., 1985: What past climates can indicate about a warmer world ? In:The Potential Climatic Effects of Increasing Carbon Dioxide, Dept. of Energy, DOE/ER-0237, Washington D.C., 237-258

FINGERPRINT DETECTION USING SPATIAL CORRELATION TECHNIQUES

by

P.D. Jones
B.D. Santer[1]
and
T.M.L. Wigley

Climatic Research Unit
University of East Anglia
Norwich NR4 7TJ U.K.

[1]Max-Planck-Institut für Meteorologie
Bundesstrasse 55
2 Hamburg 13
Germany

1. Introduction

In detecting the enhanced greenhouse effect, the key step is to be able to attribute an observed change in climate to this specific cause. Attribution almost certainly requires the identification in the observational record of a multivariate signal characteristic of (and, ideally, unique to) greenhouse-gas-induced climatic change. This type of detection approach has been called the fingerprint method (Madden and Ramanathan, 1980; MacCracken and Moses, 1982). Previous fingerprint studies have been inconclusive either because the signal has been obscured by the noise of regional-scale natural climatic variability (Barnett, 1986; Barnett and Schlesinger, 1987; Santer et al., 1991; Barnett, 1991), or because of uncertainties regarding the level and structure of both the signal and natural variability (Madden and Ramanathan, 1980; Karoly, 1987, 1989; Wigley and Barnett, 1990).

A fingerprint detection variable may be considered to be a vector whose components are either different scalar variables (e.g., temperature, precipitation, etc.) and/or the same variable measured at different points or averaged over different regions (e.g., temperatures at different locations on the Earth's surface or at different levels in the atmosphere) (Wigley and Barnett, 1990). The basic detection strategy is to compare the observed time series of this vector either with an estimate of the equilibrium greenhouse-gas signal (as may be inferred by differencing the results of equilibrium model experiments for $1xCO_2$ and $2xCO_2$), or with a time-evolving signal (inferred from a model simulation with time-dependent greenhouse-gas forcing).

2. Definition of Detection Vector and Data Used

Before performing any analysis, some consideration should be given to the choice of the detection vector. There are at least four criteria which can be used to evaluate the usefulness of a given vector for detection studies.

- First, the individual components should have high signal-to-noise ratios (SNR). Results from climate model experiments indicate that temperature and atmospheric moisture content are relatively good choices in this regard, while precipitation and sea level pressure have low SNR (Barnett and Schlesinger, 1987; Santer et al., 1991).

- Second, the signal vector should not be model-specific. If different models gave different signal vectors, then this would lower confidence in the signal from any one model.

- Third, the signal vector should be easily distinguished from (i.e., near orthogonal to) both the signals due to other forcing factors and the noise of natural internal variability on the 10-100 year time scale relevant to the enhanced greenhouse effect (Barnett and Schlesinger, 1987; Wigley and Barnett, 1990). A poor choice in this regard might be the troposphere/lower stratosphere temperature change contrast, where the expected greenhouse signal is apparently similar to observed natural variability (Liu and Schuurmanns, 1990). The spatial pattern of near-surface temperature changes, which we use here, is also less than ideal, since the feedback processes that influence the spatial character of the enhanced greenhouse signal must also operate with other external forcing mechanisms, such as changes in solar irradiance (Wigley and Jones, 1981).

- Fourth, suitable observational data must exist. Because it is the 10-100 year time scale that is of concern, long data records are needed. Only surface-based data have records exceeding ~40 years in length, and of these, only temperature data satisfy the high signal-to-noise ratio condition.

Part of the reason why detection is so difficult is because the intersection of the sets satisfying the above criteria is effectively empty. The observed data we use satisfy these criteria better than any other data set, viz. 5° x 5° gridded monthly-mean, land-based surface air temperatures and sea surface temperatures (SST). These data come from the combined land-ocean data set described by Jones et al. (1991). They are expressed as anomalies relative to 1950-79 and span the 90-year interval, 1900-1989. We only use data from 1900 because of the reduced coverage and decreased reliability of earlier data.

The data we use for defining a greenhouse-gas signal come from five separate climate model experiments. The experiments differ in terms of the models employed and the forcings applied. All experiments use an atmospheric general circulation model (AGCM) coupled to an ocean model. The ocean models vary from simple mixed-layer models in which horizontal heat transport is prescribed (viz., OSU [Schlesinger and Zhao, 1989], Geophysical Fluid Dynamics Laboratory [GFDL; Wetherald and Manabe, 1990] and United Kingdom Meteorological Office [UKMO; Wilson and Mitchell, 1987] to a model in which heat transport below the mixed-layer is parameterized as a diffusive process (Goddard Institute for Space Studies [GISS; Hansen et al., 1988]) and an ocean model intermediate in complexity between the mixed-layer models and fully-dynamic ocean general circulation models (Lawrence Livermore National Laboratory [LLNL; Gates and Potter, personal communication], which uses a two-layer ocean [Pollard, 1982]).

Results are taken from experiments with two different types of forcing: step-function and time-dependent. In the former (OSU, GFDL,

UKMO and LLNL), the steady-state response to a step function doubling of CO_2 is considered. In the latter (the GISS Transient A experiment; see Hansen et al., 1988), greenhouse forcing commences in 1958 and increases with time in a manner similar to the IPCC "Business as Usual" scenario (Houghton et al., 1990). We used control run data and information from years 81-90 of the GISS perturbed run.

3. Spatial Correlation Methods

We consider two observed data sets D1 and D2 (D = Data), and two model data sets M1 and M2 (M = Model). Each are two-dimensional (x,t) arrays where x and t are independent discrete variables running over space (grid-points) and time (years), with x = 1, ...,m and t = 1,...,n respectively. Here, D1 and D2 represent near-surface temperature averaged over two different periods of time. M1 and M2 are corresponding temperatures simulated in a control run and a greenhouse-gas experiment, respectively. If M1 and M2 are derived from equilibrium experiments, we are concerned with a signal pattern which has no time dependence viz. the equilibrium surface temperature change between a $2xCO_2$ experiment and a $1xCO_2$ control run

$$M(x) = M2(x) - M1(x) \tag{1}$$

Our detection strategy, which follows that used by Barnett (1986) and Barnett and Schlesinger (1987), is to compare the signal pattern, M(x), with a series of time-evolving observed patterns of change

$$D(x,t) = D2(x,t) - D1(x) \tag{2}$$

where the index t runs over the total number of observed mean change fields. Here, D1(x) is the time-average over some fixed reference period, defined as

$$D1(x) = [\sum_{u=t_0}^{t_0+p-1} D(x,u)]/p \tag{3}$$

where $p<n$ and t_0 is the initial year of the reference period. Similarly, D2(x,t) is defined by

$$D2(x,t) = [\sum_{u=t}^{t+p-1} D(x,u)]/p \tag{4}$$

where $t \geq t_0+p$. Here we take p=10. The use of decadal averages and overlapping observed decades yields a smoother estimate of the observed changes. In later examples we assume that $t_0=0$, corresponding to the year 1900, and therefore compare a single model-derived signal pattern with 71 observed mean fields of change relative to 1900-09 for the overlapping decades from 1910-19 to 1980-89.

If the enhanced greenhouse effect is actually becoming stronger in the observed data, and if the model-based estimate of the signal is reasonable, M(x) and D(x,t) should become increasingly similar - i.e., there should be a positive trend in the statistic used to measure similarity. Detection then becomes a problem of identifying a significant trend in the similarity statistic. The next choice, therefore, is to select an appropriate indicator of pattern similarity. In the past, as we will show below, some inappropriate choices have been made.

As a similarity indicator, Barnett and Schlesinger (1987) use a statistic, C(t), which involves the uncentered cross-moment between the M(x) and D(x,t) fields

$$C(t) = [\sum_{x=1}^{m} D(x,t)\ M(x)]/[\sum_{x=1}^{m} (M(x))^2] \qquad (5)$$

(Note that in their original usage Barnett and Schlesinger (1987) used a fixed single year to define the reference period rather than a decadal-average as used here). More conventionally, pattern similarities are quantified using a centered statistic such as a correlation coefficient (i.e., where D(x,t) and M(x) are centered on their respective spatial means). This is the approach which we adopt here. In its general form, the centered detection statistic, R(t), is defined as the pattern correlation between the observed and simulated fields.

$$R(t) = [\sum_{x=1}^{m} (D(x,t) - \bar{D}(t))\ (M(x) - \bar{M})]/[m\ S_D(t)\ S_M] \qquad (6)$$

where S_D^2 and S_M^2 are the spatial variances defined as

$$S_D(t)^2 = [\sum_{x=1}^{m} (D(x,t) - \bar{D}(t))^2]/m$$

(S_M^2 similarly) and $\bar{D}$ and $\bar{M}$ are spatial means defined as

$$\bar{D}(t) = [\sum_{x=1}^{m} D(x,t)]/m$$

($\bar{M}$ similarly). It turns out that C(t) is <u>not</u> an appropriate statistic. To see this, we need to determine the relationship between C(t) and R(t).

If we define

$$Z^2 = [\sum_{x=1}^{m} (M(x))^2]/m = S_M^2 + \bar{M}^2 \qquad (7)$$

then C(t) can be written as

$$C(t) = \{[\sum_{x=1}^{m} (D(x,t) - \bar{D}(t))(M(x) - \bar{M})] + m\ \bar{D}(t)\ \bar{M}\}/mZ^2$$

$$= [S_D(t)\ S_M\ R(t) + \bar{D}(t)\ \bar{M}]/Z^2$$

Hence

$$C(t) = a\ R(t) + b\ \bar{D}(t) \qquad (8)$$

where

$$a = [S_D(t)S_M]/Z^2 \qquad (9)$$

$$b = \bar{M}/Z^2 \qquad (10)$$

Equ. (8) shows that C(t) has an R(t) component and a component determined by the spatial-mean of the D field. If, therefore, D(t) has a trend, this will produce a trend in C(t) independent of the behaviour of R(t). It may be, therefore, that C(t) tells us little more than $\bar{D}(t)$, in which case C(t) would be of little value as a fingerprint test statistic. To determine how much the "message" in C(t) is affected by the behaviour of the spatial mean, we need to estimate the coefficients "a" and "b" in equ. (8). As a specific example, we use model data from the LLNL model and observed data from the land-ocean grid-point temperature data set. In this example we use pointwise normalized versions of D(x,t) and M(x). Normalization is achieved for both observed and modelled data using the standard deviations of the observed data. The normalization procedure is not important to the argument and qualitatively identical results would have been achieved without normalization.

The specific numerical values we require are $S_D(t)$, S_M and $\bar{M}$ (the latter two of which determine Z). For the LLNL model, M varies between 32.0 and 33.3 (mean value, 32.7) and S_M varies between 14.1 and 14.5 (mean, 14.3). Hence $Z^2 \simeq 1274$. For the observed data, $S_D(t)$ varies with time from about 1.9 initially (for the difference field 1910-19 minus 1900-09) to about 3.3 (for 1980-89 minus 1900-09). For the 71-year record of $S_D(t)$ values, S_D is near 2 for the first 30 years, near 2.5 for the next 25 years and rises steadily over the next 16 years to 3.3. The mean value is about 2.5. Inserting $Z^2 = 1274$, $S_M = 14.3$, $\bar{M} = 32.7$ and $S_D = 2.5$ into equs. (8), (9) and (10) gives

$$C(t) = 0.028\ R(t) + 0.026\ \bar{D}(t) \qquad (11)$$

Since a and b are similar, the behaviour of C(t) depends of which of R(t) and $\bar{D}(t)$ is the more variable. $\bar{D}(t)$ rises rapidly from near zero to around 2.7 over the first 25-30 years, oscillates about this value for the next 30-35 years, and then rises rapidly to 4.2 for 1980-89 minus 1900-09 - see Figure 1. R(t), however, can only vary between -1 and 1. The net effect is that the C(t) time series looks very much like the $\bar{D}(t)$ time series. Furthermore, the $\bar{D}(t)$ time series, even though the data are normalized, looks very much like the global-mean unnormalized temperature time series. In other words, use of C(t) is effectively the same as using a spatial mean.

The reason C(t) (or, probably, any uncentered statistic) doesn't succeed as a multivariate test statistic is because it does not separate the effect of the changing mean (which, for $\bar{D}(t)$, is a strongly time-varying quantity) from the true pattern similarity.

We now calculate R(t) using the observed data and the five sets of model data discussed earlier. Pattern correlations were calculated using the grid-point data (after excluding grid-points with missing observed data and the corresponding model grid-points), defined by equations (1) to (4). The results for the five models are shown in

Figure 2. In none of the five cases is there any increase of R(t) with time. In other words, there appears to be no increasing expression of any of the five model signals of the enhanced greenhouse effect in the observational record.

4. Conclusions

In this paper we have introduced a spatial pattern correlation statistic, R(t), in order to attempt to detect the GCM-based climate change signal due to increasing greenhouse gas concentrations in the observed data. We show how the statistic is related to the C(t) statistic defined earlier by Barnett and Schlesinger (1987). Uncentered statistics like C(t) are inappropriate for detection because they do not separate the effect of the changing mean from the true pattern similarity. C(t) is, in effect, equivalent to the spatial-mean time series (i.e., global-mean temperature for the data considered here).

The detection exercise using R(t) to measure the pattern similarity between the five model signals and the observed temperatures during the present century failed. There was no evidence of an increasing expression of the signal in the observed data. Elsewhere, Santer et al. (1992) consider methods to optimize detection by concentrating on regions of high signal-to-noise ratios. This is achieved by comparing fields of pointwise normalized model and observed changes as opposed to the raw changes. A number of statistical problems are highlighted by Santer et al. (1992) that seem to limit the value of normalization.

Of the set of possible spatial similarity fingerprint detection methods, it is the most obvious and basic of these (pattern correlations using the raw data) that appears to be the most useful. Alternatives, using non-centered moments and/or normalizing to highlight certain (high SNR) regions, have drawbacks that make them of lesser use. Results obtained here with the direct method (i.e., using R(t)) are negative indicating that the spatial details of the greenhouse signal are still within the noise, or that the signal itself is poorly defined.

5. Acknowledgements

We gratefully acknowledge the assistance of W.L. Gates, M.E. Schlesinger, J.F.B. Mitchell, J.E. Hansen and S. Manabe in providing model results. The work was supported by the U.S. Department of Energy, Atmospheric and Climate Research Division, and by the U.K. Department of the Environment.

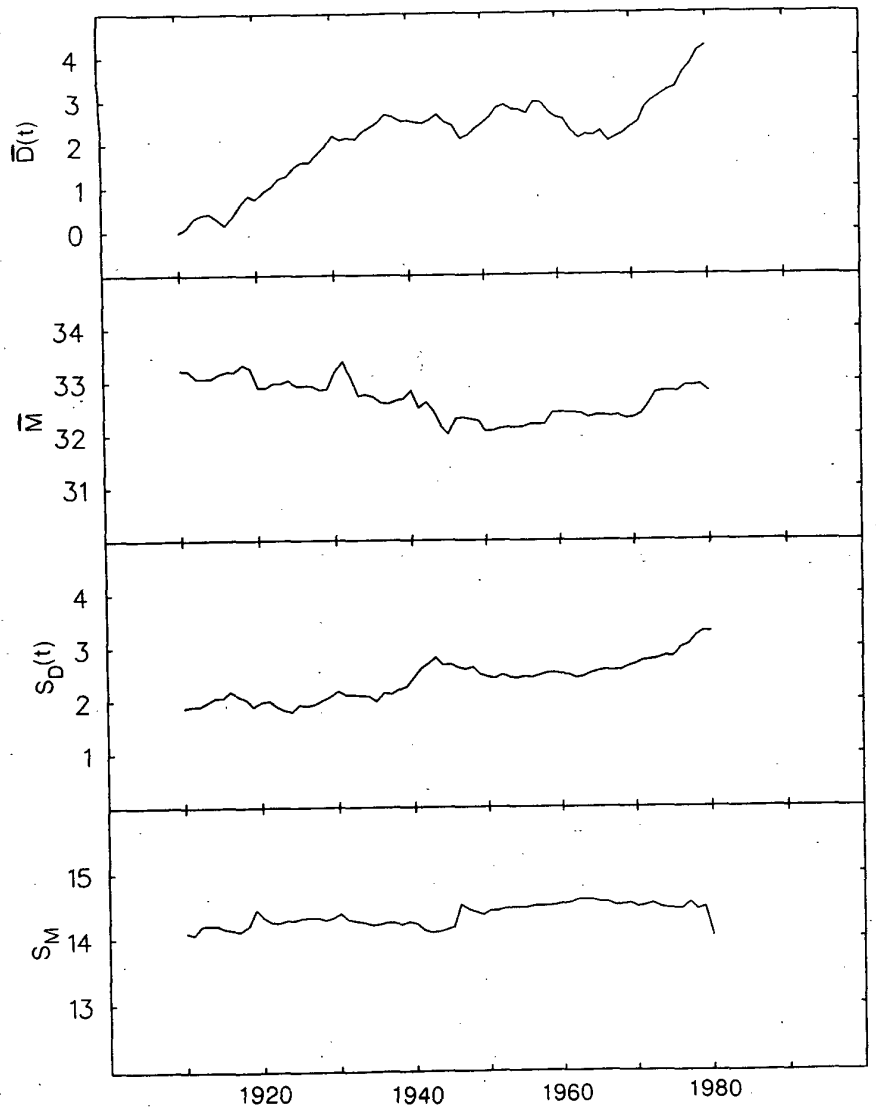


Figure 1: Observed and model-based spatial means and standard deviations of pointwise normalized temperature changes. Changes in time of M and S_M are due to slight changes in data coverage only. The data used is determined by coverage during the initial decade, 1900-09. Slight reductions in observed coverage (and hence in the model data used) occur over 1910 to 1989. Changes in D and $S_D(t)$ represent decadal-scale climate changes. Note the similarity of D to the global-mean temperature time series.

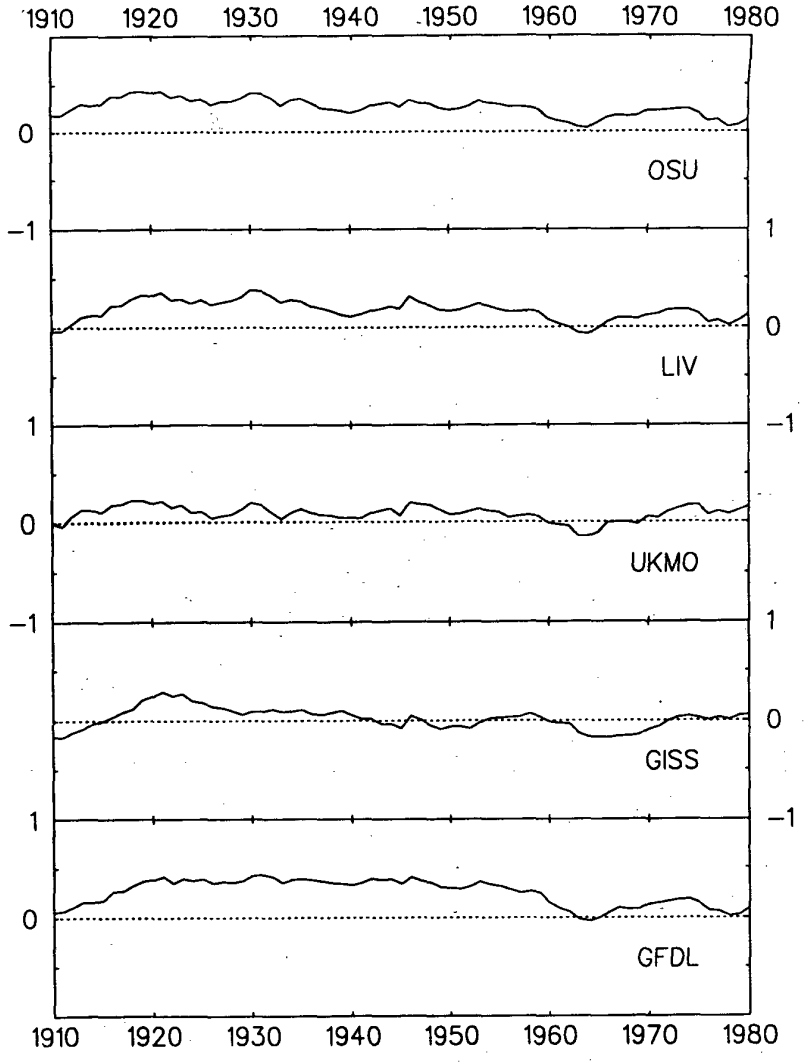


Figure 2: Model versus observed pattern correlations. Pattern correlations are calculated from grid point data using 71 overlapping decades (from 1910-19 to 1980-89) relative to a fixed reference decade (1900-09). Results are plotted on the first year of each decade.

References

Barnett, T.P., 1986: Detection of changes in the global troposphere field induced by greenhouse gases. J. Geophys. Res. 91, 6659-6667.

Barnett, T.P., and Schlesinger, M.E., 1987: Detecting changes in global climate induced by greenhouse gases. J. Geophys. Res. 92, 14772-14780.

Barnett, T.P., 1991: An attempt to detect the greenhouse-gas signal in a transient GCM simulation. In (M.E. Schlesinger, Ed.) Greenhouse-gas-induced Climatic Change: A critical appraisal of simulations and observations, Elsevier, 559-568.

Hansen, J., Fung, I., Lacis, A., Rind, D., Lebedeff, S., Ruedy, R. and Russell, G., 1988: Global climate changes as forecast by Goddard Institute for Space Studies three-dimensional model. J. Geophys. Res. 93, 9341-9364.

Houghton, J.T., Jenkins, G.J. and Ephraums, J.J. (Eds.), 1990: Climate Change: The IPCC Scientific Assessment. Cambridge University Press, 365pp.

Jones, P.D., Wigley, T.M.L. and Farmer, G., 1991: Marine and land temperature data sets: A comparison and a look at recent trends. In (M.E. Schlesinger Ed.) Greenhouse-gas-induced Climatic Change: A critical appraisal of simulations and observations, Elsevier, 153-172.

Karoly, D.J., 1987: Southern Hemisphere temperature trends: A possible greenhouse gas effect? Geophys. Res. Lett. 14, 1139-1141.

Karoly, D.J., 1989: Northern Hemisphere temperature trends: A possible greenhouse gas effect? Geophys. Res. Lett. 16, 465-468.

Liu, Q. and Schuurmanns, C.J.E., 1990: The correlation of tropospheric and stratospheric temperatures and its effects on the detection of climate changes. Geophys. Res. Lett. 17, 1085-1088.

MacCracken, M.C. and Moses, H., 1982: The first detection of carbon dioxide effects: Workshop Summary, 8-10 June 1981, Harpers Ferry, West Virginia. Bull. Am. Met. Soc. 63, 1164-1178.

Madden, R.A. and Ramanathan, V., 1980: Detecting climate change due to increasing carbon dioxide. Science 209, 763-768.

Pollard, D., 1982: The performance of an upper-ocean model coupled to an atmospheric GCM: Preliminary results. Climatic Research Institute Report 31, Corvallis, Oregon.

Santer, B.D., Wigley, T.M.L., Jones, P.D. and Schlesinger, M.E., 1991: Multivariate methods for the detection of greenhouse-gas-induced climate change. In (M.E. Schlesinger Ed.) Greenhouse-gas-induced climate change: A critical appraisal of simulations and observations. Elsevier, 511-536.

Santer, B.D., Wigley, T.M.L. and Jones, P.D., 1992: Correlation methods in fingerprint detection studies. Climate Dynamics (submitted).

Schlesinger, M.E. and Zhao, Z.-C., 1989: Seasonal climate changes induced by doubled CO_2 as simulated by OSU atmospheric GCM mixed-layer ocean model. J. Climate 2, 459-495.

Wetherald, R.T. and Manabe, S., 1990: Cited by Cubasch, U. and Cess, R.D., In J.T. Houghton, G.J. Jenkins and J.J. Ephraums Eds.) Climate Change: The IPCC Scientific Assessment, Cambridge University Press, p81.

Wigley, T.M.L. and Jones, P.D., 1981: Detecting CO_2-induced climate change. Nature 292, 205-208.

Wigley, T.M.L. and Barnett, T.P., 1990: Detection of the Greenhouse Effect in the Observations. In (J.T. Houghton, G.J. Jenkins and J.J. Ephraums Eds.) Climate Change: The IPCC Scientific Assessment, Cambridge University Press, 239-255.

Wilson, C.A. and Mitchell, J.F.B., 1987: A doubled CO_2 Climate Sensitivity experiment with a GCM including a simple ocean. J. Geophys. Res., 92, 13315-13343.

Greenhouse Climate Change Fingerprint Analysis Using Radiosonde Data

David J. Karoly

Centre for Dynamical Meteorology, Monash University
Clayton VIC 3168, Australia

1. Introduction

Over the last hundred years, there has been a significant global-mean surface warming of about 0.3–0.6°C (Folland et al., 1990), which is generally consistent with the enhanced greenhouse effect. It is possible that this recent observed warming could have been due to processes other than the enhanced greenhouse effect, including natural climate variability. It is very difficult to demonstrate a causal relationship between observed climate change and the enhanced greenhouse effect using a single variable, such as global-mean surface temperature. As stated by Wigley and Barnett (1990), " To claim detection in a useful and practical way, we must not only identify a climatic change, but we must **attribute** at least part of such a change to the enhanced greenhouse effect. ... Detection requires that the observed changes in climate are in accord with detailed model predictions of the enhanced greenhouse effect, demonstrating that we understand the cause or causes of the changes.". The method which has been proposed to address the attribution problem is the fingerprint method (Madden and Ramanathan, 1980; Wigley and Barnett, 1990); identification of a multivariate signal or 'fingerprint' unique to the enhanced greenhouse effect and then determination of a significant change in this signal in observational data.

In this study, an example of fingerprint analysis of greenhouse climate change is presented. The zonal mean atmospheric temperature change between equilibrium climate model simulations with control and double CO_2 concentrations is used as the greenhouse fingerprint. This fngerprint is compared with the observed atmospheric temperature changes over the period 1963–88 from the radiosonde-based analyses of Oort (1983, updated).

It is important that the greenhouse fingerprint should have a structure unique to the enhanced greenhouse effect to allow unequivocal attribution. However, the patterns of low frequency atmospheric temperature change associated with natural climate variability or with other forcing processes are not well known. Some processes other than the enhanced greenhouse effect which might produce this pattern are discussed in Section 3.

There may be problems with using the results of equilibrium simulations from climate models to define the greenhouse fingerprint as the climate system is not in equilibrium with the increasing concentrations of greenhouse gases in the atmosphere. There have been very few time-dependent greenhouse climate model simulations with slowly increasing concentrations of greenhouse gases (Stouffer et al., 1989, Washington and Meehl, 1989). The inclusion of ocean dynamics in these model simulations leads to slower greenhouse signal definition and greater variability within and between the model simulations. Some results for the zonal mean atmospheric temperature changes on decadal timescales in time-dependent greenhouse climate model simulations are presented in Section 4. These are compared with the equilibrium greenhouse fingerprint and provide some measure of climate noise due to natural variability. They also allow some comparison of the greenhouse fingerprint with changes due to natural variations.

2. A Fingerprint Example

The fingerprint used in this example is the zonal mean atmospheric temperature change as a function of height and latitude between double CO_2 and control equilibrium climate model simulations. This signal is the most consistent pattern of greenhouse climate change for different models (Schlesinger and Mitchell, 1987). Here, the results from 3 different climate models are used. These models are updated versions of those used in the intercomparison of equilibrium double CO_2 climate model simulations by Schlesinger and Mitchell (1987). Each of these climate models involves an atmospheric general circulation model coupled to a simple mixed layer ocean model. They have been run at the Geophysical Fluid Dynamics Laboratory (Wetherald and Manabe, 1988; referred to as the GFDL model), the National Center for Atmospheric Research (Meehl and Washington, 1990; referred to as the NCAR model) and at the UK Meteorological Office (Wilson and Mitchell, 1987; referred to as the UKMO model). The atmospheric models have rather different physical parameterisation schemes and computational formulation but similar resolution, with 9 (GFDL and NCAR) or 11 (UKMO) vertical levels and relatively low horizontal resolution (equivalent to about 5° of latitude).

Each of the climate models was run to equilibrium with control CO_2 concentrations and with double CO_2 concentrations. The difference between the time mean (10-year for GFDL, 5-year for NCAR and 15-year for UKMO) zonal mean temperature from these equilibrium simulations defines the greenhouse fingerprint for this example. The annual mean difference is shown in Fig. 1 for the GFDL model, with very similar patterns for the other models. The pattern correlations between these model temperature differences are all greater than 0.95, confirming the great similarity between the results from the different models. Seasonal mean differences for June-August (JJA) and December-February (DJF) have been considered also but are not shown as they are similar to those in Fig. 1, apart from enhanced warming at low levels near the winter pole (Schlesinger and Mitchell, 1987).

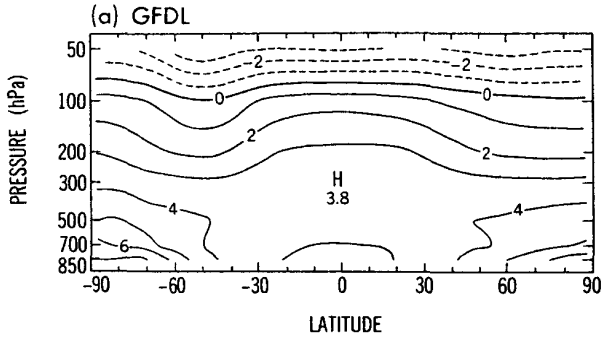


Figure 1. Zonal mean annual mean atmospheric temperature differences between equilibrium double CO_2 and control climate model simulations with the GFDL model. The contour interval is 1° C and negative contours are dashed.

There is a relatively short period of observations of atmospheric temperature above the surface as reliable radiosonde data with reasonable global coverage are available from the late 1950s only. In addition, these data have usually been used for weather analysis and forecasting. Changes in analysis methods and inclusion of different sources of temperature observations, including satellite-derived temperature soundings, have often led to a conventional analysis record which is inappropriate for investigation of climate change. Recently, Oort has updated his global analyses of atmospheric temperature prepared from radiosonde data alone from the original period, May 1963 to April 1973 (Oort, 1983) to include an additional 15 years up to April 1988. These global analyses have been prepared using the same analysis technique for the whole period. Problems for detecting climate change with these analyses may arise through changes in the spatial coverage of the radiosonde network, changes in radiosonde instrument type and less than adequate coverage over ocean areas, particularly in the Southern Hemisphere. It is beyond the scope of this paper to describe the potential impact of these problems on detecting climate change but it is important to recognise these possible sources of error. Details of the numbers of stations available for these analyses and some analysis of upper air temperature trends is presented by Oort and Liu (1992).

These analyses are used to define the evolving thermal structure of the atmosphere over the period 1963 to 1988. The zonal mean temperature is used at seven pressure levels; 850, 700, 500, 300, 200, 100 and 50 hPa (approximate altitudes of 1, 3, 6, 11, 14, 19 and 24 km), and at 5° latitude intervals from 87.5°N to 87.5°S to be compatible with the model resolution. Thus a total of 252 points are used to define the spatial patterns. The difference between the 10-year average annual means for the last 10 years (1978-88) and first 10 years (1963-73) of these analyses is the time-mean observed temperature change and is shown in Fig. 2. The annual mean is defined to be the 12 month average from May to April, rather than the more conventional January to December, because May 1963 is the first month of global analyses available. Stippling in Fig. 2 shows differences which are locally significant at the 5% level using a t-statistic test. Note the general warming in the lower troposphere and cooling in the lower stratosphere over this period. The seasonal mean temperature differences are quite similar to the annual mean pattern.

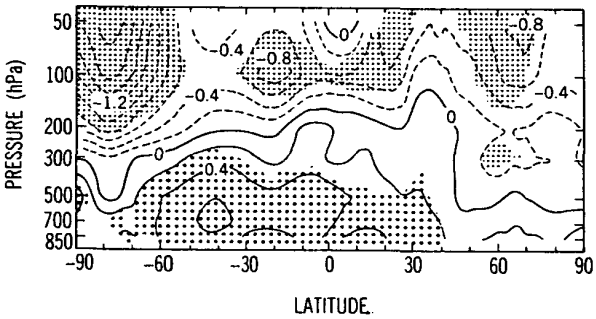


Figure 2. Observed zonal mean annual mean temperature difference between 10-year averages for periods 1978-88 minus 1963-73. The contour interval is 0.2° C and negative contours are dashed. Stippling indicates differences locally significant at the 5% level using a t-statistic test.

Multiple permutation field significance tests as described by Livezey and Chen (1983) were performed to determine if the difference fields are significantly different from zero. These showed that the annual and JJA difference fields are significant at more than the 2% level while the DJF difference is significant at the 4% level, indicating that there has been significant temperature change over this period.

Table 1: *Pattern correlations between observed decadal zonal mean temperature change and the three greenhouse climate model fingerprints. The significance levels of the pattern correlations obtained by multiple permutation tests are indicated by superscripts.*

Averaging period	GFDL	NCAR	UKMO
Annual	$0.65^{10\%}$	$0.74^{5\%}$	$0.72^{5\%}$
JJA	0.51	$0.63^{10\%}$	$0.63^{10\%}$
DJF	$0.45^{10\%}$	$0.56^{5\%}$	$0.60^{1\%}$

To test the null hypothesis that the observed and modelled fields do not differ, the observed mean temperature changes were compared with the model greenhouse fingerprints. Pattern correlations are shown in Table 1 for the annual, JJA and DJF means. Significance levels were determined by multiple resampling of the observed dataset. The observed temperature changes have significant positive pattern correlations with the greenhouse fingerprints. This indicates that the patterns have features in common, including warming in the lower troposphere and cooling in the lower stratosphere. About half the spatial variance of the observed temperature change can be explained by the greenhouse fingerprint for annual mean data and about a third for the seasonal data. The temperature changes shown in Figs. 1 and 2 appear to be less similar than indicated by the pattern correlations in Table 1, because the pattern correlation is calculated after removing the pattern mean value, which is different between the patterns. The largest difference between the observed and modelled patterns is that the level at which warming changes to cooling in the observations is much lower than in the model simulations. Other details such as the enhanced warming at high latitudes and in the tropical upper troposphere in the model results are not apparent in the observations.

3. Uniqueness of the Greenhouse Fingerprint

One of the assumptions required for greenhouse fingerprint analysis is that the greenhouse fingerprint must be unique to the enhanced greenhouse effect and not due to any other forcing mechanism. Other possible causes of the observed tropospheric warming and stratospheric cooling at all latitudes may include changes in the frequency or amplitude of El Nino-Southern Oscillation events, general increases in sea surface temperatures or decreases in stratospheric ozone concentrations. We have considered the possible role of these mechanisms in this Section. The possible role of natural climate variability in contributing to the observed changes is discussed in the next Section.

El Nino-Southern Oscillation (ENSO) events are irregular warmings of the eastern equatorial Pacific Ocean at intervals of about 3 to 6 years, which are also associated with warming of the tropical troposphere (Pan and Oort, 1983). It has been suggested that the warm episodes in 1982-83 and 1986-87 without an intervening cool episode may have contributed to the observed atmospheric warming during the 1980s. There are two reasons why ENSO events are not a significant contribution to the increasing greenhouse fingerprint in the observations. First, there are only very small differences in ENSO indices, such as east equatorial Pacific (EEP) sea surface temperature (SST), on time scales longer than 10 years, with most of the variability being on time scales of 2 to 6 years. The difference of EEP SST (averaged over the region 10°N-10°S, 180°-90°W) between the two 10 year periods, 1978-88 minus 1963-73, is about 0.2°C using updated SST analyses from Pan and Oort (1983). In addition, the pattern of atmospheric temperature anomalies associated with ENSO events is different from the observed decadal temperature changes (Pan and Oort, 1983). The largest warming associated with ENSO events is in the tropical troposphere, with much smaller changes in middle and high latitudes in the troposphere, and opposite sign changes in the stratosphere in the Northern and Southern Hemispheres. This temperature change pattern associated with ENSO events has also been found in climate model simulations and in ECMWF analyses by Meehl and Albrecht (1991).

Another possible forcing which may have contributed to the observed atmospheric temperature changes is the observed large-scale increase in SST, particularly in the Southern Hemisphere, over the data period (Folland et al., 1990). The observed variations of SST may be associated with long-term changes in deep ocean circulation and may be unrelated to the enhanced greenhouse effect. It is likely that there would be a warming of the troposphere associated with these large-scale increases of SST alone. Some studies of the response of climate models to uniform global increases of SST were carried out as part of an international model intercomparison (Cess et al., 1989). The pattern of tropospheric warming in these studies due to increased SST is remarkably similar to that for the greenhouse fingerprint but there is little stratospheric cooling. This is likely to be due to the absence of an increase in the greenhouse gas concentrations (CO_2 or water vapour) in the stratosphere in the SST increase experiments whereas radiative effects from increased CO_2 in the stratosphere contribute to the cooling there in the double CO_2 model experiments. This suggests that SST increases alone

cannot explain the observed temperature changes.

Reductions in stratospheric ozone concentrations have been observed over the last two decades (WMO, 1989). These are believed to be associated with increases in chlorine concentrations. The thermal fingerprint from climate models due to reductions in stratospheric ozone concentrations shows a pronounced cooling in the stratosphere and weak warming or cooling in the troposphere, depending on the distribution of ozone change used in the model (Kiehl and Boville, 1988). Although it has not been possible to calculate pattern correlations between the ozone reduction and greenhouse fingerprints, the weak tropospheric response to ozone reductions leads to marked differences between the temperature change patterns.

It appears that the greenhouse thermal fingerprint from the model simulations and the observed temperature changes are different from the thermal fingerprints for either ENSO events, SST increases or ozone decreases, individually. A combination of mechanisms, such as increased SST leading to tropospheric warming and decreased ozone leading to stratospheric cooling could have produced a pattern of temperature change similar to that observed.

4. Time-dependent Greenhouse Simulations

It is not possible to estimate the magnitude or pattern of temperature change on decadal timescales due to natural, internal variations of the climate system using observed upper air data because of the short period for which this data is available. Extended simulations with coupled ocean-atmosphere climate models provide a means for estimating natural climate variability. In addition, time-dependent climate simulations using coupled ocean-atmosphere models with slowly increasing concentrations of greenhouse gases may provide a better greenhouse fingerprint than equilibrium model simulations. Such simulations also allow the testing of the fingerprint analysis method on detection of greenhouse climate change in the time-dependent model simulation, in a similar way to the observed climate.

Some preliminary analysis of two such time-dependent coupled ocean-atmosphere climate model simulations is presented here. The models are essentially the atmospheric component of the equilibrium climate models for GFDL and NCAR described above, combined with global ocean circulation models (Stouffer et al., 1989, Washington and Meehl, 1989). The ocean models use a similar horizontal resolution to the atmospheric models but have a fairly coarse vertical resolution (only 4 layers for the NCAR model, 12 for the GFDL model). As with most coupled models, there is a drift of the NCAR model climate in its control simulation. This is avoided in the GFDL model through application of correction of the surface fluxes between the atmosphere and the ocean models, which eliminates the climate drift in the GFDL control simulation. For the greenhouse simulations, the model atmospheric concentrations of CO_2 are increased by 1% per year (compound increase in the GFDL model, simple increase in the NCAR model). This would mean a doubling of CO_2 concentrations after about 70 years in the GFDL model and 100 years in the NCAR model. To compare the time-dependent simulations with the equilibrium greenhouse fingerprint and the observed temperature changes, decadal mean zonal mean temperature fields from transient greenhouse and control simulations with the GFDL model (0 to 100 years) and NCAR model (26 to 70 years) were available.

First, the zonal mean temperature differences between the transient greenhouse and control simulations were compared with the equilibrium greenhouse fingerprint. The greenhouse fingerprint is apparent in the zonal mean temperature changes from the earliest period of the transient simulations, as shown in the time series of pattern correlations in Fig. 3. The temperature differences of the transient simulations from the control runs have pattern correlations with the equilibrium fingerprint greater than 0.8 for the NCAR model and greater than 0.9 for the GFDL model after about 30 years of the transient simulations.

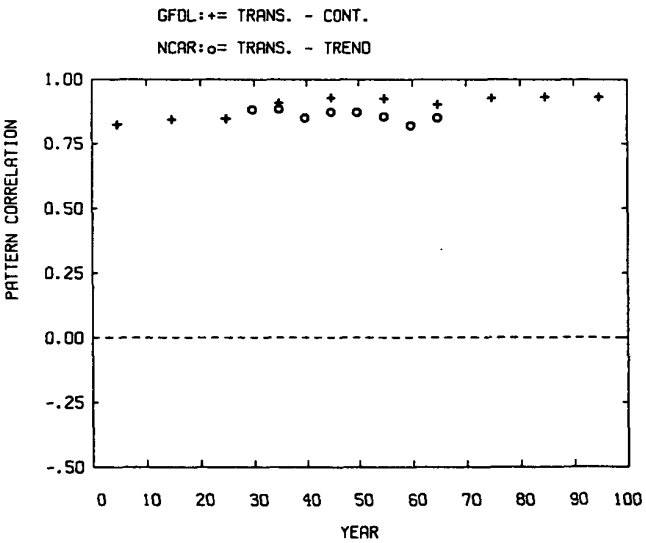


Figure 3. Time series of pattern correlations between the decadal mean temperature changes of the transient greenhouse simulations from the control runs and the equilibrium double CO_2 temperature changes for the GFDL (+) and NCAR (o) models.

To investigate the possible similarities between the greenhouse fingerprint and the patterns of temperature change due to natural variability, the temperature changes between successive ten year periods in the control simulations have been compared with the equilibrium greenhouse fingerprint. In addition, to simulate the detection of the evolving greenhouse signal in the transient simulations, temperature changes between successive ten year periods in the transient simulations have been compared with the equilibrium greenhouse fingerprint. Pattern correlations from these two comparisons are shown in Fig. 4.

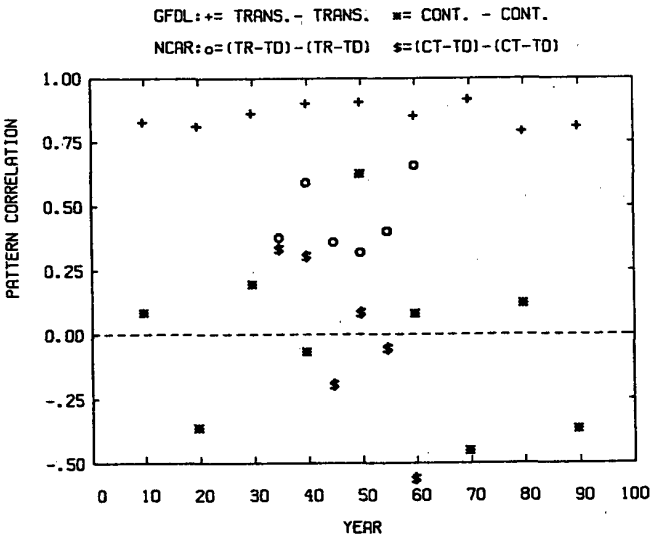


Figure 4. Time series of pattern correlations of the temperature changes between successive decades in the transient greenhouse and control simulations with the equilibrium double CO_2 temperature changes for the GFDL and NCAR models.
*GFDL transient +, control * ;*
NCAR transient o, control $.

The decadal changes in the GFDL transient simulation are very similar to the equilibrium greenhouse fingerprint, with pattern correlations greater than 0.8 in general. The patterns of decadal temperature change in the GFDL control simulation have much smaller pattern correlations than for the transient greenhouse simulation, indicating that natural variability in the GFDL model has quite distinct patterns of temperature change compared with the greenhouse fingerprint. However, for the NCAR transient greenhouse simulation, the decadal changes have smaller pattern correlations, about 0.5, with the equilibrium fingerprint. For the NCAR control simulation, the pattern correlations are usually smaller than for the transient simulation but, in several cases, are similar to those in the transient run. Hence, for the NCAR model, the greenhouse signal is less distinct from natural variability. Similar conclusions are reached when the transient and control model simulations are compared with the observed temperature change pattern.

5. Conclusions

A strategy for the detection of the enhanced greenhouse effect in observational data from the recent international review of climate change (Wigley and Barnett, 1990) has been applied. This study used the zonal mean temperature difference between double CO_2 and control equilibrium climate model simulations from three different models as the greenhouse fingerprints and compared these fingerprints with observed atmospheric temperature variations from radiosonde-station data over the last three decades. There was shown to be a significant increase in the greenhouse signal in the observational data over the period of analysis, 1963 to 1988. About half the spatial variance of the pattern of observed temperature change is common to the greenhouse fingerprints.

However, a number of cautions have to be placed on these results. Upper air data are available for a very short period, probably too short to resolve a real greenhouse climate change and definitely too short to quantify natural variability on decadal time scales. There are important differences between the greenhouse signal shown in Fig. 1 and the observed pattern of temperature change in Fig. 2. These may indicate either that different forcing mechanisms are responsible for the two patterns, that natural variability is confusing the observed pattern or that the models are not simulating the greenhouse signal correctly. It is possible that the multivariate greenhouse signal used here is not unique to the greenhouse effect and may be due to other forcing mechanisms, including decreases in stratospheric ozone concentrations, increases in sea surface temperatures or natural climate variability. The atmospheric thermal fingerprints associated with ENSO events, with uniform global increases of SST and with decreases of stratospheric ozone concentrations have been considered separately. These are distinct from the greenhouse fingerprint and, individually, have small contribution to the observed pattern of temperature change.

Preliminary analysis of two time-dependent coupled ocean-atmosphere climate models with slowly increasing CO_2 concentrations indicates that the equilibrium greenhouse fingerprint is very similar to the patterns of zonal mean atmospheric temperature change from these models and is generally distinct from patterns of temperature change associated with natural climate variability.

Acknowledgements

This study is based on a methodology presented

in the Intergovernmental Panel on Climate Change Working Group 1 Report Section 8 *Detection of the Greenhouse Effect in Observations.* The assistance of Bram Oort in providing access to his observational dataset and Ron Stouffer, Jerry Meehl and John Mitchell in providing their model results is gratefully acknowledged. Jenny Cohen and Dean Collins helped with processing the data. I wish to thank Tom Wigley for making available an unpublished manuscript prepared for the IPCC report, on which the analysis in this paper is based. A longer version of parts of this paper has been accepted for publication in *Climate Dynamics* (Karoly et al, 1992). This study was supported by grants from the Australian Research Council, the Australian Bilateral Science and Technology Program, and the National Greenhouse Advisory Committee.

References

Barnett, T. P., and M. E. Schlesinger, 1987: Detecting changes in global climate induced by greenhouse gases. *J. Geophys. Res.*, **92**, 14772-14780.

Cess, R. D., et al., 1990: Interpretation of cloud-climate feedback as produced by 14 atmospheric general circulation models. *Science*, **245**, 513-516.

Folland, C. K., T. Karl and K. Ya. Vinnikov, 1990: Observed climate variations and change. In *Climate Change, The IPCC Scientific Assessment*, (Eds. J. T. Houghton, G. J. Jenkins and J. J. Ephraums), CUP, 195-238.

Karoly, D. J., J. A. Cohen, G. A. Meehl, J. F. B. Mitchell, A. H. Oort, R. J. Stouffer and R. T. Wetherald, 1992: An example of fingerprint detection of greenhouse climate change. *Clim. Dyn.* accepted.

Kiehl, J. T., and B. A. Boville, 1988: The radiative-dynamical response of a stratospheric-tropospheric general circulation
model to changes in ozone. *J. Atmos Sci.*, **45**, 1798-1817

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., and V. Ramanathan, 1980: Detecting climate change due to increasing carbon dioxide. *Science*, **209**, 763-768.

Meehl, G. A., and B. A. Albrecht, 1991: Response of a GCM with a hybrid convection scheme to a tropical Pacific sea surface temperature anomaly. *J. Climate*, in press.

Meehl, G. A., and W. M. Washington, 1990: CO_2 climate sensitivity and snow-sea-ice albedo parameterization in an atmospheric GCM coupled to a mixed-layer ocean model. *Climatic Change*, **16**, 283-306.

Oort, A. H., 1983: *Global atmospheric circulation statistics, 1958-1973.* NOAA Prof. Paper 14, U. S. Govt. Printing Office, 108pp.

Oort, A. H., and H. Liu, 1992: Upper air temperature trends over the globe, 1958–1989. Submitted to *J. Climate.*

Pan, Y. -H., and A. H. Oort, 1983: Global climate variations connected with sea surface temperature anomalies in the eastern equatorial Pacific Ocean for the 1958-1973 period. *Mon. Wea. Rev.*, **111**, 1244-1258.

Schlesinger, M. E., and J. F. B. Mitchell, 1987: Climate model simulations of the equilibrium climatic response to increased carbon dioxide. *Rev. Geophys.*, **25**, 760-798.

Stouffer, R. J., S. Manabe and K. Bryan, 1989: Interhemispheric assymmetry in climate response to a gradual increase of CO_2. *Nature*, **342**, 660-662.

Washington, W. M., and G. A. Meehl, 1989: Climate sensitivity due to increased CO_2: experiments with a coupled atmosphere and ocean general circulation model. *Climate Dynamics*, **4**, 1-38.

Wetherald, R. T., and S. Manabe, 1988: Cloud feedback processes in a general circulation model. *J. Atmos. Sci.*, **45**, 1397-1415.

Wigley, T. M. L., and T. P. Barnett, 1990: Detection of the greenhouse effect in the observations. In *Climate Change, The IPCC Scientific Assessment*, (Eds. J. T. Houghton, G. J. Jenkins and J. J. Ephraums), CUP, 239-255.

Wilson, C. A., and J. F. B. Mitchell, 1987: A doubled CO_2 climate sensitivity experiment with a global climate model including a simple ocean model. *J. Geophys. Res.*, **92**, 13315-13343.

World Meteorological Organization, 1989: *Scientific Assessment of Stratospheric Ozone: 1989*, WMO Global Ozone Research and Monitoring Project - Report No. 20, Vol. 1, 486pp.

Orthogonality of Signal and Noise in Time-Dependent Greenhouse Warming Experiments

Benjamin D. Santer, Wolfgang Brüggemann, Ulrich Cubasch, Klaus Hasselmann, Heinke Höck, Ernst Maier-Reimer, and Uwe Mikolajewicz

Max-Planck-Institut für Meteorologie, Bundesstrasse 55, 2 Hamburg 13, FRG

12 May, 1992

Abstract

Results from a control integration and time-dependent greenhouse warming experiments performed with a coupled ocean-atmosphere model are analysed in terms of their signal-to-noise properties. The aim is to illustrate techniques for efficient description of the space-time evolution of signals and noise and to identify potentially useful components of a multivariate greenhouse-gas "fingerprint". The three 100-year experiments analysed here simulate the response of the climate system to a step-function doubling of CO_2 and to the time-dependent greenhouse gas increases specified in Scenarios A ("Business as Usual") and D ("Draconian Measures") of the Intergovernmental Panel on Climate Change (IPCC).

If signal and noise patterns are highly similar, the chances of mistaking natural variability for a greenhouse-gas signal are increased. We use the pattern correlation between the dominant EOFs of the control run and the Scenario A experiment as a measure of the orthogonality of signal and noise patterns. The EOF 1 patterns of signal and noise are least similar for the vertical structure of zonal winds, 2m temperature, precipitable water, and stratospheric/tropospheric temperature contrasts, and are most similar for sea level pressure (SLP). Despite the differences in forcing history, a highly similar EOF 1 surface temperature response pattern is established in all three greenhouse warming experiments. A large part of this similarity is due to a common land-sea contrast component of the signal.

The EOF analysis indicates that 2m temperature, precipitable water, and the vertical distribution of the zonal temperature and zonal winds are characterized by low-dimensional signals embedded in a higher-dimensional noise space. In contrast, cloud cover and precipitation rate have high-dimensional signals and noise. In order to determine the degree to which the signal is contaminated by the natural variability (and/or drift) of the control run, we projected the Scenario A data onto EOFs 1 and 2 of the control. Signal contamination by the EOF 1 and 2 modes of the noise is lowest for 2m temperature, a situation favorable for detection. The signals for precipitable water, SLP, and the vertical structure of zonal temperature and zonal winds are significantly contaminated by the dominant noise modes.

1 Introduction

Recently, considerable attention has been devoted to the problem of detecting the climatic signature of the enhanced greenhouse effect in observed data (Wigley and Barnett, 1990). Detection studies generally involve the comparison of observed changes in a climate variable, such as surface temperature (e.g. Wigley and Jones, 1981; Barnett, 1986; Barnett and Schlesinger, 1987) or troposphere/stratosphere temperature contrast (Karoly, 1987, 1989) with the changes simulated

by a climate model in response to greenhouse gas forcing. Before performing such comparisons, it is important to identify those climate variables and spatial scales which are optimal for detection purposes. By limiting the comparison of model and real-world climate changes to a subset of variables with favorable signal-to-noise properties, the chances of detecting an enhanced greenhouse effect signal in the observed data are improved. The use of a signal comprised of a number of different variables – a so-called multivariate "fingerprint" – also increases the chances of unambiguously attributing a change in climate to the enhanced greenhouse effect (Madden and Ramanathan, 1980; MacCracken and Moses, 1982).

In this study we examine the signal-to-noise properties of a control run and three time-dependent greenhouse warming experiments recently performed with a coupled ocean-atmosphere model (Cubasch et al., 1992). The aim is to illustrate techniques for efficient description of the space-time evolution of signal and noise, and to identify potentially useful components of a greenhouse gas "fingerprint".

2 Orthogonality of Signal and Noise Patterns

Consider two sets of time-evolving spatial patterns, $c(x,t)$ and $s(x,t)$, $x = 1,p$; $t = 1,n$. The indices x and t indicate (respectively) space (model grid-points) and time (years). In the following, we assume that we are dealing with only one variable at one model level, and that $c(x,t)$ and $s(x,t)$ are (respectively) data from the coupled model control integration and the Scenario A integration performed by Cubasch et al. (1992), e.g. for 2m temperature. Here $p = 2048$ (for 2-d fields), $p = 480$ (for zonally-averaged data at 15 atmospheric levels) and $n = 100$. All results are for annually-averaged data. We assume further that $c(x,t)$ and $s(x,t)$ are expressed as anomalies relative to the smoothed initial state of the control integration and that both fields have been area-weighted, as in Cubasch et al. (1992). In our notation we use the superscripts c and s to distinguish between control and signal data.

The EOF representation of $c(x,t)$ and $s(x,t)$ is

$$c(x,t) = \sum_{j=1}^{q} \alpha_j^c(t) e_j^c(x) \tag{1}$$

$$s(x,t) = \sum_{j=1}^{q} \beta_j^s(t) e_j^s(x) \tag{2}$$

$x = 1,\ldots,p\,; \quad t = 1,\ldots,n$

where $\alpha_j^c(t), \beta_j^s(t)$ are the principal component (PC) time series of the control run and the Scenario A experiment and $e_j^c(x), e_j^s(x)$ are the respective control and experiment EOFs. Note that the covariance matrices used for computing the EOFs are rank-deficient, so that the number of EOFs with non-zero eigenvalues is actually only $q = \min(n-1,p) = 99$ (see Preisendorfer, 1988). Each set of EOFs is normalized to form an orthonormal basis, i.e.,

$$\sum_{x=1}^{p} e_j^c(x) e_k^c(x) = \delta_{jk} \tag{3}$$

$j,k = 1,\ldots,q$

In order to measure the similarity of the dominant signal and noise patterns we computed r, the correlation between the EOF 1 patterns of Scenario A and the control run;

$$r = \sum_{x=1}^{p} (e_j^s(x) - \overline{e_j^s})(e_j^c(x) - \overline{e_j^c})/(\sigma_j^s \sigma_j^c) \quad (j = 1) \tag{4}$$

where $\overline{e_j^s}, \overline{e_j^c}$ and σ_j^s, σ_j^c are the spatial means and spatial standard deviations of the Scenario A and control EOF 1 patterns.

Results for r are given in Table 1 for 10 different variables. A low value of r indicates that the dominant signal and noise patterns are dissimilar, a situation favorable for detection. The lowest values of r are for the vertical structure of the zonally-averaged u-velocity component ($r = 0.29$) and 2m temperature ($r = 0.43$), whereas SLP has the highest pattern correlation ($r = 0.82$), indicating highly similar signal and noise patterns.

Table 1: Spatial correlation r between EOF 1 of the control run and EOF 1 of the Scenario A experiment for 10 different climate variables. Results are ranked from smallest to largest value of r, i.e., in order of increasing similarity between dominant signal and noise patterns.

No.	Variable	Correlation	EOF1 CTL[1] (%)	EOF1 SZA[2] (%)
1	Vertical U-Velocity	0.29	23.6	80.5
2	2 m Temperature	0.43	52.5	84.1
3	Vertical Temperature	0.46	61.4	97.6
4	Precipitable Water	0.50	27.7	89.5
5	Precipitation Rate	0.55	12.8	17.3
6	Total Cloud Cover	0.57	16.1	22.4
7	10 m U-Velocity	0.67	17.4	28.4
8	10 m V-Velocity	0.69	20.1	25.9
9	Vertical V-Velocity	0.72	29.5	56.3
10	Sea Level Pressure	0.82	27.7	36.9

[1] Explained variance for EOF 1 of control run.

[2] Explained variance for EOF 1 of Scenario A.

The control run EOF 1 pattern for the vertical profile of the u-velocity component has largest variance in the Southern Hemisphere, with a pronounced dipole structure extending throughout the model atmosphere (Figure 1a). This pattern is similar to that obtained by James and James (1989) in an experiment with a simple uncoupled atmospheric model. The dipole structure in the Southern Hemisphere is probably related to interannual variations in the latitudinal location of the subtropical jet. In contrast, the EOF 1 pattern for Scenario A (Figure 1b) has a strong signature in the lower stratosphere, where intensive cooling occurs and the intensity of the jet cores increases in both hemispheres and in both seasons.

A pronounced vertical temperature contrast (stratospheric cooling and tropospheric warming) is a common feature of the greenhouse-gas signals simulated in a number of different model experiments (e.g. Schlesinger and Mitchell, 1987; Cubasch et al., 1992). Recently, Karoly (1987, 1989) has compared such signals with observations, and suggests that vertical temperature

contrasts could be useful for detecting greenhouse-gas-induced climatic change. However, as Liu and Schuurmanns (1990) and Wigley and Barnett (1990) have pointed out, this characteristic signal may resemble the observed pattern of vertical temperature changes associated with low-frequency internally-generated natural variability. Our results suggest that the overall structure of the dominant control run EOF for vertical temperature changes is only weakly correlated (r = 0.46) with the EOF 1 pattern of Scenario A (Table 1). There is, however, some correspondence of signal and noise patterns in the tropical troposphere and stratosphere, which requires further investigation (see Figures 1c,d). In the case of the noise EOF this is probably due to changes in convective activity and the height of the tropopause associated with the overall cooling trend (-0.4°C in globally-averaged 2m temperature) in the control run.

The dominant noise and signal EOFs of vertically-integrated precipitable water (Figures 2a,b) are also only weakly correlated (r = 0.50). The signal EOF is spatially coherent, and is characterized by hemispheric and zonal symmetry, with highest variance in the tropics and low variance towards the poles. This primarily describes the change in mean state (increases in precipitable water at virtually all grid-points, with the largest increases over the tropical oceans). The first EOF of the control run is less coherent spatially. In contrast, the first noise and signal EOFs of SLP (Figures 2c,d) are both characterized by strong zonal structure in the vicinity of the Antarctic Circumpolar Trough, an unfavorable situation for signal detection.

The spatial patterns of the dominant surface temperature EOFs show a number of interesting features (Figure 3). First, the EOF 1 patterns of three greenhouse warming experiments (IPCC Scenarios A and D, and a 2xCO_2 experiment) are highly similar despite the differences in forcing history (see Table 2). This similarity is primarily due to a common land-sea contrast component of the signal.

Table 2: Spatial correlation matrix for EOFs 1 and 2 of Scenarios A, D, 2xCO_2 and control run. Results are for annually-averaged 2 m temperature. Correlations indicating $\geq$ 50% common spatial variance are indicated by an asterisk.

			EOF 1				EOF 2			
EXP	**EOF**	**EV(%)[1]**	**SZA[2]**	**SZD[3]**	**TCO[4]**	**CTL[5]**	**SZA**	**SZD**	**TCO**	**CTL**
SZA	1	84.1	1.00	0.82*	0.97*	0.43	0.30	0.22	0.33	0.35
SZD	1	46.8		1.00	0.75*	0.81*	0.65	0.03	0.38	0.26
TCO	1	87.8			1.00	0.35	0.22	0.65	0.34	0.31
CTL	1	52.5				1.00	0.74*	0.16	0.37	0.06
SZA	2	3.1					1.00	0.47	0.14	0.24
SZD	2	12.4						1.00	0.65	0.86*
TCO	2	2.7							1.00	0.59
CTL	2	18.6								1.00

[1] EV: Explained variance.
[2] SZA: Scenario A.
[3] SZD: Scenario D.
[4] TCO: 2xCO_2.
[5] CTL: Control.

Second, the EOF 1 patterns in the Scenario A and 2xCO_2 experiments are only weakly correlated

with the dominant noise mode ($r = 0.43$ and $r = 0.35$, respectively; see Table 2). We note, however, that the variability exhibited in the region of the Antarctic Circumpolar Current (ACC) is at least qualitatively similar in the first EOFs of the control run and the three greenhouse warming integrations. In the control run – unlike Scenario A and $2xCO_2$ – the variability in the Ross and Weddell Seas is generally out of phase with variability in the tropics and in the Northern Hemisphere (Figure 3). The higher correlation between EOF 1 patterns of the control and Scenario D ($r = 0.81$) is in part due to similar phase relationships between changes in the Arctic and in the ACC. Third, the high pattern correlation between EOF 2 of Scenario A (which explains only 3.1% of the variance) and EOF 1 of the control run ($r = 0.74$; Table 2) suggests at least a small common component of natural variability and/or drift in the two experiments. There is no evidence for such a common component in the $2xCO_2$ experiment ($r = 0.37$ for EOF 1 control versus EOF 2 of $2xCO_2$).

3 Signal-to-Noise Ratios

The correlations given in the previous section measure only the pattern similarity of the normalized signal and noise EOFs, and provide no information about the relative amplitudes of signal and noise. Various metrics can be defined in order to provide such information.

We first consider the number of EOFs required in order to explain 95% of the variance of the control run (k_c) and Scenario A data (k_s; see Table 3). k_c and k_s are measures of the spatio-temporal coherence and amplitude of signal and noise. k_c is largest for variables with limited spatial coherence, such as precipitation rate ($k_c = 80$) and total cloud cover ($k_c = 84$), and smallest for variables with highly coherent spatial structures, e.g. vertical temperature contrasts ($k_c = 24$). In the Scenario A data, k_s is smallest for variables with spatially-coherent signals and large signal amplitude, e.g. vertical temperature contrast ($k_s = 1$) and 2m temperature ($k_s = 19$). These results indicate that vertical temperature contrast, the vertical structure of zonal winds, 2m temperature and precipitable water are characterized by low-dimensional signals embedded in a much higher-dimensional noise space.

How well can we represent the Scenario A (control) data in the truncated space of the first 99 control run (Scenario A) EOFs? In order to address this question we form the projections

$$\alpha_j^s(t) = \sum_{x=1}^{p} s(x,t) e_j^c(x) \tag{5}$$

and

$$\beta_j^c(t) = \sum_{x=1}^{p} c(x,t) e_j^s(x) \tag{6}$$

$j = 1, \ldots, q\,; \quad t = 1, \ldots, n$

and then compute the ratio of the variance of the projection relative to the variance of the original data set, e.g., for the projection of the control data onto the signal EOFs

$$\text{VAR1} = \left[\sum_{j=1}^{q} \sum_{t=1}^{n} \left[\beta_j^c(t)\right]^2 \right] / \left[\sum_{x=1}^{p} \sum_{t=1}^{n} c(x,t)^2 \right] \tag{7}$$

Table 3: Explained variance for projection of control data onto Scenario A EOFs, and Scenario A data onto control EOFs. Results are for k_c, the number of control run EOFs required in order to explain 95% of the control variance (bold) and k_s, the number of Scenario A EOFs required to explain 95% of the Scenario A variance. VAR1 is a measure of how well the control data (bold) can be represented in the truncated space of the first 99 Scenario A EOFs and how well the signal data can be represented in the space of the first 99 control EOFs. VAR2 and VAR3 are analagous to VAR1, but now the projections are restricted to EOF 1 of the control run (bold) and Scenario A and EOF 2 of the control run (bold) and Scenario A.

No.	Variable	k_c[1]	k_s[2]	VAR1[3]		VAR2[4]		VAR3[5]	
1	Vertical Temperature	**24**	1	**89.1**	96.6	**31.5**	20.6	**2.5**	18.4
2	Vertical U-Velocity	**24**	9	**98.2**	98.9	**13.3**	9.4	**28.0**	18.7
3	Vertical V-Velocity	**35**	27	**93.1**	94.1	**30.7**	17.2	**8.8**	6.6
4	Sea Level Pressure	**48**	44	**82.0**	86.5	**26.8**	18.3	**4.4**	9.4
5	Total Cloud Cover	**84**	84	**32.4**	32.8	**8.7**	5.4	**0.4**	5.7
6	10 m U-Velocity	**64**	62	**73.7**	73.3	**14.6**	10.6	**7.9**	6.5
7	10 m V-Velocity	**70**	69	**64.6**	64.4	**13.9**	10.4	**3.5**	6.6
8	2 m Temperature	**40**	19	**66.4**	81.8	**2.6**	1.2	**2.1**	30.2
9	Precipitable Water	**75**	21	**78.1**	57.5	**37.7**	12.3	**0.3**	10.5
10	Precipitation Rate	**80**	80	**43.0**	45.0	**6.5**	3.9	**2.1**	5.5

[1] Number of control run EOFs required in order to explain $\geq$ 95% of the total control run variance.

[2] Number of Scenario A EOFs required in order to explain $\geq$ 95% of the total Scenario A variance.

[3] Percentage of total **Scenario A** (control run) variance explained by projection of **Scenario A** (control run) anomaly data onto first 99 **control run** (Scenario A) EOFs.

[4] Percentage of total **Scenario A** (control run) variance explained by projection of **Scenario A** (control run) anomaly data onto EOF 1 of **control run** (Scenario A).

[5] Percentage of total **Scenario A** (control run) variance explained by projection of **Scenario A** (control run) anomaly data onto EOF 2 of **control run** (Scenario A).

VAR1 is a measure of the overall orthogonality of signal and noise data sets. The results are a little surprising. For total cloud cover, for example, only 32.4% (32.8%) of the total Scenario A (control) variance can be explained by the projection onto the first 99 control (Scenario A) EOFs, while for the zonally-averaged u-velocity component, virtually all of the variance of the original data sets can be explained by the projection onto either signal or noise EOFs. Does this imply that the 'detection potential' of changes in cloud cover is high? Probably not. We know from the k_c and k_s results (Table 3) and from visual inspection of the EOF patterns that cloud cover is extremely noisy spatially in both Scenario A and control data sets. A time-history PC selection rule (Rule Q; Preisendorfer, 1988) also revealed that the control and signal PC time series for cloud cover (for PCs higher than 2) were indistinguishable from white noise. For variables such as total cloud cover and precipitation rate, where both k_c and k_s are large, VAR1 is not a meaningful indicator of the true orthogonality of signal and noise. This is due to the problem of estimating the EOFs of a high-dimensional signal and noise in a situation where the spatial dimension p is many times larger than the number of time samples n. VAR1 is only a useful indicator of orthogonality where a relatively small number of EOFs characterize signal and noise (i.e., where k_c and k_s are much smaller than n), and the EOFs can be estimated reliably.

The VAR1 results are not highly sensitive to whether the noise is projected onto the signal EOFs or the signal is projected onto the noise EOFs. The most notable exceptions are 2m

temperature and precipitable water. In the former case, the explained variance is much lower for the projection of the signal onto the noise EOFs (66.4% vs. 81.8%). This difference is due to the fact that the dominant EOF 1 pattern of Scenario A (see Figure 3), which is spatially coherent and represents warming at almost all grid-points, is not represented in the control run. In the case of precipitable water, the VAR1 results differ because the EOF 1 pattern of the control run is more strongly represented in the signal than the EOF 1 pattern of the signal is represented in the noise (see Figures 2a,b).

These results are confirmed by VAR2 and VAR3, which are analagous to VAR1, but in which the projection is restricted to EOF 1 and EOF 2 of signal and noise. VAR2 and VAR3 provide insights into the extent to which the signal is contaminated by the dominant noise modes. For precipitable water, SLP and the vertical signal in temperature and meridional winds, the signal is strongly contaminated by the EOF 1 mode of the control run. This situation is unfavorable for detection purposes. Contamination of the signal by the first noise mode is lowest for 2m temperature (2.6%). The fact that VAR3 is large (30.2%) for 2m temperature when the noise is projected onto Scenario A EOF 2, but small when the signal is projected onto the control EOF 2 (2.1%) indicates that there is a common natural variability or drift component, but its amplitude is much larger in the control run than in Scenario A.

4 Conclusions

For detection studies, we would like to focus on variables with signal and noise patterns which are highly dissimilar (Barnett and Schlesinger, 1987). If signal and noise patterns are not orthogonal, the chances of mistaking natural variability for a greenhouse-gas signal are increased. Our results suggest that vertical changes in zonally-averaged winds (u-component), 2m temperature, vertical temperature contrasts, and vertically-integrated precipitable water have favorable orthogonality relationships, with less than 25% common variance between signal and noise EOF 1 patterns. Some features require further investigation, in particular the correspondence of signal and noise patterns for vertical temperature contrasts in the tropical troposphere and stratosphere, and the similarity between signal and noise 2m temperature patterns in the vicinity of the Antarctic Circumpolar Current. Sea level pressure has highly similar signal and noise EOF 1 patterns, and is unlikely to be a useful component of a multivariate 'fingerprint' detection vector.

The EOF analysis indicated that 2m temperature, precipitable water, and the vertical distribution of zonal temperature and zonal winds are characterized by low-dimensional signals embedded in a higher-dimensional noise space. In contrast, cloud cover and precipitation rate have high-dimensional signals and noise. In order to determine the degree to which the signal is contaminated by the natural variability (and/or drift) of the control, we projected the Scenario A data onto EOFs 1 and 2 of the control. Signal contamination by the EOF 1 and 2 modes of the noise was lowest for 2m temperature, a situation favorable for detection. The signals for precipitable water, SLP, and the vertical structure of zonal temperature and zonal winds were significantly contaminated by the dominant noise modes.

It should be stressed that the signal-to-noise results presented here are subject to considerable uncertainty. The uncertainties stem from two separate areas – the so-called "cold start" problem (Hasselmann et al., 1992) and the problem of discriminating between drift and natural variability.

The work presented here does not explicitly consider the time dimension of the signal-to-noise problem. In a separate investigation (Santer et al., 1992), we have applied the method of

Bloomfield and Nychka (1992) in order to determine the significance of a 100-year linear trend in a greenhouse warming experiment relative to the variance of linear trends on the century time scale, as estimated from a long "ocean only" stochastic forcing experiment (Mikolajewicz and Maier-Reimer, 1990). Future signal-to-noise studies should employ some form of "optimal detection" technique. In the context of time-dependent greenhouse warming experiments, this will involve optimization of the combined space-time structure of a multivariate signal vector relative to the noise (Hasselmann, 1992).

References

Barnett, T.P., 1986: Detection of changes in global tropospheric temperature field induced by greenhouse gases. *Journal of Geophysical Research* **91**, 6659-6667.

Barnett, T.P. and Schlesinger, M.E., 1987: Detecting changes in global climate induced by greenhouse gases. *Journal of Geophysical Research* **92**, 14,772-14,780.

Bloomfield, P. and D. Nychka, 1992: Climate spectra and detecting climate change. *Climatic Change* (In press).

Cubasch, U., Hasselmann, K., Höck, H., Maier-Reimer, E., Mikolajewicz, U., Santer, B.D. and Sausen, R., 1991: Time-dependent greenhouse warming computations with a coupled ocean-atmosphere model. *Climate Dynamics* (In press).

Hasselmann, 1992: Optimal signal detection (In preparation).

Hasselmann, K., R. Sausen, E. Maier-Reimer, and R. Voss, 1992: On the cold start problem in transient simulations with coupled ocean-atmosphere models. (Submitted to *Climate Dynamics*).

James, I.N. and James, P.M., 1989: Ultra-low-frequency variability in a simple atmospheric circulation model. *Nature* **342**, 53-55.

Karoly, D.J., 1987: Southern Hemisphere temperature trends: A possible greenhouse gas effect? *Geophysical Research Letters* **14**, 1139-1141.

Karoly, D.J., 1989: Northern Hemisphere temperature trends: A possible greenhouse gas effect? *Geophysical Research Letters* **16**, 465-468.

Liu, Q. and Schuurmanns, C.J.E., 1990: The correlation of tropospheric and stratospheric temperatures and its effect on the detection of climate changes. *Geophysical Research Letters* **17**, 1085-1088.

MacCracken, M.C. and Moses, H., 1982: The first detection of carbon dioxide effects: Workshop Summary, 8-10 June 1981, Harpers Ferry, West Virginia. *Bulletin of the American Meteorological Society* **63**, 1164-1178.

Madden, R.A. and Ramanathan, V., 1980: Detecting climate change due to increasing carbon dioxide. *Science* **209**, 763-768.

Mikolajewicz, U. and E. Maier-Reimer, 1990: Internal secular variability in an ocean general

circulation model. *Climate Dynamics* 4, 145-156.

Preisendorfer, R.W., 1988: Principal component analysis in meteorology and oceanography. *Developments in Atmospheric Science* **17**, Elsevier, Amsterdam, 425 pp.

Santer, B.D., W. Brüggemann, U. Cubasch, K. Hasselmann, H. Höck, E. Maier-Reimer, and U. Mikolajewicz, 1992: Signal-to-noise analysis of time-dependent greenhouse warming experiments: Standard error estimates. (In preparation).

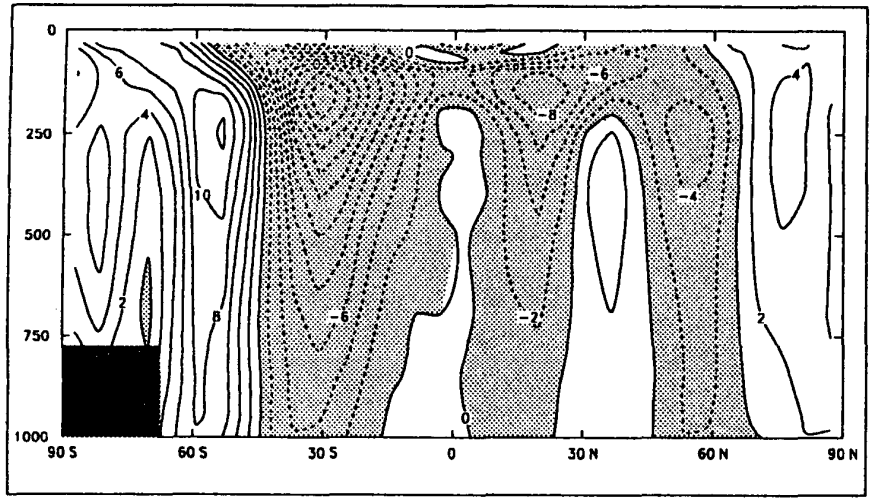
Schlesinger, M.E. and J.F.B. Mitchell, 1987: Climate model simulations of the equilibrium climatic response to increased carbon dioxide. *Reviews of Geophysics* **25**, 760-798.

Wigley, T.M.L. and Barnett, T.P., 1990: Detection of the greenhouse effect in the observations. In: *Climate Change. The IPCC Scientific Assessment* (Eds. J.T. Houghton, G.J. Jenkins and J.J. Ephraums). Cambridge University Press, Cambridge, pp 239-256.

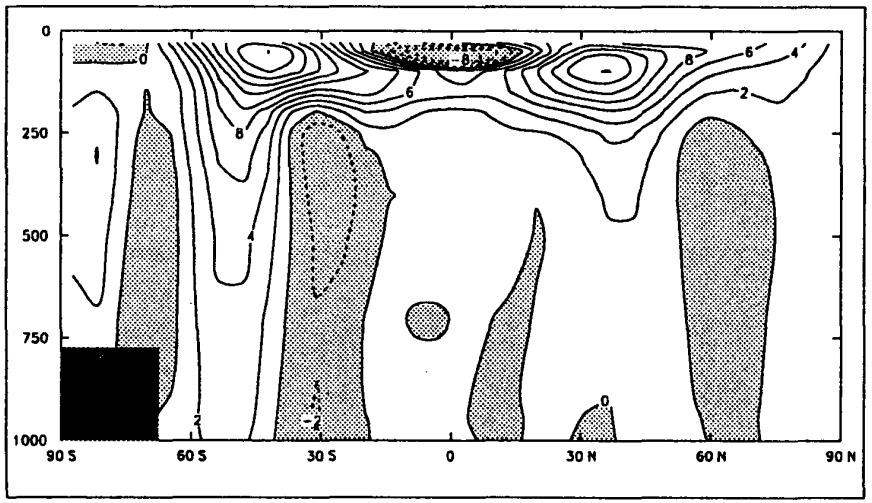
Wigley, T.M.L. and Jones, P.D., 1981: Detecting CO_2-induced climatic change. *Nature* **292**, 205-208.

Figure 1: EOF 1 of the annually-averaged changes in the vertical structure of zonal wind for the control run (a) and Scenario A (b) and for vertical temperature contrast for the control run (c) and Scenario A (d).

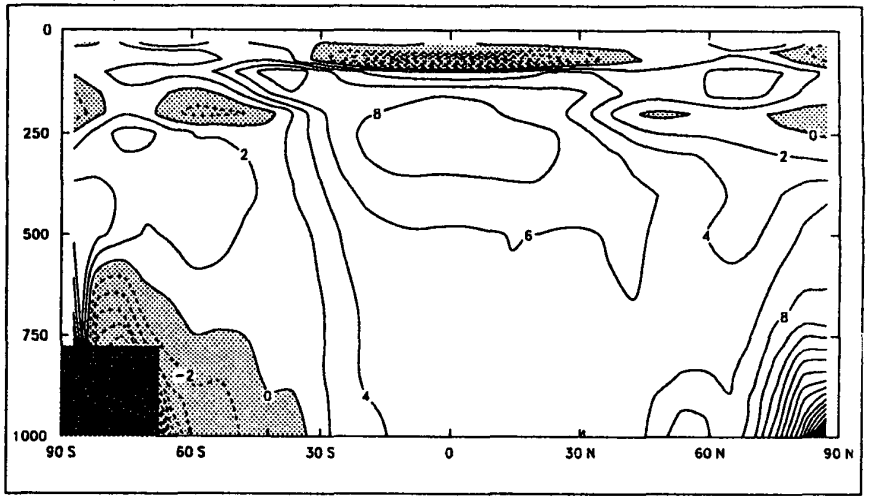
a) EOF 1 ZONAL WIND: CONTROL (EV 23.6%)



b) EOF 1 ZONAL WIND: SCENARIO A (EV 80.5%)



c) EOF 1 TEMPERATURE: CONTROL (EV 52.5%)



d) EOF 1 TEMPERATURE: SCENARIO A (EV 97.6%)

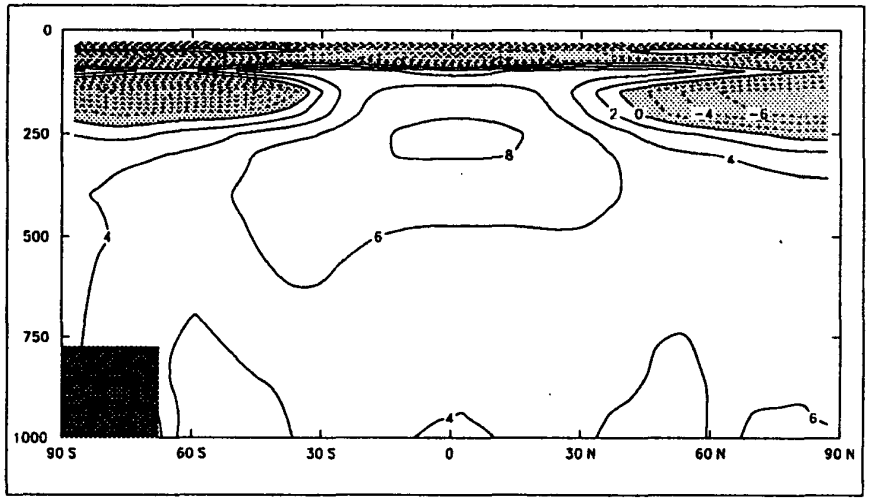
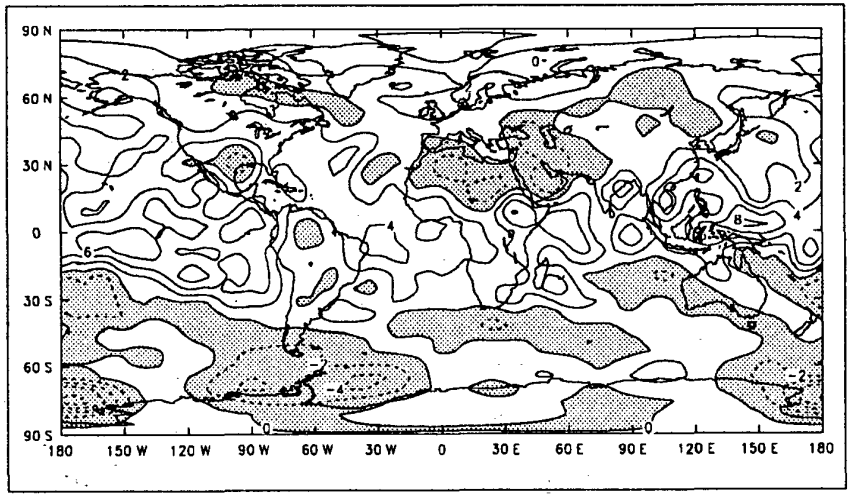
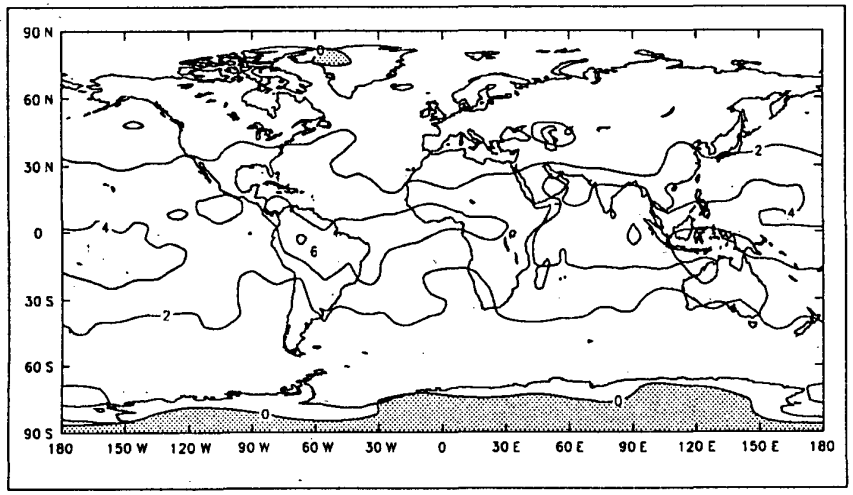


Figure 2: EOF 1 of the annually-averaged changes in the vertically integrated preciptable water for the control run (a) and Scenario A (b) and for sea level pressure (SLP) for the control run (c) and Scenario A (d).

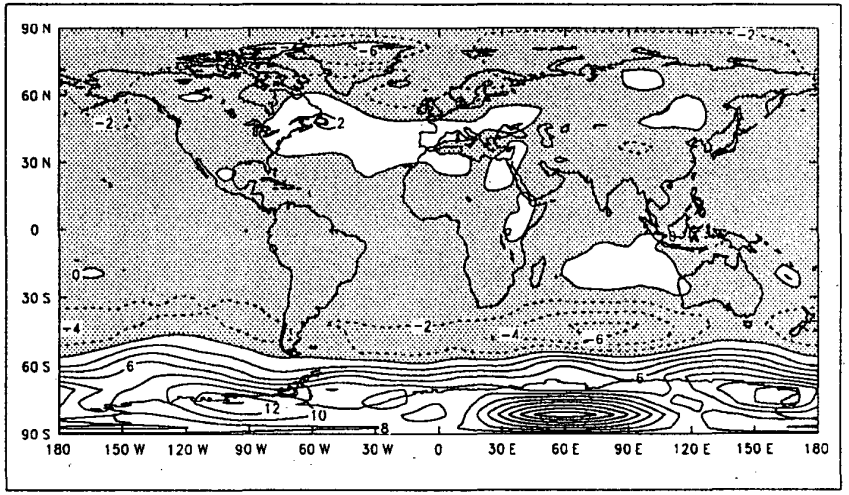
a) EOF 1 PRECIPITABLE WATER: CONTROL (EV 27.7%)

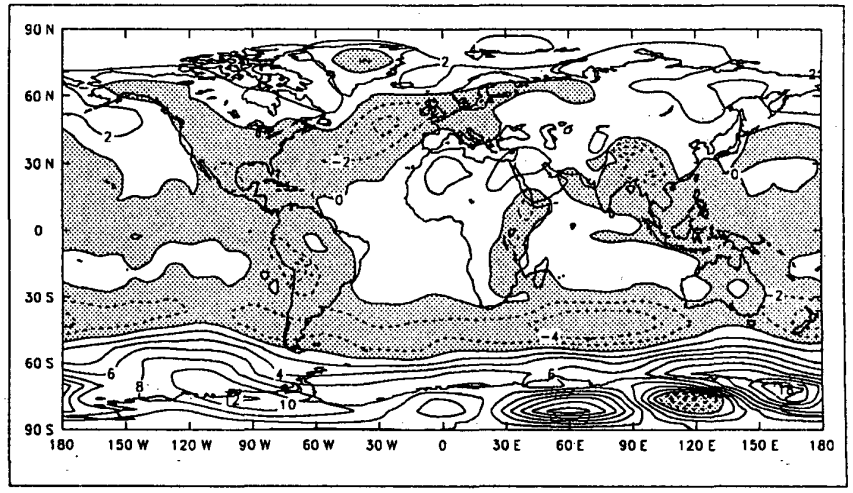




b) EOF 1 PRECIPITABLE WATER: SCENARIO A (EV 89.5%)

c) EOF 1 SLP: CONTROL (EV 27.7%)

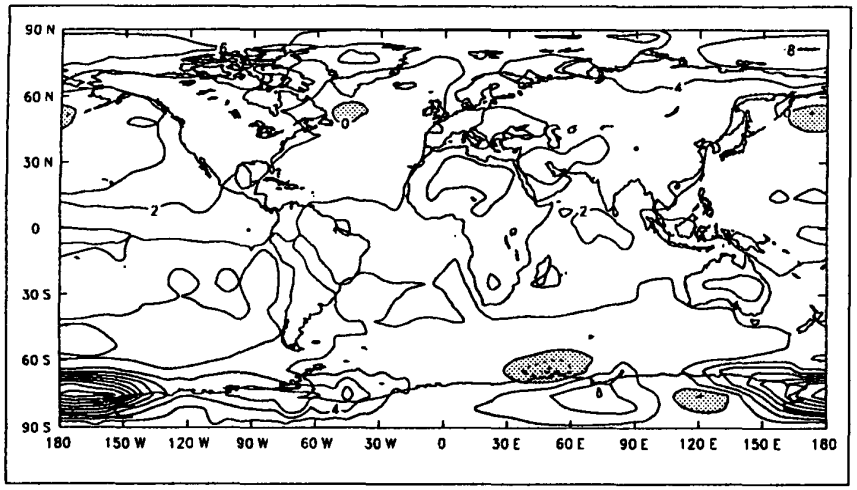




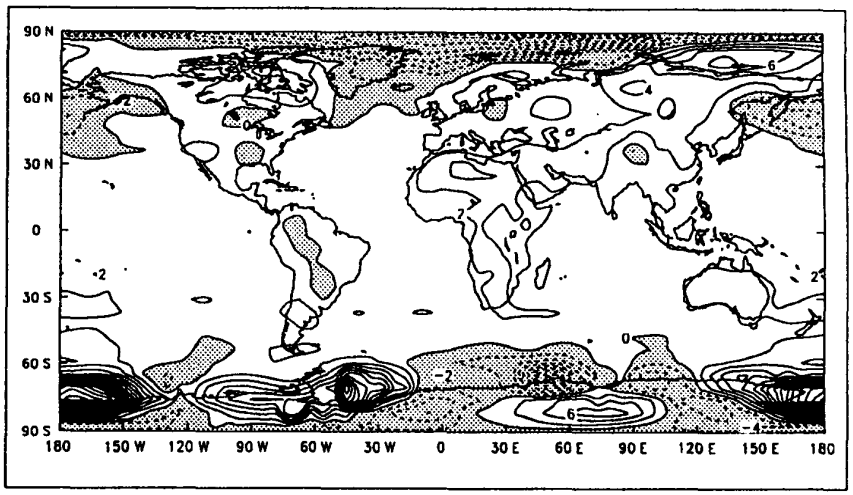
d) EOF 1 SLP: SCENARIO A (EV 36.9%)

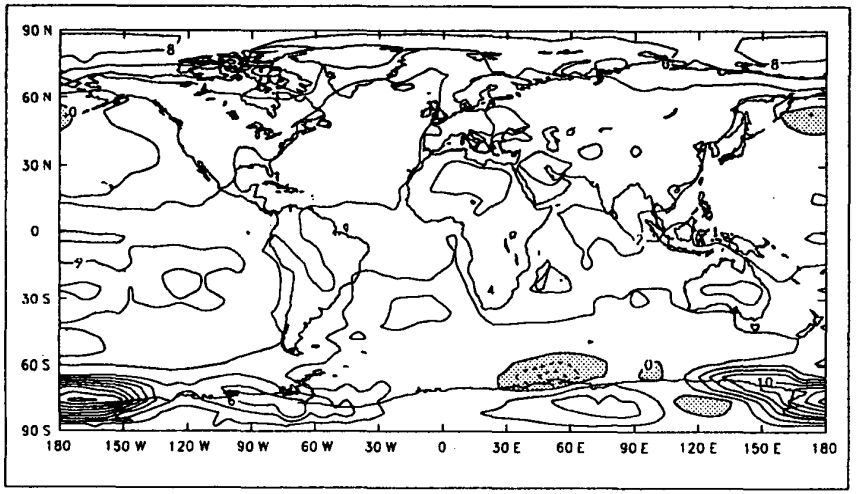
Figure 3: EOF 1 of the annually-averaged 2m temperature changes for Scenario A, Scenario D, 2xCO2 and control run. Anomalies used for computing EOFs were defined relative to the smoothed initial state of the control run.

a) EOF 1: SCENARIO A (EV 84.1%)

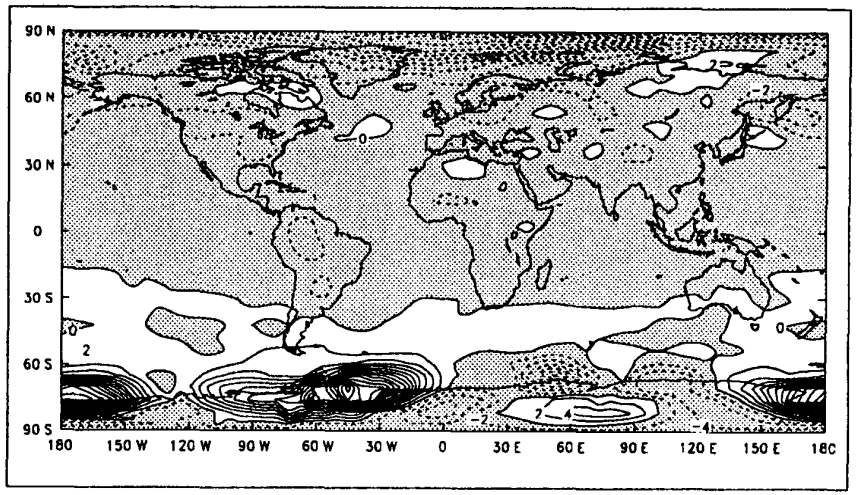


b) EOF 1: SCENARIO D (EV 46.8%)





c) EOF 1: 2 x CO2 (EV 87.8%)



d) EOF 1: CONTROL (EV 52.5%)

Noise Response in a Simple Stochastic Coupled Climate Model

KWANG-YUL KIM

Applied Research Corporation, College Station, Texas

GERALD R. NORTH

Climate System Research Program, Department of Meteorology, Texas A&M University, College Station

ABSTRACT

This paper considers the geographical distribution of various second-moment statistics and the Karhunen-Loève basis functions (principal components) of the surface temperature fluctuations of a simple coupled climate model forced by noise. The model is a surface energy balance model with explicit two-dimensional geography coupled to a diffusive-upwelling ocean. A comparison with observations shows that the detailed geography of variance, spatial and temporal correlations look close to those of observations in the frequency bands of interest to climatologists without too much sensitivity to the form of the driving noise and the prescribed model parameters. The model also provides some insight into the sensitivity and benchmark solution of the Karhunen-Loève basis functions which are a useful piece of information in detecting and predicting climate change.

1. Introduction

One key factor in developing and improving detection and prediction techniques is to understand the natural variability of our climate system. Extracting a wide spectrum, both in space and time, of natural fluctuations from observations is difficult because observational record is rarely long enough, observational data are contaminated by sampling errors, and inadequate spatial coverage biases the spectrum. Therefore, natural variability is often inferred from stochastic climate models. Earlier studies using stochastic climate models include Hasselmann (1976), Lemke (1977), Robock (1978), North et al. (1981), North and Cahalan (1982). These studies did not consider that the model could be used more than in a schematic sense.

Recent studies suggest that noise-forced energy balance models (EBMs) can also be used in quantitative studies of the natural variability of the climate system. Leung and North (1991) demonstrated that a noise-forced EBM can be used to fit the surface temperature fluctuations of a simplified version of the Community Climate Model. North et al. (1992) showed that the response of the same general circulation model (GCM) is essentially linear thereby justifying the use of linear stochastic models. Kim and North (1991) showed that a noise-forced EBM with explicit two-dimensional geography produces various second-moment statistics of the surface temperature fluctuations close to those of observations. In recognition of its importance in modeling interdecadal to century scale natural fluctuations, Kim and North (1992a) developed a simple coupled climate model. The development of a simple coupled stochastic climate model with explicit two-dimensional geography is an important contribution to our limited understanding of the natural fluctuations in the presence of the deep ocean (e.g., Wigley and Raper, 1990a, 1991; Mikolajewicz and Maier-Reimer, 1990; Zebiak and Cane, 1991).

In this study, we summarize the noise response characteristics of a simple coupled climate model developed by the authors. The second-moment statistics and the Karhunen-Loève (hereafter KL) basis functions (principal components) of the surface temperature fluctuation of the model are compared with those of observations. The surface component of the model is essentially the energy balance model used in many earlier studies. Extensive comparisons between EBMs and GCMs indicate that the two have similar sensitivities for many different types of forcing (e.g., Hyde et al., 1989, 1990; Crowley et al., 1991; Kim and North, 1991; North et al., 1992). The surface component is coupled to the deep ocean. The thermodynamics of the deep ocean is approximated by an upwelling-diffusion model (Hoffert et al., 1980; Hoffert and Flannery, 1985; Lebedeff, 1988; Watts and Morantine, 1990; Wigley and Raper, 1990; Morantine and Watts, 1990, Kim et al., 1992; Kim and North, 1992a, 1992b). The one-dimensional upwelling-diffusion model has widely been used in future projections of the temperature trend due to greenhouse warming, aerosols, solar variability and emission reductions (e.g., Hoffert and Flannery, 1985; Wigley and Raper, 1987, 1989, 1990b, 1990c; Bretherton et al., 1990; Hoffert, 1991) and also have been used for studying long-term natural fluctuations of a climate system (Wigley and Raper, 1990a, 1991). Readers interested in the model physics and solution techniques are referred to our previous studies (Kim and North, 1991, 1992a; Kim et al., 1992).

2. Result and discussion

Observational data used here is the subset (1950–1989) of the United Kingdom data archieved at the National Center for Atmospheric Research. Detailed information on the ob-

Corresponding author address: Dr. K.-Y. Kim, Applied Research Corporation, 305 Arguello Drive, College Station, TX 77840.

servational data processing and the construction of various second-moment statistics and KL functions is found in Kim and North (1991, 1992a, 1992b).

a. Second-moment statistics

Figure 1 shows the maps of variance in the 2 month to 10 year band. The resemblance between the distribution of variance and the land/sea geography (essentially the heat capacity in the model) is remarkable, both in the model and observations. The lack of consistence over Antarctica may be due to topographical influence not in the model and/or the sparcity of data over the area. The variance distribution in the 2 month to 1 year band is similar to those shown in Figure 1.

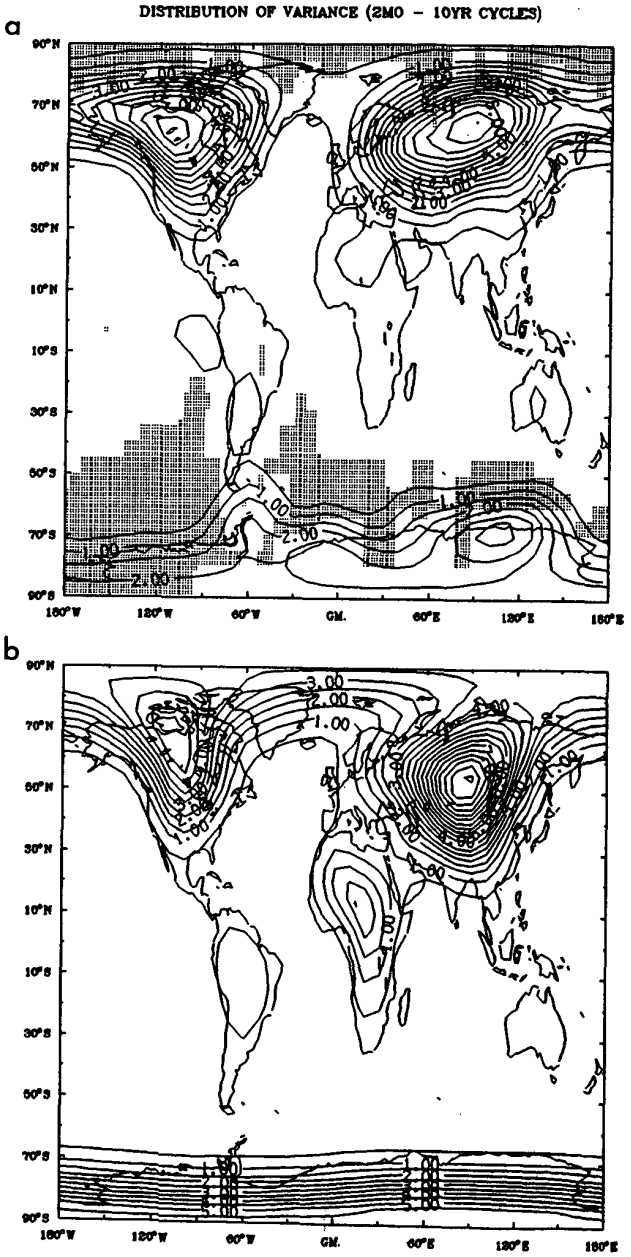


Fig. 1. Geography of the distribution of variance ($°C^2$) of surface temperature fluctuations within 2 month to 10 year band: (a) observations, (b) model. Shaded areas indicate regions with less than 30 years of data available.

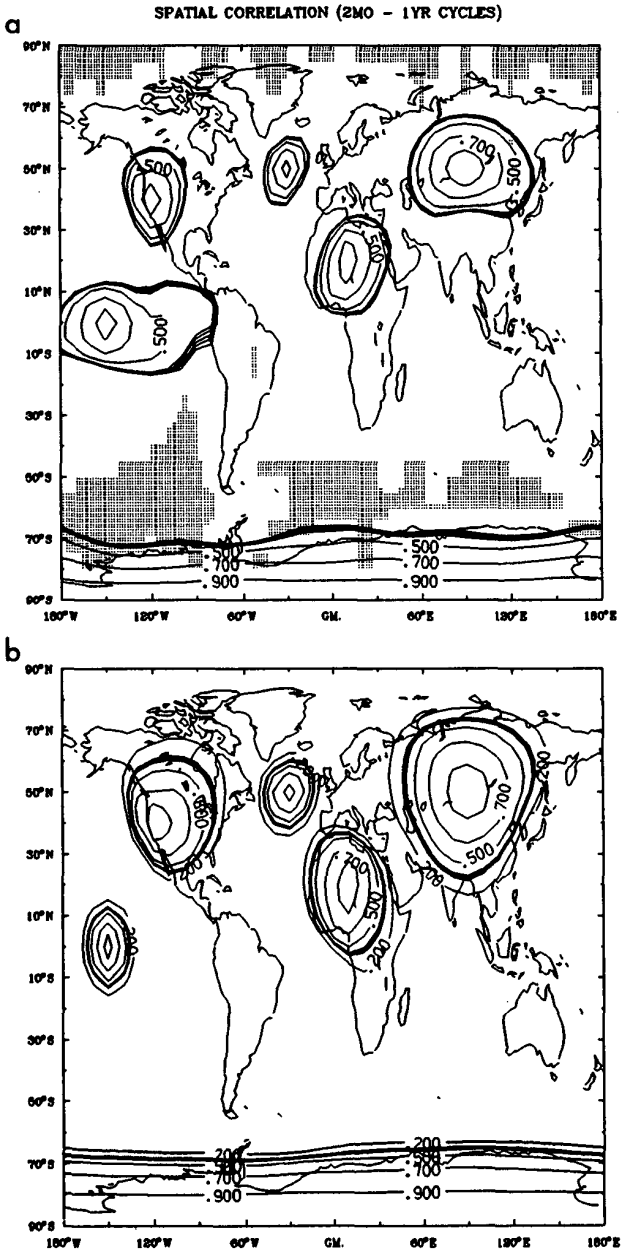
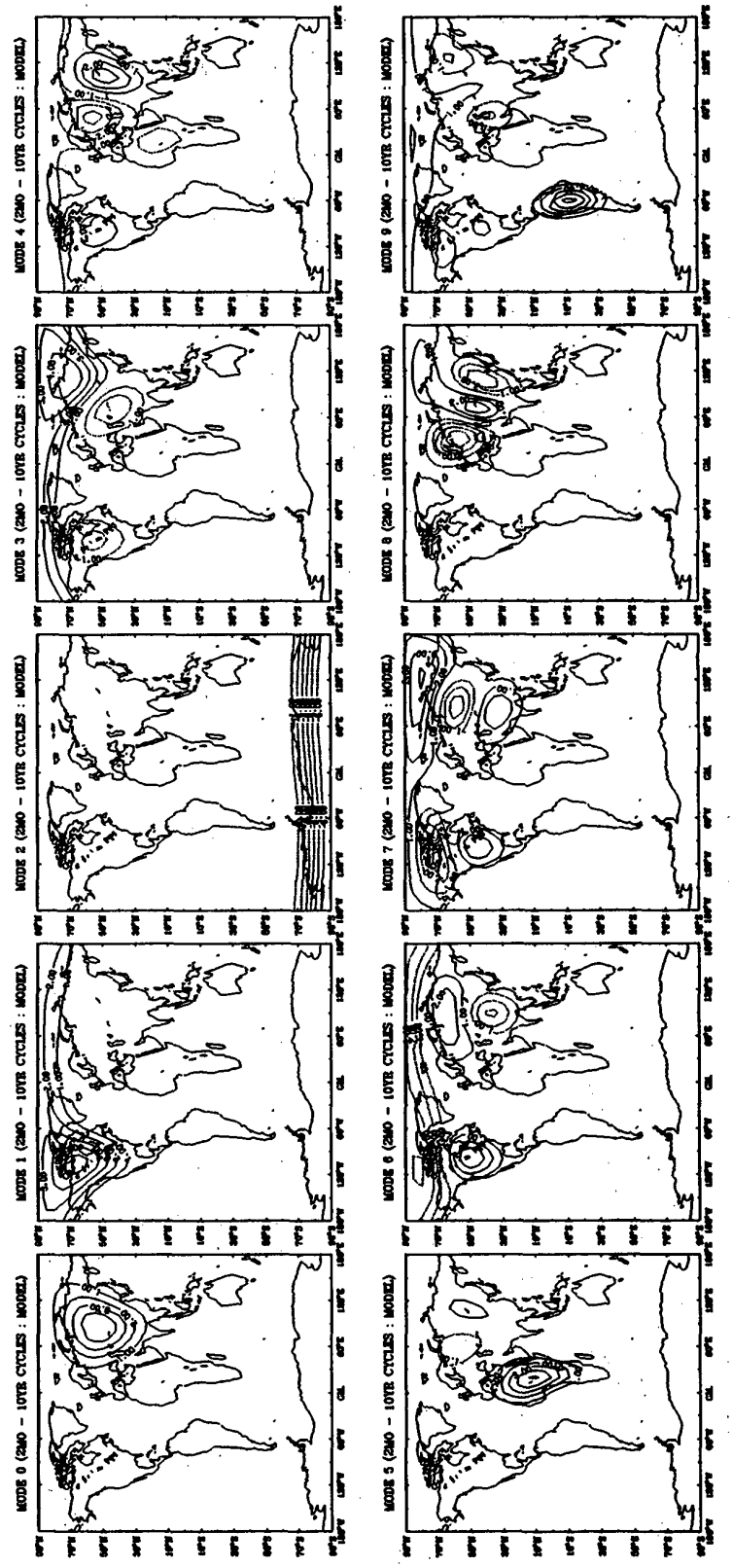


Fig. 2. Spatial correlation of surface temperature fluctuations at six selected test points within 2 month to 1 year band: (a) observation, (b) model. Shaded areas indicate regions with less than 10 years of data available. The test points are (30°W, 50°N), (90°E, 50°N), (120°W, 40°N), (10°E, 20°N), (150°W, 0°N), (0°E, 90°S).



a

Fig. 3. The first 10 eigenfunctions of the covariance matrix in the 2 month to 10 year band: (a) model (Karhunen-Loève basis functions), (b) observations (estimated Karhunen-Loève basis functions).

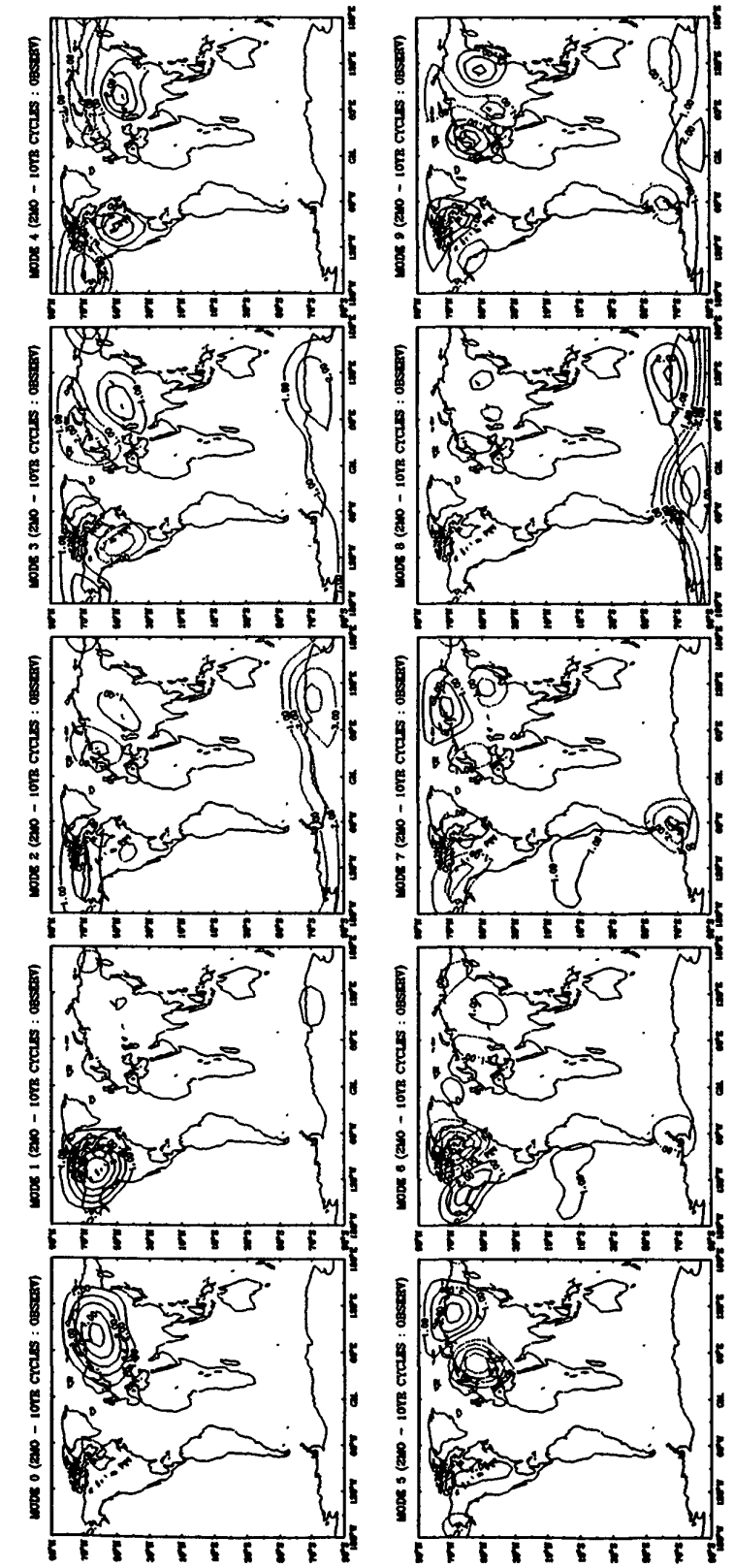


Fig. 3. (continued)

The distribution of variance in the 1 year to 10 year band is more uniform over the globe (figures not shown). One major difference between the model and observations is the variability associated with El Niño. The model does not include this effect.

Figure 2 shows the maps of spatial correlation functions at 6 test points for the 2 month to 1 year band. The model spatial correlation scale is comparable to that of observations except for the obvious interference by El Niño. The spatial correlation scale over land is longer than that over the ocean. This is because the very definition of high- versus low-frequency depends on the relevant relaxation time of the surface medium. The spatial correlation length scale of the interannual (1 year to 10 year) variability, henceforth, is larger than that in Figure 2 (figures not shown).

The model has also been shown to reproduce reasonable temporal correlation scale of surface temperature fluctuations (figures not shown). The temporal correlation scale over land is much shorter than that over the ocean, both in the model and observations. The model somewhat underestimates the correlation time scale over the equatorial oceans.

b. Karhunen-Loève basis functions

Figure 3 shows the KL functions of surface temperature fluctuations in the 2 month to 10 year band. A comparison shows that the first 5 KL functions are consistent with their estimates from observations with high correlation (> 0.5). The similarity between the KL functions and their estimates generally decreases with the mode number. The discrepancy at high mode numbers reflects the difference in the noise response between the model and our climate system, sampling errors, and bias due to spontaneous spatial filtering. The KL functions in the 2 month to 1 year band are similar to those in Figure 3.

The KL functions do not agree well with their estimates from observations in the 1 year to 10 year band. The discrepancy is due partly to the difference between the model and our climate system and partly to sampling errors. Considering that there are only 4 ensembles (40-year data) available for the estimation of the KL functions, the effects of sampling errors could be large. Earlier studies suggest that an accurate estimation of the first few KL functions require about 1000 ensembles (e.g., North et al., 1982). It is not attempted in this paper, however, to estimate the effects of sampling errors (see Kim and North, 1992b).

Figure 4 shows the plots of eigenvalues versus the mode number for three interesting frequency bands. Eigenvalues decrease very slowly both in the model and observations. This closeness of any two adjacent eigenvalues has an important implication on the estimation of the KL functions. Namely, two eigenfunctions whose eigenvalues are close to each other can easily coalesce in the presence of sampling errors (Preisendorfer and Barnett, 1978; North et al., 1982). Figure 4 suggests that an accurate estimation of the KL functions may be difficult from observations contaminated with sampling errors.

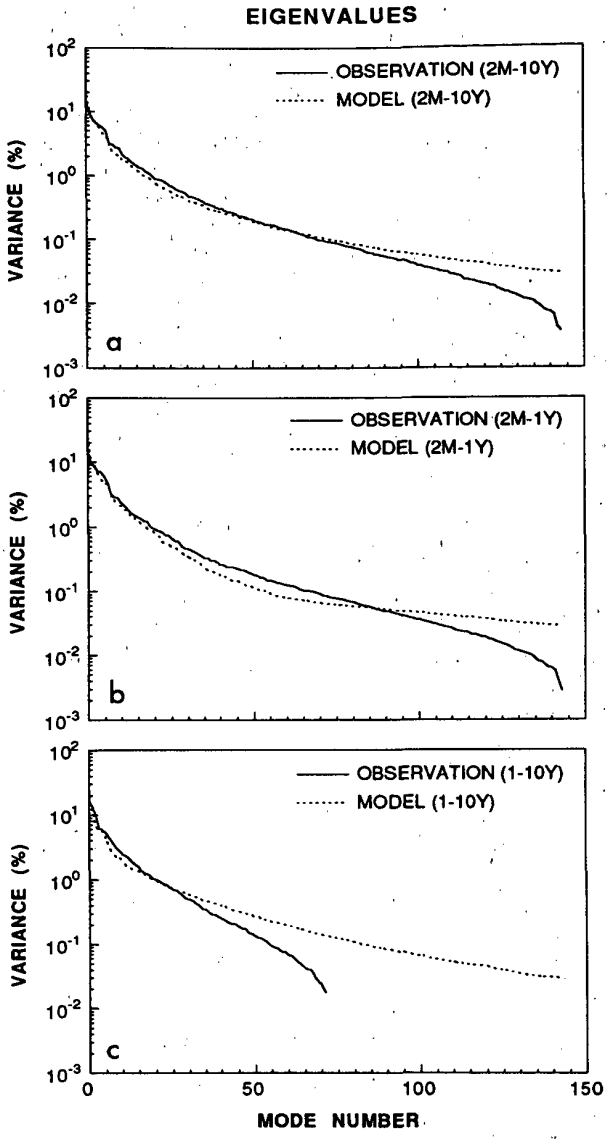


Fig. 4. Plot of eigenvalues versus the mode number: (a) 2 month to 10 year band, (b) 2 month to 1 year band, (c) 1 year to 10 year band. The solid line represents observations and the dashed line represents the model.

3. Summary and concluding remarks

An attempt was made to model various noise-response characteristics of our climate system using a simple coupled climate model. The model is a two-dimensional surface energy balance model coupled to an upwelling-diffusion ocean. The extra dimension allows geography to be explicitly included and greatly enhances the utility of this class of models for climate studies.

The model reproduces faithfully the two-dimensional geography of variance, and spatial and temporal correlation functions of surface temperature fluctuations in the frequency bands interesting to climatologists. Note that the second-moment statistics are important measure of the sensitivity of the model. The land/sea geography (essentially the heat capacity in the model) is vital in the distribution

of various second-moment statistics. It is remarkable that dynamical and topographical considerations play almost no role in the maps. The temporal and spatial correlation scales of the model are also close to those of observations.

The KL basis functions are an important piece of information for developing and improving detection and prediction techniques. The KL functions are difficult to estimate from observations because (1) the observational record is rarely long enough, (2) observations are contaminated by sampling errors, and (3) inadequate spatial coverage of observations introduces a bias. The KL functions were modeled in three frequency bands. The first few (gravest) modes of the KL functions agree with their estimates, the variance of which is likely to be much larger than that of sampling errors. Comparisons between the KL functions and their estimates from observations also provide some insight into the sensitivity to sampling errors and the benchmark solution of the KL functions.

Acknowledgements. We thank the Department of Energy for its support of this work via a grant to Applied to Research Corporation through CHAMMP (DE-FG05-91ER61221). The Department of Energy does not necessarily endorse any of the conclusions drawn in the paper. We have benifited from many discussions with T. J. Crowley.

REFERENCE

Bretherton, F. P., K. Bryan, and J. D. Woods, 1990: Time-dependent greenhouse-gas-induced climate change. *Climate Change, The IPCC Scientific Assessment*, Cambridge University Press, 173–193.

Crowley, T. J., S. K. Baum, and W. T. Hyde, 1991: Climate model comparison of Gondwanan and Laurentide glaciations. *J. Geophys. Res.*, **96**, 9217–9226.

Hasselmann, K., 1976: Stochastic climate models, part I. Theory. *Tellus*, **28**, 473–484.

Hoffert, M. I., 1991: The effects of solar variability on climate. *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations*, Elsevier, 413–428.

——, A. J. Callegari, and C.-T. Hsieh, 1980: The role of deep sea heat storage in the secular response to climatic forcing, *J. Geophys. Res.*, **85**, 6667–6679.

——, and B. F. Flannery, 1985: Model projections of the time-dependent response to increasing carbon dioxide. *Projecting the Climate Effects of Increasing Carbon Dioxide*, U.S. Dept. of Energy, Washington D. C., Report DOE/ER-0237, 149–190.

Hyde, W., T. J. Crowley, K.-Y. Kim and G. R. North, 1989: Comparison of GCM and energy balance model simulations of seasonal temperature changes over the past 18000 years. *J. Clim.*, **2**, 864–887.

——, K.-Y. Kim, T. J. Crowley, and G. R. North, 1990: On the relationship between polar continentality and climate: Studies with a nonlinear seasonal energy balance model, *J. Geophys. Res.*, **95**, 18,653–18,668.

Kim, K.-Y., and G. R. North, 1991: Surface temperature fluctuations in a stochastic climate model, *J. Geophys. Res.*, **96**, 18,573–18,580.

——, and ——, 1992a: Seasonal cycle and second-moment statistics of a simple coupled climate system, *J. Geophys. Res.*, submitted.

——, and ——, 1992b: Principal component analysis of surface temperature fluctuations in a simple coupled climate system, *J. Clim.*, submitted.

——, ——, J. Huang, 1992: On the transient response of a simple coupled climate system, *J. Geophys. Res.*, in press.

Lebedeff, S. A., 1988: Analytical solution of the box diffusion model for a global ocean, *J. Geophys. Res.*, **93**, 14,243–14,255.

Lemke, P., 1979: Stochastic climate models, part 3, Application to zonally averaged energy models, *Tellus*, **29**, 385–392.

Leung, L., and G. R. North, 1991: Atmospheric variability on a zonally symmetric all land planet, *J. Clim.*, **4**, 753–765.

Morantine, M., and R. G. Watts, 1990: Upwelling diffusion climate model: Analytical solutions for radiative and upwelling forcing, *J. Geophys. Res.*, **95**, 7563–7571.

North, G. R., and R. F. Cahalan, 1982: Predictability in a solvable stochastic climate model, *J. Atmos. Sci.*, **38**, 504–513.

——, ——, and J. A. Coakley, 1981: Energy balance climate models, *Rev. Geophys. Space Phys.*, **19**, 91–121.

——, K. Yip, L. Leung, and R. Chervin, 1992: Forced and free variations of the surface temperature field in a general circulation model, *J. Clim.*, **5**, 227–239.

Robock, A., 1978: Internally and externally caused climate change. *J. Atmos. Sci.*, **35**, 1111–1122.

Watts, R. G., and M. Morantine, 1990: Rapid climatic change and the deep ocean. *Clim. Change*, **16**, 83–97.

Wigley, T. M. L., and S. C. B. Raper, 1987: Thermal expansion of sea water associated with global warming. *Nature*, **330**, 127–131.

——, and ——, 1990a: Natural variability of the climate system and detection of the greenhouse effect, *Nature*, **344**, 324–327.

——, and ——, 1990b: Climate change due to solar irradiance changes. *Geophys. Res. Lett.*, **17**, 2169–2172.

——, and ——, 1990c: Future changes in global-mean temperature and thermal-expansion-related sea level rise. *Climate and Sea level Change*, Cambridge University Press, pp. .

——, and ——, 1991: Internally generated natural variability of global-mean temperatures. *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations*, Elsevier, 471–482.

TEMPORAL AND SPATIAL CHANGES IN LARGE DAILY PRECIPITATION EVENTS ACROSS THE SOUTHEASTERN UNITED STATES

David Changnon, Clay Lawson, and David J. Smith
Southeast Regional Climate Center
Columbia, South Carolina

1. INTRODUCTION

For many years, information resulting from the study of heavy precipitation events has been widely used to design water control structures such as storm sewers, dams, culverts, and bridges. In recent years extreme events studies, such as large rainstorms, have also been used to monitor climate trends and change (Huff and Changnon, 1987). The Huff and Changnon study indicated significant changes in the number and spatial distribution of heavy precipitation events in Illinois from an earlier period to a later period. Because there are continual changes in regional climates, the most updated climate information should be made available and used in the design of important structures.

The Southeast Regional Climate Center, in association with the National Weather Service (NWS) Office of Hydrology, is initiating a process to update rainfall frequencies across six southeastern states - Alabama, Florida, Georgia, North Carolina, South Carolina, and Virginia. This process will incorporate information from meteorological, climatological, hydrological, and statistical resources within the region.

To justify the update of rainfall frequencies developed in NOAA Technical Paper Number 40 (Hershfield, 1961), analyses of daily (24-hour) precipitation amounts were conducted to identify changes in the temporal and spatial distribution of the heavy rainstorms. Various users that require information on extreme precipitation events have suggested that in addition to updating rainfall frequencies they need to know the frequency that various rainfall thresholds are surpassed. Also, there is need to have seasonal as well as annual information for design of water control structures. This paper will present the results of these analyses.

2. DATA AND METHODOLOGY

Daily precipitation data were obtained for NWS cooperative stations with 60 years (1930-1989) of digitized record. In the Southeast, 89 stations were found to meet this requirement and have less than two percent of the daily values missing during the 60 years. These digital data were provided by the National Climatic Data Center (NCDC). Although the number of stations with digital records beginning in 1948 was larger, to analyze the temporal trends a longer record was required. Figure 1 shows the location of the stations and the climatic divisions in the six southeastern states. The number of stations per state varies greatly with eight available in Alabama, 14 in Florida, 22 in Georgia, 20 in North Carolina, 14 in South Carolina and 11 in Virginia.

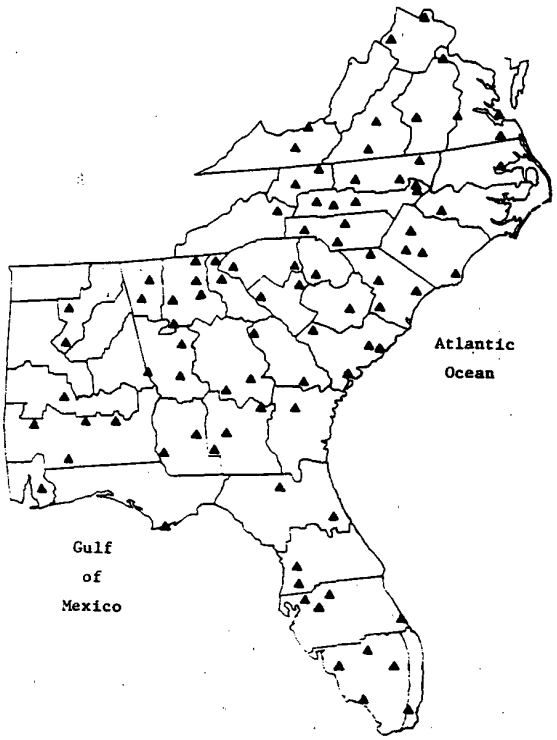


Figure 1. Location of NWS cooperative stations used in the analyses.

Two basic questions were determined from discussions with practicing civil engineers in the region: 1) what is the annual and seasonal frequency of exceeding a given rainfall threshold and 2) how

has the rainfall amount for a given return interval changed. A 24-hour rainfall amount that exceeded 50 mm was determined to be an important threshold in the design of various water control structures and was examined both annually and seasonally. Results described in a study by Robinson and Walsh (1991) indicated that further analysis of precipitation events above critical design thresholds was important in providing user groups with needed information.

To analyze the occurrence of temporal changes in heavy precipitation events, the annual time series and decadal averages of the heavy events were determined. The natural variability in the annual and seasonal time series was large in most cases and made it difficult to identify any long-term trend. The short period of record (60 years) made it difficult to discern reliable information about climatic trends in the decadal averages. These results are similar to those described in a study of Midwest heavy precipitation events (Huff and Angel, 1990) and indicated that the comparison of two long periods may give the most useful information about regional climatic trends. Thus, the frequency distributions of 24-hour storms were analyzed for two 30-year periods, 1930-59 and 1960-89. The frequency distributions were determined both annually and seasonally from the partial duration series of 24-hour events at the 89 stations. The five-year return rainstorm value was determined from the data at each station. The change between the two periods was determined by dividing the five-year return storm value for the later period (1960-1989) by the five-year return storm value for the earlier period (1930-1959). If the ratio of these two values was greater than one, then the five-year storm value during the most recent period was greater than the five-year storm value during the earlier period. The results of annual and seasonal temporal changes will be described.

3. RESULTS AND DISCUSSION

The seasonal and annual number of daily precipitation amounts greater than 50 mm (big rain events) were determined for each station. The total number of big rain events through the 60-year period is shown in Figure 2 and ranged from greater than 300 (an average of five or more per year) in southern Florida, southwestern Alabama, and northeastern Georgia to less than 50 (average of less than one big rain event a year) in western Virginia. The number of big rain events in each season as a percent of the 60-year was determined. The maps for summer and winter are shown in Figures 3a-b.

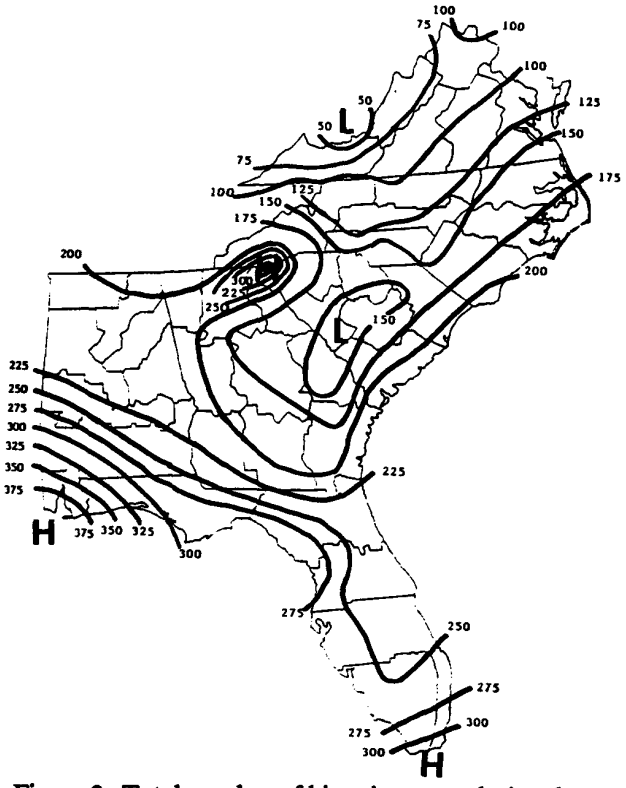


Figure 2. Total number of big rain events during the 60-year period.

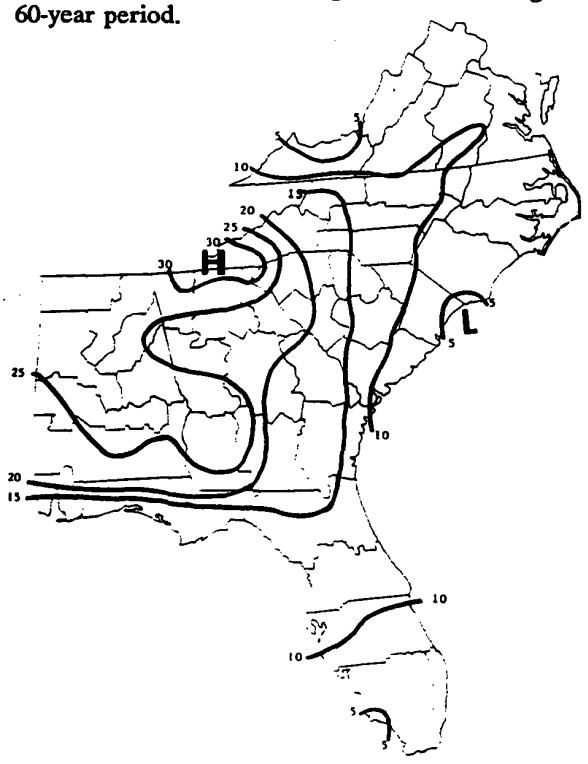


Figure 3a. The winter (Dec.-Feb.) percentage of big rain events during the 60-year period.

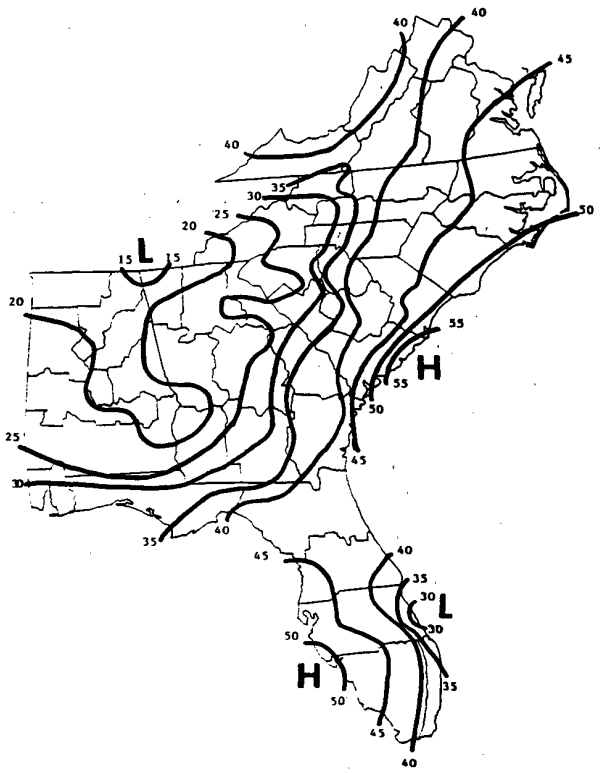


Figure 3b. The summer (Jun.-Aug.) percentage of big rain events during the 60-year period.

The seasonal differences in spatial distribution of big rain events identifies three heavy precipitation regimes, that exist outside of Florida, in 1) the Appalachian mountains, 2) the coastal zone (inland from the coast line 60 km), and 3) areas between these two regions (Changnon, 1992). When maximums occur along the coast minimums occur in the mountains (summer) and the opposite is true in winter. There is almost no mountain-to-coast gradient in the number of big rain events during spring or autumn, however maximums are generally located in the region between the mountains and the coast during these seasons.

The temporal results, determined from the ratio of the five-year, 24-hour storm amounts from the 1960-1989 period over the 1930-1959 period, identified patterns of increase and decrease in rainfall intensity across the region. Figure 4 shows the ratios based on the 60-year period of record. Regions that exhibited a decrease of 10 percent or more included southern Florida, southern parts of Alabama and Georgia, and areas from central South Carolina north to Virginia. Western Alabama, northern Georgia, eastern parts of South Carolina and southeastern Virginia exhibited increases of 10 percent or more. The small pockets of significant change may be related to meso-scale or localized effects; however, all broad areas of increase or decrease are justified by at least three stations having similar anomalies. Based on the Huff and Angel (1990) results in the Midwest, the increases in northern Georgia may be explained by the increased urban influence of Atlanta.

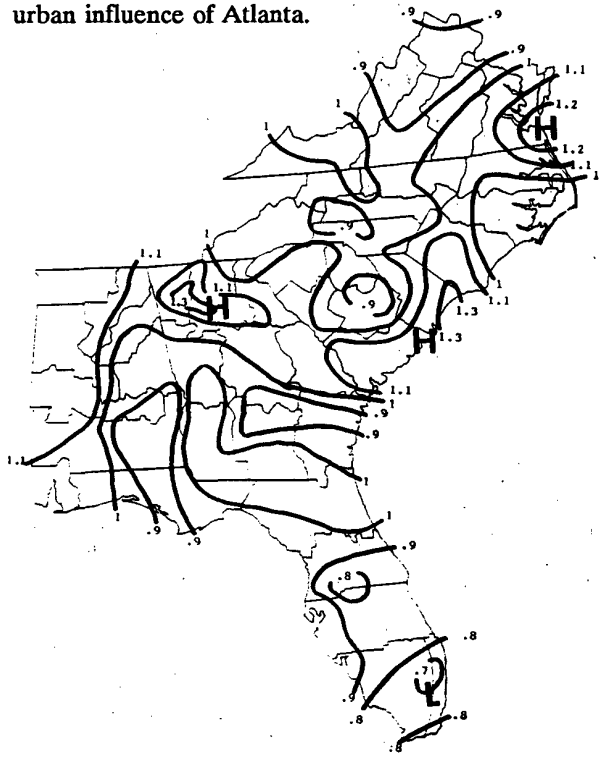


Figure 4. The ratio for the two 30-year periods for the five-year, 24-hour storm (annual).

To further understand the anomalies on the annual pattern of 24-hour, five-year storm ratios, seasonal 24-hour, five-year storm ratios were determined. The summer pattern (see Figure 5a) of ratios identified significant maximums, with increases greater than 30 percent in areas around Atlanta and along the South Carolina coast. Figure 5b shows the autumn ratios. This map shows 20 to 30 percent decreases across Florida and southern parts of Alabama and Georgia. These decreases may be related to findings in a recent study (Gray, 1990) that indicated a sharp decrease in the number of tropical storms and hurricanes since 1970. These changes in seasonal frequency of heavy rainstorms may indicate changes in large-scale atmospheric circulations. The spatial patterns for different seasons are closely related to the annual patterns shown in Figure 4 and indicate that seasonal information must be incorporated into updating rainfall frequencies.

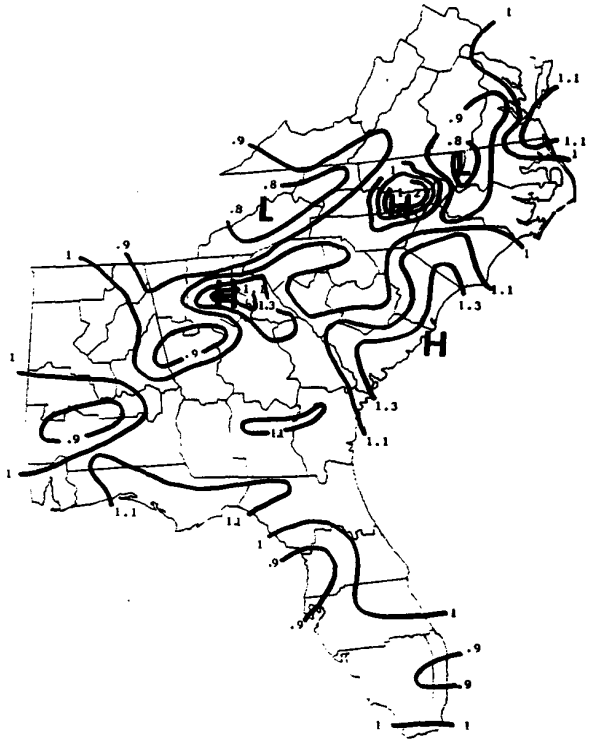


Figure 5a. The ratio for the two 30-year periods for the five-year, 24-hour storm during summer.

Figure 5b. The ratio for the two 30-year periods for the five-year, 24-hour storm during autumn.

4. CONCLUSIONS

This study analyzes the spatial and temporal changes in large precipitation events annually and seasonally at 89 NWS stations across the Southeast U.S. Analysis of daily events >50 mm by season indicates that three heavy precipitation regimes exist across the region outside of Florida: 1) mountains, 2) coastal zone, and 3) in between. The significant seasonal differences in big rain events can have implications for design and management of water related structures. The temporal analyses indicated broad significant increases and decreases across the region. Further research using synoptic analyses may help explain the differences identified. These results, based on climatic analyses, indicate that seasonal differences in heavy rainstorms must be understood to help develop updated rainfall frequency values across the region.

5. REFERENCES

Changnon, D., 1992: Regional variations in large daily precipitation amounts: Information for planners in a changing climate. Submitted to Int. J. of Climatology for review summer 1992.

Gray, W.M., 1990: Strong association between West African rainfall and U.S. landfall of intense hurricanes. Science, 249, pp. 1251-1256.

Hershfield, D.M., 1961: Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years. Technical Paper No. 40, Weather Bureau, U.S. Dept. of Commerce, Washington, D.C., 115 pp.

Huff, F.A. and J. Angel, 1990: Fluctuations in the frequency distributions of heavy rainstorms in the Midwest. Preprint from Symposium on Global Change Systems and the Special Sessions on Climate Variations and Hydrology, Feb. 5-9, 1990, Anaheim, CA, pp. 189-192.

Huff, F.A. and S. A. Changnon, 1987: Temporal changes in design rainfall frequencies in Illinois. Climatic Change, 10, 195-200.

Robinson, P.J. and S.J. Walsh, 1991: Precipitation regime changes associated with climatic changes. Department of Geography, University of North Carolina, Chapel Hill, NC, 124 pp.

TEMPORAL VARIABILITY OF FLOOD AND HEAVY PRECIPITATION EVENTS IN THE MIDWEST

R.T. Shealy, K.E. Kunkel, and S.A. Changnon
Office of Applied Climatology
Illinois State Water Survey
Champaign, Illinois

1. INTRODUCTION

We investigated temporal changes in the frequencies of floods and heavy precipitation events in the Midwest for the period 1921-1985. This period was chosen because most streamgaging records begin around 1920 (most precipitation records begin earlier, so the streamgage records were the limiting factor). The area of analysis encompassed the nine-state region of Illinois, Indiana, Iowa, Kentucky, Michigan, Minnesota, Missouri, Ohio, and Wisconsin. This area has a relatively homogeneous climate with similar causal mechanisms for heavy precipitation events, and by implication, for flooding events.

Flood and heavy precipitation events were defined by means of a partial duration series of a one-year recurrence of daily streamflow and precipitation values (U.S. Water Resources Council, 1981). A partial duration series consists of the largest events in the time series in decreasing order of magnitude. We tested event occurrence times for temporal uniformity; any significant departure from a uniform distribution would suggest that climatic changes have occurred over the period of record.

If a partial duration series is temporally homogeneous, and the total number of events in the period of record is fixed, then it is a Poisson process with constant intensity function equal to the inverse of the average recurrence (Karlin and Taylor, 1975). Therefore, testing for temporal homogeneity in event location is analogous to testing the null hypothesis of constant event intensity. Chang (1987) analyzed sets of event locations by estimating the intensity function, using an estimator of Cox and Lewis (1966). In this paper, we employ statistical tests specifically designed to detect various types of nonconstant intensity functions.

2. METHODS

Streamgaging data were obtained from the EarthInfo Hydrodata CD-ROM. Stations were selected based on the criteria that (1) there were no significant control structures upstream of the station, and (2) less than 25% of the data was missing over the 65-year period. A few additional stations were chosen for areas that were underrepresented. A total of 79 stations were selected, and their locations are indicated on Figure 1. Partial duration series for flood events were constructed as in Buishand (1989).

Precipitation station data were obtained from a 1500-station database on the Midwestern Climate Information System (Kunkel et al., 1990). Stations were chosen that included a 1921-1985 period and had less than 20% missing data for this period, with 6 additional stations added to fill sparse areas. A total of 242 stations were selected (see Figure .1). Precipitation partial duration series were computed in the same manner. Several event lengths (1-, 3-, 5-, 7-, and 10-day) were investigated. It was found that the longer event lengths corresponded more closely with flood occurrences. Subsequently, an event length of 7 days was chosen for analysis (see Kunkel et al., 1992).

Based on the nature of midwestern flood events, we chose two seasonal periods of interest for separate analysis: a cold season (December 1-April 30) when snowmelt can contribute and a warm season (May 1-November 30) when snowmelt is not important. The present study focuses on the warm season. The partial duration series was obtained,

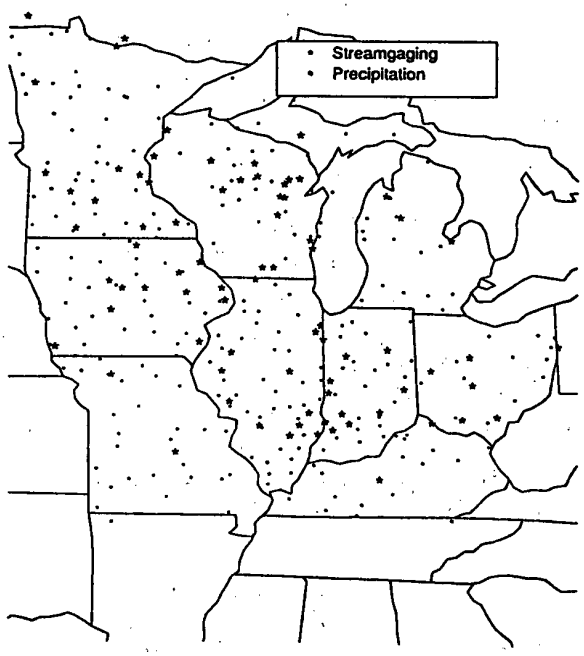


Figure 1. Locations of streamgaging and precipitation stations used in the study.

then, by restricting attention to the daily record in the warm season only; in reality, we are deriving a series with a recurrence interval of 1 season.

Each partial duration series was examined for nonuniformity of event occurrence over the period of record; if it is uniform, then the proportional event times (event times divided by total record length) are order statistics from a Uniform (0,1) distribution. Consequently, a national test to use is the Kolmogorov-Smirnov (K-S) test for goodness of fit to a uniform distribution:

$$D = \max_{1<i<=n}[\max[i/n - T_i/T_0, T_i/T_0 - (i-1)/n]]$$

where i is the event number, n is the total number of events, T_i is the event time in days relative to the start of the record, and T_o is the record length in days. Quantiles of the null distribution of D are derived using an algorithm of Kolmogorov (Birnbaum, 1952).

Once the overall homogeneity had been assessed it was desired to test for a trend in event intensity. Two tests for general increasing trend were used. The first test is based on the statistic

$$Z = 2 \sum_{i=1}^{n} \ln(T_o/T_i)$$

which has approximately a chi-squared distribution with 2n degrees of freedom if the $\{T_i\}$ are uniformly distributed. The hypothesis is rejected in favor of an increasing trend in relative frequency for small values of Z. In a power study by Bain et al. (1985), Z was most powerful among all considered tests in detecting a smoothly increasing (including a linear) trend. The second test was employed to detect rapid jumps in intensity; it uses large values of Laplace's statistic

$$L = \sum_{i=1}^{n} T_i/T_o$$

to reject in favor of such alternatives. L is approximately Normal (n/2, n/12). Again, Bain et al. (1985) showed that a test based on L was most powerful in detecting jumps in relative frequency among the studied tests. To test for decreasing trend, the tests were reversed; rejection occurs for large values of Z and small values of L.

As no events are being considered from December to April, these months are excluded from the time line; otherwise, the eventless December-April gaps will cause the event series to exhibit nonuniformity because of these gaps, confounding the nonuniformity due to possible climate fluctuations. In addition, if there is missing data in a record (in the warm season), this portion is also excised from the record.

3. RESULTS

It was found that 32% of streamgaging stations have nonuniform flood intensity, using the K-S test at the 10% level. The level is set at 10% because the K-S test is inherently a 2-sided test, testing for both positive and negative trends simultaneously, as well as nontrend fluctuations. In contrast, only 10% of the precipitation stations showed nonuniformity at the 10% level.

When the two trend tests are applied, an interesting pattern emerges. Referring to Table 1, it is seen that there is no evidence of a decreasing trend in any of the streamgaging or precipitation station records. However, both tests detected substantial percentages of increasing trends in both types of records, well above the nominal 5% level. In fact, the percentages detected were virtually the same for the two tests, and roughly equivalent to the 2-sided 10% level percentages that the K-S test detected.

Table 1. Percentage of stations significant at the 5% level, using the two trend tests.

Trend Tests		
Trend	Z	L
Increasing flood intensity	39.2	35.4
Decreasing flood intensity	0	0
Increasing precipitation intensity	18.6	13.6
Decreasing precipitation intensity	0	0

Contour maps of the p-values obtained by the K-S test and the Z-test are given in Figure 2. The spatial extent of the streamgaging stations with increasing intensity using the Z-test (Fig. 2b) is nearly identical to the region of general nonuniformity found by the K-S test (Fig. 2a): namely, central Minnesota, western Iowa, and southern Lake Michigan.

However, the extent of precipitation event inhomogeneity was not as clear. There is very little spatial coherence among the stations detected with event inhomogeneity with the K-S test (Fig. 3a; contour 0.1) or those detected with increasing intensity using the Z-test (Fig. 3b; contour 0.05). However, although such regions are fragmented, there is some compatibility between the two tests, indicating again that the majority of the inhomogeneity detected is in the form of increasing intensity over the period of record. If the significance levels are doubled for each test (0.2 for the K-S test, 0.1 for the Z-test), then more spatial coherence is evident; however, at best, the "signal" of increasing intensity is much weaker than that in the flood occurrence patterns. The regions labeled by the tests include upper Minnesota, western Iowa, Kentucky, eastern Ohio, and Lake Michigan.

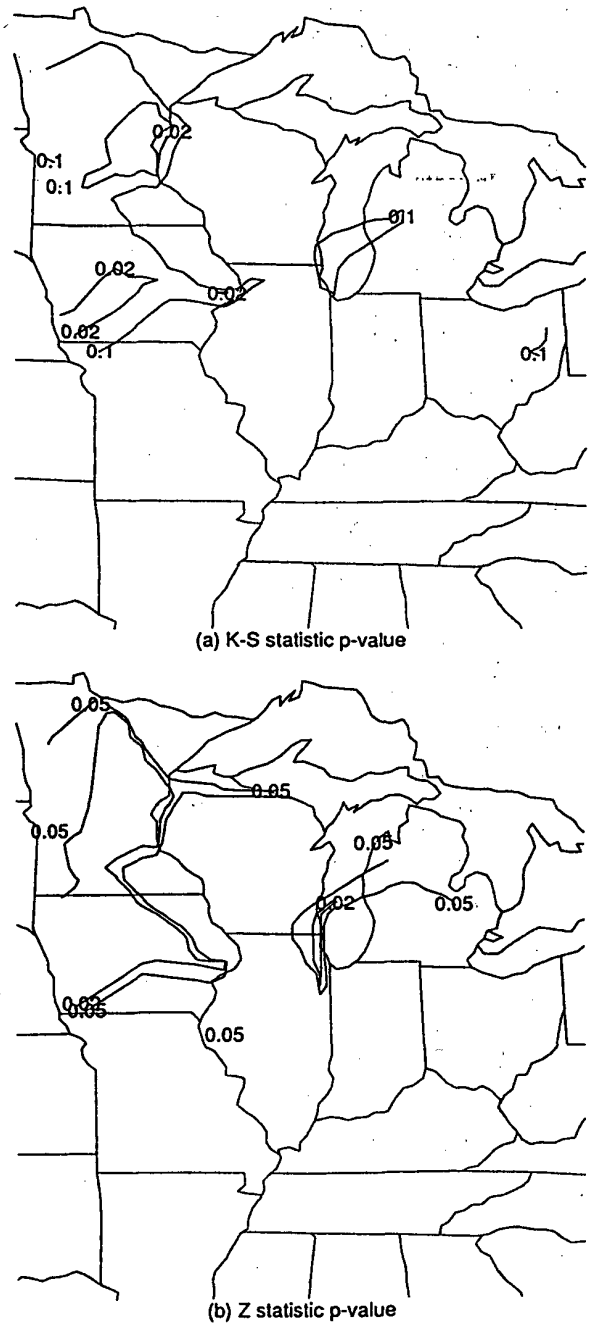


Figure 2. Contours of the p-values for the K-S test (a) and Z-test (b) for inhomogeneous intensity in flooding events in the warm season. The maximum contour in (a) is .10, twice the maximum contour in (b), since the K-S test is a 2-sided test of trend.

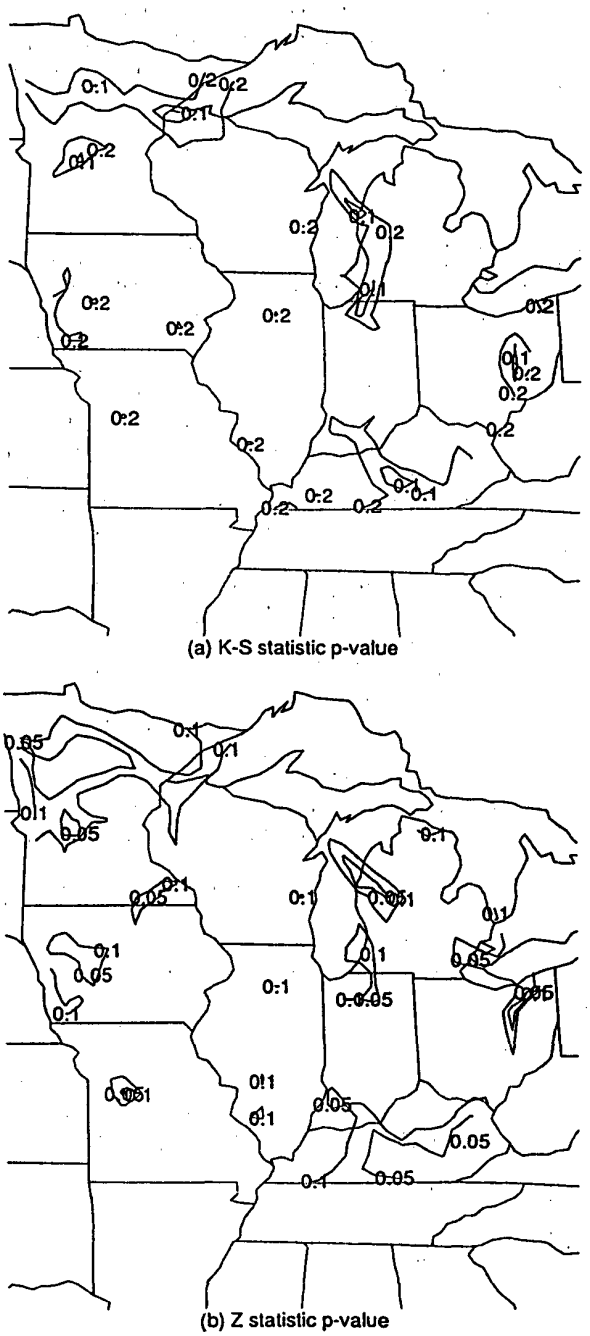


Figure 3. Contours of the p-values for the K-S test (a) and Z-test (b) for inhomogeneous intensity in precipitation events in the warm season. Since the precipitation intensity is more constant temporally, a higher maximum contour than in Figure 2 (.2 in (a); .1 in (b)) is used to detect a weaker nonuniform intensity.

Comparing Figures 2b and 3b, it can be seen that only 3 areas emerge with upward trends in both flood and precipitation event frequencies: central to upper Minnesota, a small part of western Iowa, and the area near the southern extent of Lake Michigan.

4. DISCUSSION

Based on the results, it is possible that the increases in precipitation and flood intensities concurrently observed in the northwestern region of the study area indicate a linkage of increased flooding conditions in response to increased frequency of heavy precipitation events. However, the signal of increasing precipitation intensity is weak at best, suggesting two possibilities:

1. Since runoff is non-linearly related to precipitation, a small increase in precipitation intensity and/or magnitude could result in a larger increase in runoff and streamflow.

2. Increasing flooding frequencies may be only partly the result of an increase in precipitation event intensity. Another source is human modification of the drainage basin. Although these basins were carefully chosen based on advice of state experts regarding minimal human influence, there has undoubtedly been at least some modification or land use changes in every basin and runoff may be more sensitive than believed to these changes.

The indications are clearer in the remainder of the region: no tendency of increasing or decreasing trend in either flood or precipitation event frequencies, over the entire 65-year period of record, is evident.

ACKNOWLEDGEMENTS

This work was partially supported by the U.S. Geological Survey under award number 14-08-0001-G1731. We thank Jim Zandlo, Fred Nurnberger, Kenneth Scheeringa, and Pam Naber Knox for providing precipitation data. We also thank Melvin Lew, Richard Novitzky, and the USGS district offices in Illinois (Wayne Curtis), Indiana (Ronald Thompson), Iowa (Phil Soenksen), Michigan (Stephen Blumer), Minnesota (George Carlson), Ohio (Harold Shindel), and Wisconsin (William Krug and Dianne Maertz) for providing information on streamflow stations. We thank Arthur Scott and Ken Wahl for their assistance in our interactions with the USGS District Offices. Finally, we thank Jean Dennison for her professional preparation of the manuscript.

5. REFERENCES

Bain, L., Engelhardt, M., and F.T. Wright, 1985: Tests for an increasing trend in the intensity of a Poisson process: A power study. J. Amer. Statist. Assoc., **80**, 419-422.

Birnbaum, Z., 1952: Numerical tabulation of the distribution of Kolmogorov's statistic for finite sample size. J. Amer. Statist. Assoc., **47**, 425-441.

Buishand, T.A., 1989: The partial duration series method with a fixed number of peaks. J. Hydrology, **109**, 1-9.

Chang, T.J., 1987: Analysis of 3-Component Floods in the Ohio River Basin. In: Hydrologic Frequency Modeling, V.P. Singh (Ed.), 583-594.

Cox, D.R., and R.A.W. Lewis, 1966: The Statistical Analysis of Series of Events. Methuen: London.

Karlin S., and H. Taylor, 1975: A First Course in Stochastic Processes. Academic Press: New York.

Kunkel, K.E., S.A. Changnon, C.G. Lonnquist, and J.R. An gel, 1990: A real-time climate information system for the midwestern United States. Bull. Amer. Meteor. Soc., 71, 1601-1609.

Kunkel, K.E., S.A. Changnon, and R.T. Shealy, 1992: Applications of Statistical Methods to the Study of Climate and Flooding Fluctuations in the Central United States. Final Report, USGS Grant 14-08-0001-G1731, Illinois State Water Survey.

U.S. Water Resources Council, 1981: Guidelines for Determining Flood Flow Frequency. Bulletin #17B, 2120 L Street, N.W., Washington, D.C.

Regional Flood Frequency Study for Kerala Region in India

E.J. JAMES
Centre for Water Resources Development and Management, Kozhikode, India

G. RANGANNA & M.R. MOHAN
Karnataka Regional Engineering College, Surathkal, India

ABSTRACT

A regional flood frequency study was carried out for the Kerala region in India; Gumbel's extreme value distribution was applied, considering its advantages. Attempts were made to establish a satisfactory relationship between annual mean peak flood and basin characteristics, such as area, stream frequency and slope.

1. Introduction

The Kerala region with about 40000 sq.km. area has 41 west-flowing and 3 east-flowing rivers with a minimum length of 15 km and a maximum length of about 250 km. These short rivers are monsoon-fed and flow through a steep topography; the average width of the region between the Western Ghats and the Arabian Sea is only 70 km (Fig.1).

The regional analysis attempted overcomes the sampling error and its essential feature is the grouping of data to improve the quality of the result. An attempt was also made to establish a relationship between annual mean peak flood and basin characteristics for predictive purposes at ungauged sites.

2. Data Base

The peak flood data available from 85 stations have been considered. The areas covered by gauging stations vary from 28 sq. km. to 4000 sq.km. These stations are maintained by agencies like Public Works Department and Kerala State Electricity Board.

Stations yielding more than seven years data were only considered; data showing discrepancy when compared to the rainfall data or the discharge of neighbouring stations were avoided; data from those stations influenced by impoundments were not included. Most of the stream gauging stations use stage-discharge method; some sites have well defined cross-sections and others have weirs or other artificial structures. The stations are calibrated using current meters and rating curves have been established.

3. Regional Flood Frequency Analysis

The methodology followed is based on the Flood Studies Report (NERC 1975). The different steps involved in such an analysis are given below:

(i) Assembling annual peak flood data, Q for all the stations in the region;

(ii) Ranking the data of each station in ascending order;

(iii) Estimating the mean of each stations' data set as the arithmetic average of the annual maxima, the exception to this being when the highest ranking flood exceeds three times the median;

(iv) Dividing Q of each station data by $\bar{Q}$ (mean annual peak flood)

to get the $Q/\bar{Q}$ values (non-dimensional);

(v) Using the extreme value distribution of Gumbel (1942), the corresponding plotting position Y (reduced variate of the return period) for the $Q/\bar{Q}$ values are obtained for which Gringorten's (1963) approximation given below is used:

$$F(i,N)=(i-0.44)/(N+0.12)$$
$$Y(i,N)= -\log_e(-\log_e F)$$

where i=1 to i=N, N being the number of years of data.

(vi) Merging and assembling the entire Y and $Q/\bar{Q}$ values into class intervals (Y being divided into intervals of width 0.5) and identifying and assembling corresponding values of $Q/\bar{Q}$ for intervals of Y;

(vii) Finding out the aveage Y and $Q/\bar{Q}$ values in each interval from the coordinates of the regional frequency curve and dividing the stations in the region into several groups with each group having around 100 station-years of data;

(viii) Extracting the highest values of $Q/\bar{Q}$ regardless of the station and associating these with the plotting position for i=M-3,M-2, M-1 and M, where M is the total number of station-years of data;

(ix) Placing the extracted highest values in class intervals which will facilitate extension of the curve for higher return periods;

(x) Plotting the coordinates obtained in steps (vii) and (ix) and drawing the eye-guided best fitting curve;

(xi) Arriving at the return period from the relationship:

$$Y_T = -\log_e \log_e \frac{T}{T-1}$$

where Y_T is the reduced variate for the return period T.

When the highest ranking flood exceeds three times the median, it indicates an outlier and the arithmetic average is raised and the $Q/\bar{Q}$ depressed. Under this condition an estimator, K times the median flood, is used as $\bar{Q}$. The adjustment factor, K, is obtained empirically by arriving at a relation between median and arithmetic average by plotting a curve, median versus arithmetic averages of all stations in the region with $Q_{max}/Q_{med} \leq 3$; this avoids biasing the estimate of Q when sample includes outliers.

The entire area of study was divided into 13 flood frequency zones considering the nearness of basins, their sizes and data availability. The curve obtained for central zone is given in Fig.2. Since the entire area studied seemed to be more or less homogeneous from the point of view of the frequency study, a common curve for the entire Kerala region was evolved (Fig.3).

4. Relationship between Mean Annual Peak Flood and Basin Characteristics

The prediction of flood frequency for ungauged basins mainly relies on statistical relationship between flow and certain characteristics of sample basins. These characteristics may broadly be classified based on morphometric, cover and climatic properties. Since the climatic properties in the study area are homogeneous and since sufficient data on cover properties are not available, only morphometric aspects have been considered. The major properties supposed to have an influence on the peak flood, namely basin area, stream frequency and stream slope have been used.

The relationships for the three physiographic zones of Kerala namely, the eastern highland, western lowland and the midland in between the two, as well as for the east-flowing river zone (Fig.1) are given below:

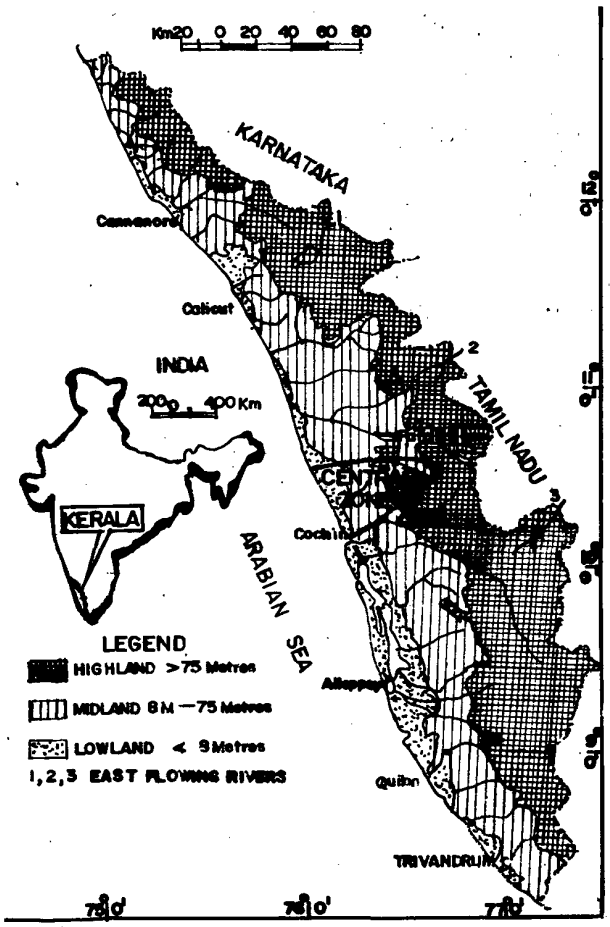


FIG.1. Kerala- physiography and drainage

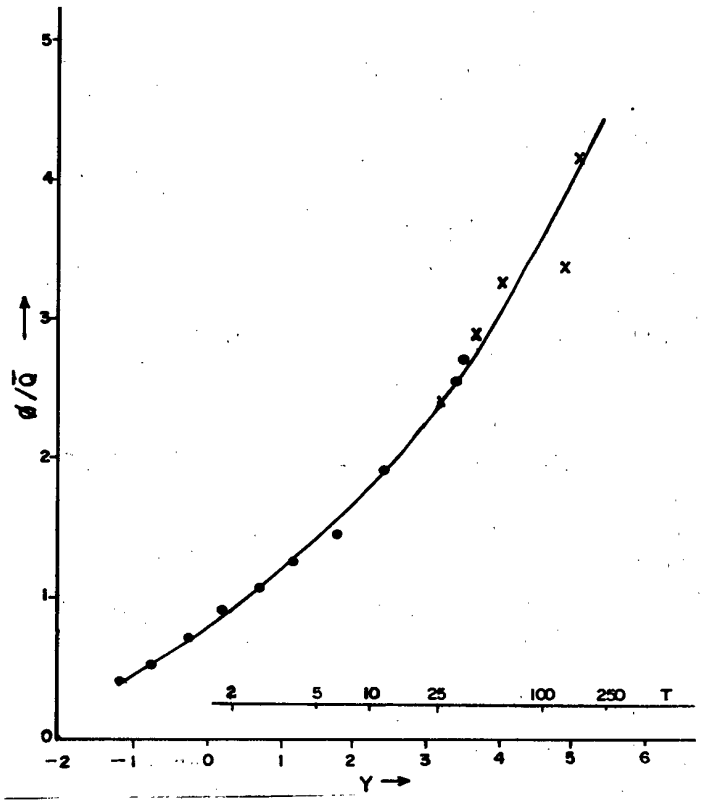


FIG.2. Flood frequency curve - central zone

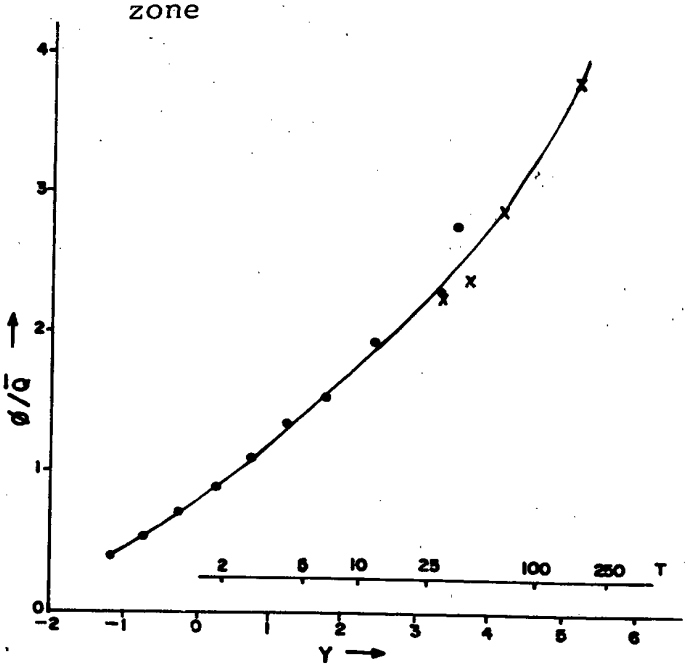


FIG.3. Flood frequency curve - Kerala region

For highland:

$\bar{Q}$ = 184.77+0.51 AREA-20.39 STMFRQ 0.42 S 1085 (R=0.82)

$\bar{Q}$ = 192.99 +0.50 AREA-20.09 STMFRQ (R=0.82)

$\bar{Q}$= 126.05+0.51 AREA +0.28 S1085 (R=0.82)

For midland:

$\bar{Q}$ = 260.32+0.63 AREA -15.75 STMFRQ +0.024 S1085 (R=0.93)

$\bar{Q}$ = 235.26+0.63 AREA -0.47 S1085 (R=0.93)

For lowland;

$\bar{Q}$ = 2.19 +0.58 AREA -60.94 STMFRQ + 11.91 S 1085 (R=0.98)

$\bar{Q}$ = 132.53+0.47 AREA-49.22 STMFRQ (R=0.91)

$\bar{Q}$ = 126.42+0.58 AREA +10.32 S1085 (R=0.98)

For east-flowing rivers:

$\bar{Q}$ = -183.91+0.32 AREA+154.32 STMFRQ -7.39 S 1085 (R=0.89)

$\bar{Q}$ = 32.55 +0.34 AREA +49.89 STMFRQ (R=0.66)

$\bar{Q}$ = 102.19+0.36 AREA +0.60 S1085 (R=0.46)

where Q is the mean annual peak flood, AREA the area of the basin, STMFRQ the stream frequency, S1085 the stream slope (between 10 and 85 percentiles of the mainstream length measured upstream from the site) and R the coefficient of correlation.

5. Conclusions

The flood frequency curve prepared for the entire Kerala region possesses sufficient accuracy for the purposes of user departments and agencies.

Significant correlations were obtained for all the equations of the three physiographic zones. In the case of the zone of east-flowing rivers, the results obtained using all the three characteristics give better results than those obtained using area and one more characteristic. The correlation obtained with area alone was not satisfactory.

REFERENCES

Gringorten, I.I., 1963: A plotting rule for extreme probability paper, J. Geophys. Res., vol.68, no.3, 813-814.

Gumbel, E.J., 1942: Statistical control-curves for flood discharges, Trans.,Am.Geophys. Union, vol.23, 489-500.

NERC, 1975: Flood studies report, National Environment Research Council, U.K., 5 vols.

INHOMOGENEITIES IN TIME SERIES OF EXTREME RAINFALL

W.D. Hogg

Canadian Climate Centre - AES

4905 Dufferin Street

Downsview, Ontario, Canada, L3Y 5N7

1. INTRODUCTION

Millions of dollars are spent each year in Canada, on the design and construction of structures to carry water runoff from small catchments. In the absence of adequate streamflow information, the design engineer frequently turns to rainfall data to aid in the synthesis of peak flows. Inherent in this approach is the need to estimate the frequency with which the capacity of the structure will be exceeded. The reliable estimation of the probability of extreme rainfall events is an important climatological service.

Regularly recorded climatological measurements have been made in Canada for more than 150 years and are archived for about 2800 locations at present. However, there are only about 50 stations across the country with more than 100 years of data. The majority of these stations use the AES Type B standard rain gauge. The Type B is essentially a can with a funnel and interior clear plastic liner in which the height of accumulated rainfall is manually read once a day. But information on the probability of short, intense bursts of rainfall is often critical for design purposes. Data from a much smaller (500 gauges) network of tipping-bucket recording rain gauges are used for this purpose. The majority of these gauges have only been in operation for 20-30 years. Thus, there is a requirement to estimate short duration rainfall with recurrence intervals as long as 100 years, usually based upon records of less than 30 years. (The 100 year return period amount is the amount expected to be equalled or exceeded, on the average, once every 100 years.)

Recent work on long time series of annual extremes of daily rain (Hogg, 1991) provided an excellent opportunity to evaluate estimates of long recurrence interval statistics based on the much shorter recording rain gauge record lengths. Samples from the extended daily extreme data were used to estimate 100 year return period rainfall amounts. Variability in these estimates is a measure of the degree of homogeneity of the annual maximum rainfall series. Samples were selected by choosing contiguous years of data of varying sample size, to better simulate the start-up and closing of recording gauge stations. The following discusses some of the findings.

2. AES FREQUENCY ANALYSIS

AES currently uses extreme value statistics procedures to estimate the frequency of rare rainfall events. Since the frequency of the heavier rainfalls is the main item of interest, procedures which permit reliable interpolation and extrapolation of the probability of infrequent events, are desirable. Historically, annual maxima of rainfall in various duration categories were selected for analysis, to help focus on extremes and for ease of data handling. As described in Hogg and Carr (1985), AES assumes that the double exponential Extreme Value Type I Distribution (Gumbel) fits extreme rainfall, for all individual stations analysed. The Gumbel is a two parameter distribution. There is a host of three and higher parameter distributions, that have been proposed as alternatives to Gumbel. As might be expected, distributions with more parameters determined by the sample, often fit the sample better than Gumbel. Whether they produce better estimates of the long term probability of extreme rainfall remains a topic of debate.

3. DATA ANALYSIS

Time series of annual extremes of daily rainfall were analysed in an earlier study (Hogg, 1991). Stations with long sequences of once a day manual

observations of rainfall were selected for analysis in an attempt to identify trends or cycles in the time series. It was hoped that time series of extremes might be less susceptible to changes in observing programs than would time series of monthly or annual accumulations of rainfall. There was some evidence of cycles for some west coast stations, but trend evidence was inconclusive. However, there was ample evidence that standard estimates of the probability of extreme rainfall events associated with very large uncertainties.

Time series of annual extremes of daily rainfall were compiled for 13 stations across Canada. The results for 8 of these stations were shown in Hogg (1991). These stations were not subjected to rigorous homogeneity testing but were the survivors of a larger selection of long period stations whose annual accumulation time series were scrutinized for step changes, and whose station history files were searched for major moves, to eliminate unsuitable candidates.

In addition to annual extremes and 10-year running means, 100 year return period estimates were made using sample windows of data. This return period amount is directly dependent on both the mean of the extremes and their year to year variability. AES fits the Gumbel distribution to the annual extreme series using either a method of moments or method of maximum likelihood fitting procedure. For Hogg (1991), time series of the 100 year return period amounts (R100) were produced using a moving window of the previous 20 years of data and, separately. For the results shown here, moving windows of 30 and 40 years duration are added.

4. RESULTS

For the two stations shown (Fig. 1), there are two sets of points. The set with lower values contains points related to the annual extreme daily rainfall as measured at that station. The bars represent the extreme daily rain for each year. A second set of points depicts the 10-year running mean of the extremes to smooth the data. The set with higher values contains points pertaining to the 100-year return period estimate of the daily extreme rainfall (R100). Separate curves show R100 as determined from the previous twenty years, thirty years and forty years and. The straight line depicts R100 as calculated using all available data.

The one thing obvious from the R100 time series for both stations is that 20 years of data is woefully inadequate to estimate the 100-year return period daily rainfall value. Estimates vary by 30-40% according to the specific 20-year period used as a data sample. Examination of the R100 curves based on longer samples suggests that even 40 years of data are not enough to obtain a reasonably stable estimate of the 100-year return period daily rainfall. In fact, Hogg (1991) showed that 20 years of data are not even enough to generate stable estimates of the 10-year return period values.

This is not a question of poor fit to the Gumbel distribution but is a sampling problem. Figures 3 and 4 adequately illustrate this. Figure 3 is a plot of the full 150 years of data for Toronto on a double exponential (Gumbel) graph using the Gringorten plotting position. The straight line represents the method of moments best fit to the sample. The fit is good. Smaller samples from the Toronto data are fit equally well by the Gumbel. Figure 4 shows R100 estimates based on twenty year wide windows of data for Agassiz, assuming different distributions and fitting techniques. For Agassiz, there is little difference between the Method of Moments versus Method of Maximum Likelihood fits of the Gumbel Distribution. The R100 estimates assuming that the 20-year samples come from a General Extreme Value Distribution, show slightly greater variability than the Gumbel estimates. The difficulty of using what might appear to be reasonable historical records to predict the frequency of extreme rainfall events, is well illustrated.

5. DISCUSSION

Currently used extreme value analysis techniques produce very unstable estimates of the frequency of heavy rainfall events. The problem seems to be a sampling one. Different sub-samples produce very different estimates of rainfall frequency. This suggests that assumptions of homogeneity and stationarity, inherent in the extreme value analysis techniques, are not valid for annual maximum series of rainfall.

Other approaches to the problem have been suggested which may, at least partially, overcome some of these difficulties. Zwiers and Ross (1991) suggest a technique which increases the sample size by including all rainfall observations throughout the year, after suitably removing seasonal cycles. Pilon and Adamowski (1992) make a case for the value of including regional information using the method of L-Moments. Others have tried to use weather radar data to make use of this instruments increased information on the spatial characteristics and probabilities of extreme rainfall events. Although rainfall in an

individual event is very discontinuous, intuitively at least, the expectation is that extreme rainfall events will be similar in frequency and characteristics, over fairly large regions. This would seem to offer hope that some sort of regional approach might improve the frequency analysis of rainfall.

Even these improvements may not reduce uncertainties to acceptable levels, however. Time series of extreme rainfall suggest that there is memory in the climate system for periods much longer than a year. There are indications of 20-30 year cycles, and trends, perhaps associated with climate change, will have to be dealt with. Now is the time to evaluate and adopt improved procedures for estimating the frequency of occurrence of extreme rainfall events.

5. REFERENCES

Hogg, W.D., 1991. Time Series of Daily Rainfall Extremes. Preprints, Fifth Conf. on Climate Variations, 86-89 AMS Boston Ma.

Hogg, W.D. and D.A. Carr, 1985. Rainfall Frequency Atlas For Canada. 89 pp. AES Downsview.

Pilon, P.J. and K. Adamowski, 1992. The Value of Regional Information to Flood Frequency Analysis using the Method of L-Moments. Cdn. J. Civ. Eng., Feb. 1992.

Zwiers, F.W. and W.H. Ross, 1991. An Alternative Approach to the Extreme Value Analysis of Rainfall Data. Atmosphere-Ocean **29** : 437-461.

FIGURE 1. Time series of annual maximum of daily rainfall for Agassiz B.C. (near Vancouver). The bars and short horizontal ines represent the annual extremes and a 10-year running mean of annula extremes. The straight line depicts the 100 year return period rainfall as estimated from the entire period of record. The points plotted about this line represent the 100 year return period rainfall as estimate using the 20, 30 or 40 years of data immediately prior to the plotted point.

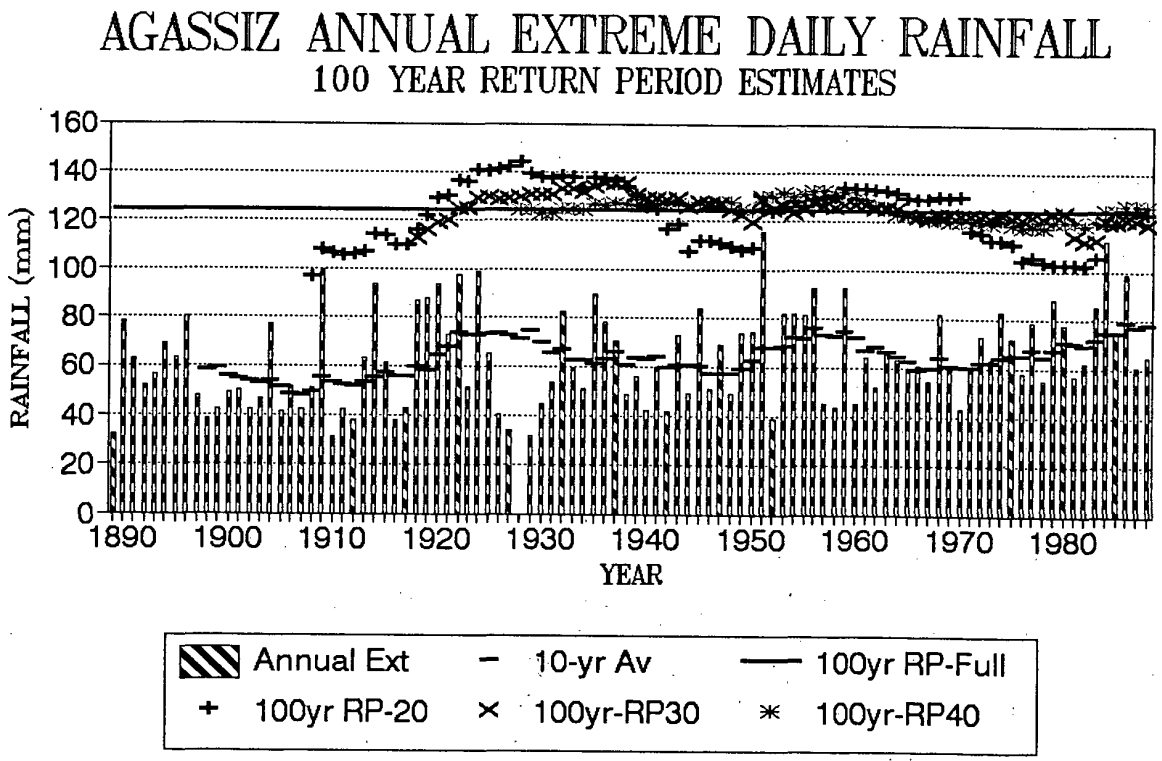


FIGURE 2. Same as Figure 1 but for Toronto, Ontario.

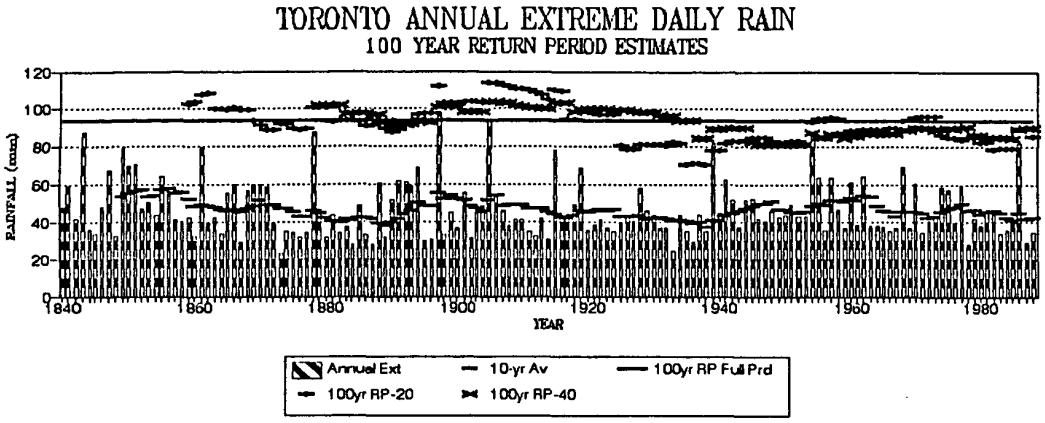


FIGURE 3. Gumbel (EV1) plot of annual *extreme daily rainfall for Toronto* using method of moments fitting.

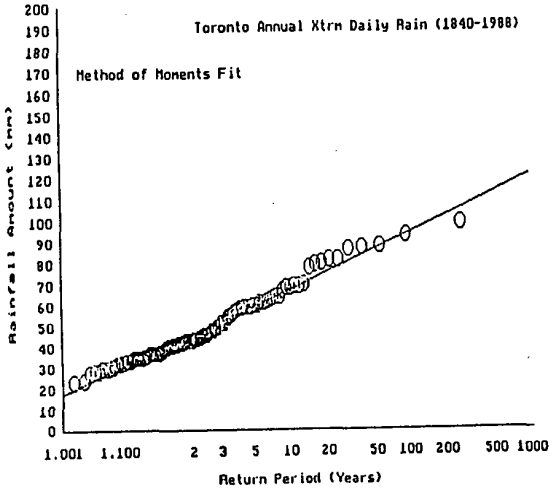
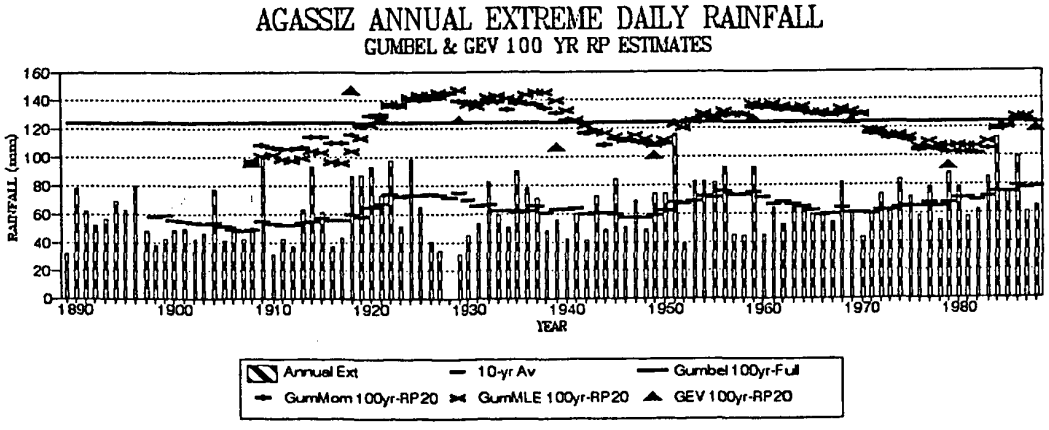


FIGURE 4. Same as Figure 1 except that 100 year return period rainfall estimates are all based upon 20 years of data fit. The + and X depict points fit to the Gumbel distribution using the method of moments and method of maximum likelihood respectively while the triangle show the same estimates assuming a General Extreme Value (GEV) distribution.



Does an internal time structure of extreme summers in Central Europe exist?

Friedrich- Wilhelm Gerstengarbe
Potsdam-Institut für Klimafolgenforschung
Telegrafenberg, O-1561 Potsdam, Germany

The external time structure

1. Introduction

The complexity of nearly all meteorological events demands appropriate complex methods for their analysis. This means a general withdrawal from the ususal linear investigation of single values and a change towards the investigation and description of interaction between a great number of quantities. The use of multivariate statistical methods seems to be a practicable way to describe meteorological events. In this way a climatological event can be described for a special time period by a fixed number of parameters. Within the present investigation a method has been developed to define meteorological extremes and to describe their temporal behaviour. The procedure is presented by the analysis of the occurence of hot and cold summers in Central Europe.

2. Data

Daily values are used for all investigations of European long-time data sets (1901-1980). The meteorological parameters used for the definition of hot and cold summers are the daily maximum and the daily mean of the air temperature. Within a preliminary examination it could be proved that only these parameters supply an important contribution to the description of extreme summers. All data have been investigated for homogeneity by various methods. It is one of the findings that some of the data series show non-homogeneities within the period from 1940 to 1945 caused by a lack of measurements. These non-homogeneities have been taken into consideration for the following investigations and interpretations.

3. Selection of the parameters

In general every meteorological event must be described not only by one parameter. Such an event is the result of a large number of parameters which are in mutual interaction. A sensible definition and selection of a defined number of these parameters contributes essentially to the quality of the description of the events. Therefore, it is appropriate to start with a large number of parameters in order to select essential contributors to the description of the events. To estimate the extreme events "hot" respectively "cold" summer, seven parameters have been selected on the base of the meteorological time series mentioned above (see Table 2). The period May-September was chosen to cover all situations discribing these events. The description of the mean conditions was limited to the months June-August to prevent possible falsification of the results e.g. by cool month of May or September. It is evident that the selected parameters are more or less intercorrelated. This fact must be considered for instance using weighted parameters. The distribution-free fourfold test was used for the estimation of the correlation between the parameters. It could be stated that the parameter "largest duration" and "difference mean-median" show the smallest inter-correlation at all investigated stations, whereas the weight of the other parameters can be considered to be similar. Thus the question arises as to the influence of these parameters on the characterisation of hot and cold summers. As a result of a further examination it could be shown that the influence of these parameters is negligibly small. Therefore, it is not necessary to consider them any further.

4. The cluster-analysis

As mentioned above, multivariate statistical methods are necessary to solve the given problems. In the present investigations, the method of cluster-analysis was applied. After an extensive test of different variants of cluster-analysis the iterative minimum-distance-methode (Forgy, 1965) was selected. With this method an attempt is made to sort the parameters for a defined group number. The basic idea is to find the best grouping by an iterative shifting of the parameters between the groups starting from a given initial position. The optimum depends on the chosen criterion of quality. As a criterion of quality the variance criterion can be used. Because of the fact that this criterion presumes the distance to be euclidic, the used parameters must be uncorrelated and comparable in their units. The present parameters do not always meet this condition. So the parameters must be prepared by a suitable transformation before the real cluster-analysis. The further steps of investigation can be described as follows:

1. Definition of the starting partition
2. Calculation of the group centroids
3. Shifting of the elements in the group with the nearest centroid
4. Iteration of the calculation steps 2. and 3. up to the point where no group changes occur any more.

5. Execution of the calculation

For the calculation, the following steps have been made:

1. Calculation of the parameters for the summer time of each year and each station.
2. Estimation of the starting point:

To cover as many of the extreme summers within one cluster as possible it was considered useful to estimate the group number. Resulting from the assumption that at maximum 10 % of all investigated summers are extreme ones, a best group number $k < 10$ results with a supposed regular

distribution within the clusters. Group numbers between k=5 and k=11 were tested. The best results showed up for k=9. Therefore, this group number was taken as basis for all further investigations.

3. Cluster analysis:
On the basis of the described methods the time series of each station (defined by the parameters) are clustered.

6. Results

The following three important facts can be stated:

1. There are years in which the majority of the investigated stations show an extreme summer of the same category.
2. There are time periods in which extreme summers of the same category are accumulated.
3. There are cases in which single stations show a contrary course to the general trend.

For the confirmation of these facts the extreme summers of 18 stations are counted and presented graphically in Figure 1. Three subperiods with a different character can be defined within the total time period:

1. 1901-1927 - An outstandingly great number of cold and cool summers.
2. 1928-1953 - An outstandingly great number of hot and warm summers.
3. 1954-1980 - Nearly equal distribution of extreme summers.

Additionally, the deviation of the annual means from the northern hemisphere mean temperature is presented in *Figure 1. The deviation is 10-year-lowpass filtered.* It can be seen that the frequency of the occurence of extreme summers in Central Europe is correlated with the behaviour of the temperature on the northern hemisphere. A check with the fourfold test is significant for an error of probability of 1 %. The pertinent correlation coefficient amounts to r=0.40. The reason for this relatively small value can be explained by the fact that some of the stations have a contrary course to the usual trend.

The following conclusion can be drawn:
The results described above can be used to confirm the hypothesis that the behaviour of extreme events is a good indicator of climate changes. On the basis of the evident statistical connection there is a possibility to conclude from the trend of the frequency of extreme summers in Central Europe to the trend of mean temperature of the northern hemisphere.

The internal time structure

1. Introduction

The investigation of extreme events is of importance in two respects: On the one hand, the question must be answered if a connection exists between the occurence of extremes and the mean climatic development (see above). On the other hand, the question arises if internal structures exist within longtime extreme events from which the development of the event can be deduced already at a relatively early stage. Information of this kind provides an essential contribution to the limitation of damages in *case of climatological extremes.*

2. Method

Within the further discussion an extreme event will be defined as a long-lasting anomaly of one weather situation. Of further importance for the consideration is that the internal time structure of an extreme event can be understood as the temporal development within a defined *event (i.e., not the description of distribution structures or similar things!).* Of course, a certain (minimum) number of extreme events must be available to save the statistical statements. If this condition is fulfilled, one can proceed as follows to prove internal time structures:

1. Check if there are enough extreme events.
2. All periods of the event are devided into partial intervals in such a way that the partial interval i includes all time periods Δt between 1 and i. The number of partial intervals is equal to the length of a period of an event (expressed by units of Δt).
3. For all partial intervals of each event the necessary parameters to characterize this event must be estimated
4. Estimation of the rank of each parameter of the partial interval of one event with reference to all events.
5. Estimation of the rank sums of each partial interval by summing-up the ranks of all parameters. This estimation of the rank sums takes place for all partial intervals and parameters.
6. Calculation of the sum-curve of the rank sums for each event according to the temporal sequence of partial intervals.

From the family of rank-sum curves obtained, evidence can be derived for the internal time structure of extreme events. The most important question is: Is there a moment t_E within the development of all events starting from which the family of rank-sum curves of the extreme events is significantly different from all others? The earlier this moment can be found, the better a definition of the development of extreme events is possible, and the clearer its own internal structure is evident.

The principal course of the family of rank-sum curves is shown in Figure 2.
For a defined day i within the summer period of a defined year, the value of the sum-curve of the rank-sum $RS_i(a)$ is fixed. To save the statistical evidence, a quantil range$\Delta RS_i(a)$ is defined around $RS_i(a)$. For the end of the summer, quantil ranges named hot - warm - moderate - cool - cold are defined. The statistical evidence now reads as follows: What is the probability of the sum curves going through$\Delta RS_i(a)$ to meet a defined quantil at the end of the summer?

3. Results

Within the investigation above, the necessary estimation of the extreme summers for all stations were conducted. To realize a complexe view, the following parameters were estimated from the daily mean and the daily maximum of the air temperature: Heat sum - summer days - hot days - summer mean. To use the greatest possible number of data $\Delta t = 1$ was chosen. The beginning and the end of the summer were defined by 1st May and 30th September.
Figure 3 presents the results (probabilities for the whole time period shown in a continuous pattern) of hot summers for Prague station and two different time periods. Four conclusions can be drawn:

1. There is a continuous increase of the probabilities of hot summers.
2. The probabilities for the occurence of hot and cold summers turn out inversely.
3. The probabilities of the longer time period are much greater than those of the shorter one.
4. Maximum information is obtained from 5th-14th June.

The inverse case can be investigated in the same way. As shown in Figure 4, the curves of the probabilities of cold summers have for Potsdam and Fanoe station a similar course as descibed in the case above.

4. Conclusions

From the presented results the following conclusions can be drawn:

1. Extreme summers in Central Europe have a typical internal time structure during their development.
2. The method presented here shows a practicable way to describe the internal time structure of extreme events.
3. For some stations in Central Europe it is possible to give a statement on the probability of the occurrence of an extreme summer at a relatively early stage.
4. The longer the investigated time series is the more reliable the statements become.

References

Forgy, E.W.(1965): Cluster analysis of multivariate data: Efficiency versus interpretability of classifications. Biometrics 21.

Gerstengarbe, F.-W., Werner, P.C. (1987): Einige Anmerkungen zur Extremwertproblematik - Z. Meteorol. 37,5.

-- (1988): A method for the statistical definition of extreme value regions and their application to meteorological time series - Z. Meteorol. 38,4.

-- (1991): Some critical remarks on the use of extreme value statistics in climatology - Theor. Appl. Climatol. 44.

Table 1. *European meteorological stations with long-time data sets (1901—1980)*

COUNTRY	STATION
Austria	Sonnenblick Vienna
Belgium	Uccle
ČSFR	Prague
Denmark	Fanoe Island
Federal Republic of Germany	Frankfort/M. Hamburg Hohenpeißenberg Jena Munich Potsdam Stuttgart
Great Britain	Oxford
Portugal	Lisbon
Russia	St. Petersburg
Switzerland	Geneva Zurich
Ukraine	Kiew

Table 2. *Selected parameters of the air temperature for the definition of extreme events by the cluster-analysis.*

PARAMETER	DEFINITION	PERIOD
Number of summer days	$T_{Max} \geq 25°C$	May — September
Number of hot days	$T_{Max} \geq 30°C$	May — September
Heat sum	$\sum T_{Max} \geq 20°C$	May — September
Summer mean	T_{Mean} / n	June — August
Extreme value mean	$(T_{M1}+T_{M2}+T_{M3})/3$	June — August
Longest duration	$T_{Max} \geq 25°C$	May — September
Difference mean-median	$T_{MaxM} - T_{Max50}$	May — September

n = number of days June-August
T_{M1}, T_{M2}, T_{M3} = monthly maximum of the air temperature June, July, August
T_{MaxM} = mean maximum of the air temperature for the summer
T_{Max50} = median of the daily maxima of the air temperature for the summer

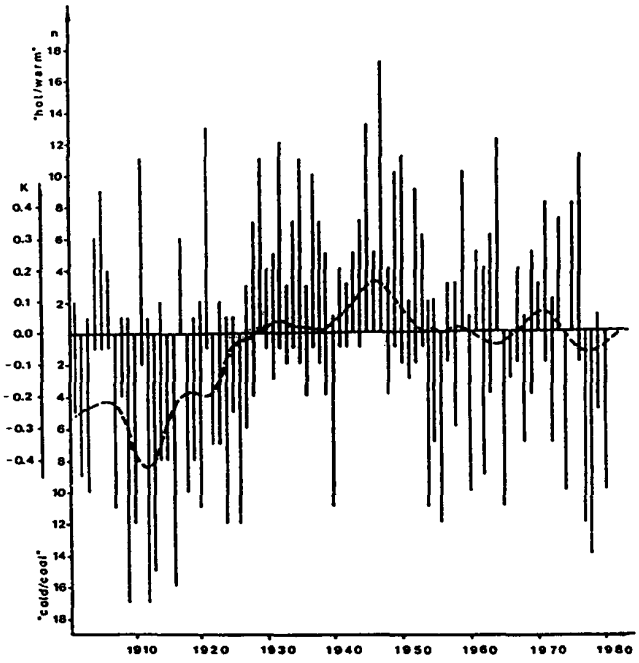


Fig. 1 Number of stations with hot/warm and cold/cool summers and 10-year lowpass filtered deviations of the annual mean values of the air temperature from the long time mean of the northern hemispheric air temperature (*dashed line*) *n = number of stations*

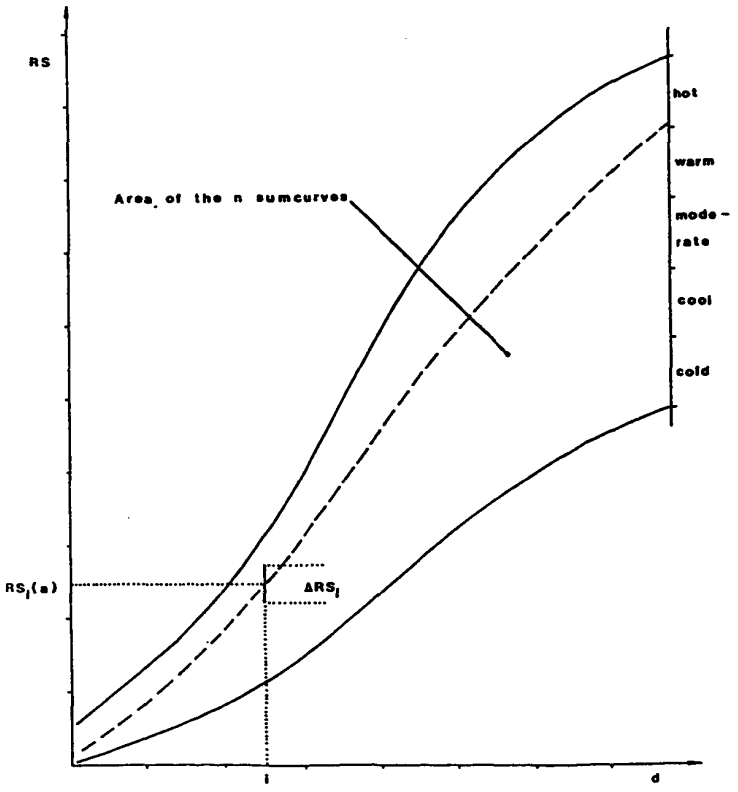


Fig. 2 Principal scheme for the representation of the sum curves of the rank sums

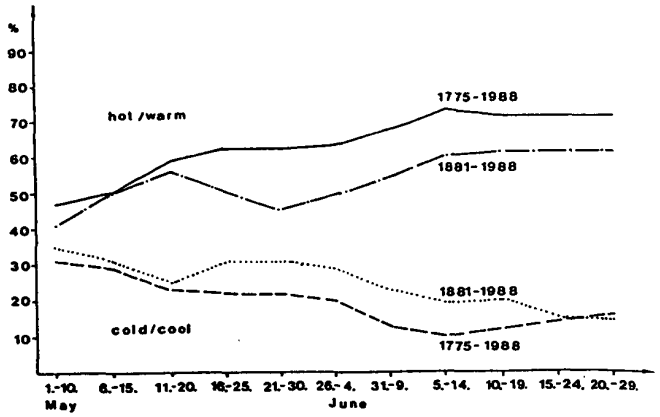


Fig. 3 Probabilities of hot/warm and cold/cool summers of Prague station and different time periods

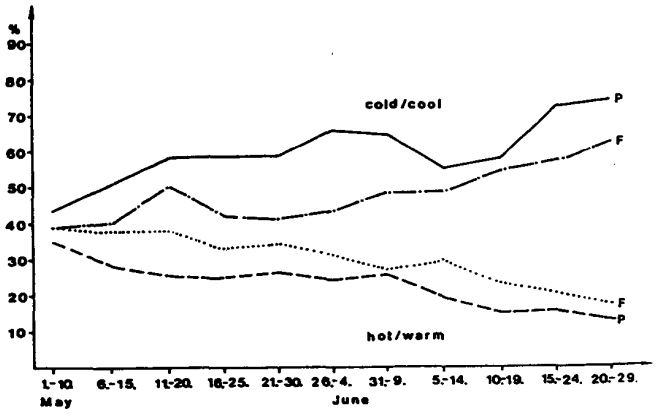


Fig. 4 Probabilities of cold/cool and hot/warm summers of the stations Potsdam (P) and Fanoe (F)

Dependence between Extreme Precipitation Amounts and Rainfall Depth-Duration-Frequency Curves

T. A. BUISHAND
Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands

Introduction

Extreme-value analysis of climatic data is an area of wide interest. For precipitation there is often a need for information about extreme amounts for various durations, e.g. the 1-, 2- and 3-day annual maxima. Quantiles of the distributions of these maxima can be presented as rainfall depth-duration-frequency curves. Each curve gives a specific quantile as a continuous function of the duration D.

There is a rather strong dependence between the annual maxima for two different durations. For instance, in the Netherlands the correlation coefficient between the 1- and 2-day annual maximum amounts is about 0.85. Up to now this dependence has been ignored in studies on rainfall depth-duration-frequency curves.

The present paper deals with the correlation between the estimated parameters of the annual maximum distributions and its consequences for quantile estimation from rainfall depth-duration-frequency curves. Gumbel distributions are fitted to the annual maxima in a 72-yr record at De Bilt for D = 1, 2, 4, 7 and 10 days. Correlation coefficients between the estimated Gumbel parameters are presented. The derivation of rainfall depth-duration-frequency curves from these correlated parameter estimates and the determination of the standard errors of the quantile estimates are discussed in the second part of the paper.

Correlation between estimated Gumbel parameters

The Gumbel distribution is defined by:

$$F(x) = \Pr(X \le x) = \exp\left\{-\exp\left[-(x-\mu)/\sigma\right]\right\}. \quad (1)$$

For the annual maxima at De Bilt the maximum likelihood (ML) estimates $\hat{\mu}$, $\hat{\sigma}$ of μ, σ are presented in Table 1. For these maxima it is convenient to introduce the reparametrization: $\tau = \ln \mu$ and $c = \sigma/\mu$. The ML estimates $\hat{\tau}$, $\hat{c}$ of τ, c are practically uncorrelated (Buishand 1992). In addition, their standard errors are almost constant over the considered range of durations.

Let $\theta = (\tau_1, c_1, \ldots, \tau_k, c_k)'$ denote the vector of Gumbel parameters for k different durations and let $L(\theta)$ be the combined log-likelihood for all durations under the assumption of independence:

$$L(\theta) = \sum_{i=1}^{k} L_i(\tau_i, c_i) \quad (2)$$

TABLE 1. Maximum likelihood estimates of Gumbel distribution parameters for D-day annual maximum precipitation amounts at De Bilt (1906-1977).

D	$\hat{\mu}$(mm)	$\hat{\sigma}$(mm)	$\hat{\tau} = \ln \hat{\mu}$	$\hat{c} = \hat{\sigma}/\hat{\mu}$
1	29.3	7.04	3.38	0.241
2	38.8	8.87	3.66	0.228
4	49.3	11.35	3.90	0.230
7	62.5	14.00	4.14	0.224
10	75.6	16.40	4.33	0.217

with L_i the log-likelihood of the maxima for the i-th duration. The ML estimates $\hat{\tau}_i$, $\hat{c}_i$ for the individual durations maximize $L(\theta)$ with respect to θ. These estimates form the vector $\hat{\theta}$. A recent procedure of Smith (1992) for spatially dependent data can be used to estimate the covariance matrix of $\hat{\theta}$. The estimate is derived from the following approximation to cov $\hat{\theta}$:

$$\operatorname{cov} \hat{\theta} \approx H^{-1} V H^{-1} \quad (3)$$

where H is the information matrix, $H = -E[\nabla^2 L(\theta)]$, and $V = \operatorname{cov}[\nabla L(\theta)]$; ∇ and ∇^2 denote gradient and Hessian respectively. For the Gumbel distribution the matrix H is easily obtained. Because of the correlation between the annual maxima the equality V = H no longer applies. To derive an estimate of V, the gradient of $L(\theta)$ is written as:

$$\nabla L(\theta) = \sum_{j=1}^{J} \nabla h_j(\theta) \quad (4)$$

where $h_j(\theta)$ is the contribution of year j to $L(\theta)$. Since the maxima in separate years are approximately independent and identically distributed, it follows that:

$$V = \operatorname{cov}\left[\nabla L(\theta)\right] = J \operatorname{cov}\left[\nabla h_j(\theta)\right] \quad (5)$$

and an estimate of V is obtained by replacing the covariance matrix of $\nabla h_j(\theta)$ by the empirical covariance matrix of the observed $\nabla h_1(\hat{\theta})$, ..., $\nabla h_J(\hat{\theta})$.

Table 2 shows the resulting correlation matrices of the ML estimates $\hat{\tau}$, $\hat{c}$ for the annual maxima at De Bilt. There is a very strong correlation between the estimates of τ at different durations. Even for the 1- and 10-day annual maxima the correlation coefficient is

TABLE 2. Estimated correlation matrices of the Gumbel parameters τ (under the diagonal) and c (above the diagonal) for D-day annual maximum precipitation amounts at De Bilt (1906-1977) according to Smith's method, the jackknife and the bootstrap.

D	Smith					Jackknife					Bootstrap				
1	*	0.72	0.43	0.23	0.15	*	0.72	0.41	0.22	0.14	*	0.70	0.40	0.21	0.15
2	0.83	*	0.56	0.37	0.23	0.83	*	0.54	0.36	0.22	0.80	*	0.52	0.35	0.23
4	0.68	0.80	*	0.78	0.62	0.68	0.80	*	0.79	0.62	0.67	0.80	*	0.79	0.64
7	0.53	0.66	0.90	*	0.86	0.53	0.66	0.90	*	0.86	0.53	0.64	0.89	*	0.86
10	0.50	0.60	0.81	0.90	*	0.50	0.59	0.81	0.90	*	0.50	0.60	0.83	0.93	*
D	1	2	4	7	10	1	2	4	7	10	1	2	4	7	10

still about 0.5. Although correlation is weaker for the estimates of the parameter c, it is still considerable for the maxima at two neighbouring durations.

Table 2 also presents the jackknife and bootstrap estimates of the correlation coefficients. The jackknife estimate of cov $\hat{\theta}$ reads

$$S_{JACK} = \frac{J-1}{J} \sum_{j=1}^{J} \left(\hat{\theta}_{(j)} - \hat{\theta}_{(.)} \right) \left(\hat{\theta}_{(j)} - \hat{\theta}_{(.)} \right)' \qquad (6)$$

where $\hat{\theta}_{(j)}$ is the vector of ML estimates for the case where year j is deleted, and $\hat{\theta}_{(.)} = \Sigma\, \hat{\theta}_{(j)}/J$. In the bootstrap method, B new samples of J years are drawn from the original set of annual maxima by choosing years randomly with replacement. The bootstrap estimate of cov $\hat{\theta}$ is then obtained as

$$S_B = \frac{1}{B-1} \sum_{b=1}^{B} \left(\hat{\theta}_b - \hat{\theta}_{(.)} \right) \left(\hat{\theta}_b - \hat{\theta}_{(.)} \right)' \qquad (7)$$

where $\hat{\theta}_b$ is the vector of ML estimates for the b-th bootstrap sample and $\hat{\theta}_{(.)}$ is now the average of $\hat{\theta}_1, \ldots, \hat{\theta}_B$. The bootstrap estimates in Table 2 are based on 2500 bootstrap samples. From the table it is seen that the jackknife and bootstrap estimates are almost identical to those from Smith's method. An explanation for this striking resemblence is given in Buishand (1992).

Consequences of correlation for statistical inference about quantiles

For the Gumbel distribution the quantile with exceedance probability 1/T ("T-year event") is given by

$$x_T = \mu + \sigma y_T = e^{\tau} (1 + c y_T) \qquad (8)$$

where $y_T = -\ln[-\ln(1-1/T)]$. To present these quantiles as rainfall depth-duration-frequency curves, it is necessary to describe the change of the parameters τ and c with duration by a continuous function. These two parameters can be treated separately because their ML estimates are practically uncorrelated. For the annual maxima at De Bilt a linear relationship was assumed between τ, c and the logarithm of duration:

$$q_i = a + b \ln (D_i/\dot{D}), \qquad i = 1, \ldots, k \qquad (9)$$

Here q_i represents either τ_i or c_i at the i-th duration D_i; $\dot{D}$ is the geometric mean of $D_1, \ldots, D_k$. The regression coefficient a refers to the mean level of $q_1, \ldots, q_k$.

Estimates of the coefficients a and b can be obtained from an ordinary least squares fit to the ML estimates of $q_1, \ldots, q_k$. Despite the rather strong correlation between these ML estimates, the use of generalized least squares does not lead to more accurate estimates of a and b (Buishand 1992). This is due to the special configuration of the explanatory variables.

The estimated regression coefficients with their standard errors are presented in Table 3. For T = 2, 10 and 100 yr the resulting rainfall depth-duration-frequency curves are shown in Fig. 1. For the Gumbel parameter c the regression coefficient b is not significant and therefore this parameter was assumed to be the same for all durations in the derivation of the quantile estimates in Table 3 and Fig. 1.

For the calculation of the standard errors of $\hat{a}$ and $\hat{b}$ the correlation between the estimated Gumbel parameters has to be taken into account. These standard errors can be obtained from the expression for the covariance matrix of ordinary least squares estimates in the linear regression model with correlated disturbances (Draper and Smith 1981, p. 110) or by a resampling technique. The values in Table 3 are based on the approximated covariances from Smith's method. The table shows that negation of correlation leads to a serious underestimation of the standard error of $\hat{a}$. The standard error of $\hat{b}$ is, however, reasonably reproduced under the independence assumption.

The correlation between the estimated Gumbel parameters also affects the standard errors of the quantile estimates from the rainfall depth-duration-frequency relationships. From a Taylor expansion the following approximation to the variance of the estimated T-yr event is obtained:

TABLE 3. Estimates of the coefficients a and b in Eq. (9) and the estimated 10-yr event for D = 1, 4 and 10 days with their standard errors for the 1906-1977 precipitation record at De Bilt. The standard errors in parentheses assume independence between the annual maxima at different durations.

	Estimate	Standard error	
Parameter τ			
a	3.879	0.025	(0.013)
b	0.403	0.013	(0.015)
Parameter c			
a	0.228	0.015	(0.009)
b	-0.009	0.012	(0.011)
10-yr event (mm)			
D = 1 day	43.9	1.7	(1.2)
4 days	76.9	2.6	(1.4)
10 days	111.2	4.0	(2.7)

$$\text{var}\,\hat{x}_T \approx x_T^2\,\text{var}\,\hat{\tau} + y_T^2 e^{2\tau}\,\text{var}\,\hat{c}. \qquad (10)$$

When $\hat{\tau}$ and $\hat{c}$ are based on the least squares predictions from model (9), their variances follow from

$$\text{var}\,\hat{q} \approx \text{var}\,\hat{a} + \left[\ln(D/\dot{D})\right]^2 \text{var}\,\hat{b}. \qquad (11)$$

The covariance between $\hat{a}$ and $\hat{b}$ has been neglected here. This is allowed because the D_i's have been standardized by their geometric mean $\dot{D}$ in Eq. (9).

Table 3 shows that the standard error of the 10-yr event is too low when the effect of dependence is ignored. The value of var $\hat{a}$ in Eq. (11) is then incorrect. The most severe underestimation of the standard error of $\hat{x}_T$ occurs at D = 4 days, because the second term in the right-hand side of Eq. (11) is relatively unimportant for this duration.

The T-year event can also be estimated from the annual maxima of the required duration only. Because the information from the maxima at other durations is now discarded, this estimate has larger standard error than that from the rainfall depth-duration-frequency relationships. For the 10-year event this increase in standard error is only about 20% for D = 4 days and is even less for the 1- and 10-day annual maxima (Buishand 1992). For other quantiles the increase in standard error is of the same order of magnitude.

Conclusions

When Gumbel distributions are fitted to annual maximum amounts of different durations there is a rather strong correlation between the estimated parameters. A recently proposed procedure by Smith (1992) to assess the effect of spatial dependence on standard errors of parameter estimates can be used to approximate the correlation coefficients between the estimated Gumbel parameters at different durations. This correlation has to be taken into account when standard errors of quantile estimates from rainfall depth-duration-frequency curves are needed. For the annual maxima at De Bilt these standard errors differ very little from those of the quantile estimates from the annual maxima at the required duration only. The rainfall depth-duration-frequency curves should therefore primarily be viewed as a tool to obtain the quantiles at any desired duration.

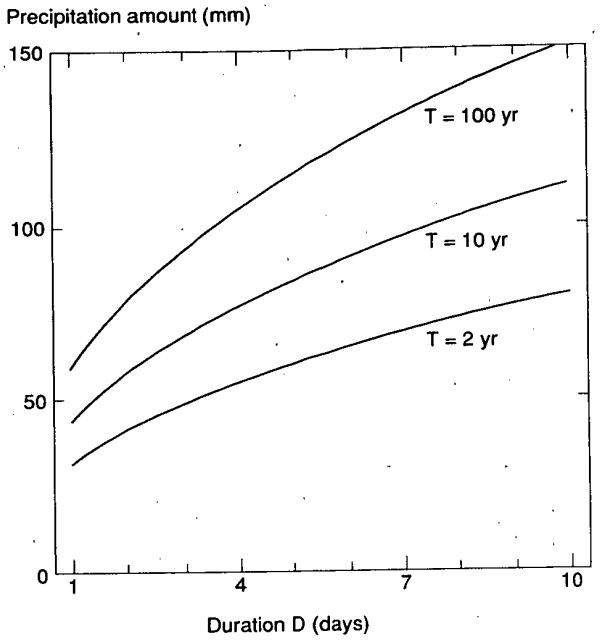


FIG. 1. Rainfall depth-duration-frequency curves for De Bilt.

Acknowledgment. I would like to thank Professor R.L. Smith for a copy of his manuscript "Regional Estimation from Spatially Dependent Data".

REFERENCES

Buishand, T.A., 1992: Rainfall depth-duration-frequency curves: A problem of dependent extremes. *Statistics in the Environment*, V. Barnett and K.F. Turkman, Eds., John Wiley and Sons (to be published).

Draper, N.H., and H. Smith, 1981: *Applied regression analysis, 2nd edition*. John Wiley and Sons, 709 pp.

Smith, R.L., 1992: Regional estimation from spatially dependent data. *Appl. Statist.* (in press).

A TECHNIQUE FOR INCORPORATING INFORMATION ON STORM STRUCTURE INTO EXTREME WIND AND WAVE ESTIMATES

D.T. Resio[1], V. Swail[2], and M. Sager[1]

Resio (1985) applied a method of extremal estimation based on spatial envelopes to the problem of estimating extremal distributions of wind speeds at a site. This technique appeared to provide improved estimates of extreme wind conditions over its region of application (Scotian Shelf and Grand Banks) and showed significant potential for application to spatial data sets, such as typical of data returned from remote sensing applications. Recently, some conventionally estimated extreme values of wave heights for a potentially important oil field development have been dramatically exceeded by a large extratropical storm (with some mixed tropical characteristics). This storm occurred during the last part of October and early November and has been termed the Halloween storm of 1991.

In this paper we first examine spatial aspects of storm structure and its effects on exceedance probabilities. The data set for this analysis is a set careful kinematically-analyzed wind fields for the 70 largest storms to occur in the Scotian Shelf/Grand Banks area during the interval 1959 through 1991 (including the Halloween storm). By selecting subregions within a storm as a basis for sample stratification, it is shown that estimates of extremes based on combinations of sample subsets can deviate significantly from estimates based on assuming winds and waves from all 70 storms are drawn from a single population. Furthermore, it is shown that stratifications based on simple synoptic classifications can also influence the extrapolated results.

Finally, we used this approach, along with a survey of severe storms from 1899 through 1991, to characterize the estimated return period for the Halloween storm. Since this storm has been termed an "outlier" in conventional analyses, it has been difficult to obtain a straightforward estimate of its return period in order to estimate the risk that such a storm might occur again during the lifetime of the offshore oil field operations.

[1] Department of Oceanography, Ocean Engineering, and Environmental Science, Florida Institute of Technology, Melbourne, Florida
[2]Canadian Climate Centre, Atmospheric Environment Service, Downsview, Ontario

A COMPOUND DOUBLE EXPONENTIAL DISTRIBUTION APPLICATION TO THE DETERMINATION OF EXTREME VALUE DISTRIBUTIONS FROM RIGHT-SIDE CENSORED SAMPLES

by

R. Sneyers and M. Vandiepenbeeck

Institut royal météorologique de Belgique, Brussels, Belgium

Abstract

The compound double exponential distribution considered is a distribution from which location and scale parameters are themselves random variables.

In the case of right-side censored random samples from the same parent population, the censored samples being not homogeneously distributed, due to internal correlation, the compound double exponential distribution gives a means of representing the distribution of the elements of the samples.

If F(x) is this distribution and if N is the size of the censored samples, the paper gives the parameters of the double exponential distribution approximating F(x) and of the distribution function $\{F(x)\}^N$ of the extreme values.

As an example, the case of a short period of observations (five years) is considered for the determination of the distribution of the annual maximum windgust in the Belgian network.

1. The one sample problem.

Let $x_1, x_2, \ldots, x_N$, be N i.i.d. elements taken from a population with distribution function F(x). The distribution function of the largest of these N elements is then $\{F(x)\}^N$.

Moreover, if we have

$$F(x) = \exp(-e^{-u}), \text{ with } u = (x-\mu)/\sigma, \quad (1)$$

the distribution function of the largest value x_M becomes

$$\{F(x)\}^N = \exp(-e^{-u'}), \text{ with } u' = u - \log N. \quad (2)$$

Given a sample of this set of variables, it is well known that the ML estimates of the location and scale parameters μ and σ are biased. For this reason, best linear unbiased estimates (BLUE) or minimum variance unbiased estimates (MVUE) are preferred to the preceding ones (Buishand, 1989), among which the ones based on Kimball's estimate of σ may be taken into consideration (Sneyers and Van Isacker, 1972), the loss of efficiency being negligible as far as N is sufficiently large.

The estimates of μ and σ based in this way leading to an estimate of F(x), this gives with (2) the extreme distribution function.

2. The multiple sample problem.

If we have a sequence of n samples of a set of N i.i.d. elements, this estimation procedure may only be applied to the total sample obtained by grouping together the n subsamples, if the n subsamples are themselves i.i.d..

This means that, before proceeding to estimation, both the independence and the homogeneity of the n samples has to be tested.

In the case where homogeneity is rejected, but if independence may be assumed, a solution to the extreme

distribution estimation problem may be found by considering that for each set of sampled elements, the scale and location parameters are themselves i.i.d. variables.

It is obvious that the form of the resulting distribution depends on the ones of the parameters. However, as far as their variances remain small relatively to the corresponding means, it is to be expected that the resulting distribution will remain well approximated by a double exponential distribution.

If this is the case, the computation of moments, with the use of (2), will give the solution of the estimation problem of the extreme value distribution.

From (1) and (2), we have thus

$$x = \mu+u\sigma \quad \text{and} \quad x_M = \mu+(u'+\log N)\sigma. \quad (3)$$

where u and u' have indentical distributions.

Differentiating equations (3), we find for the means and variances

$$E(x) = E(\mu)+E(u).E(\sigma);$$
$$\text{varx} = \text{var}\mu+E(u^2).\text{var}\sigma+E(\sigma^2).\text{var}u;$$

$$E(x_M) = E(x)+\log N.E(\sigma);$$
$$\text{var}x_M = \text{varx}+[2\log N.E(u')+\log^2 N]\text{var}\sigma; \quad (4)$$

where in these relations

$$E(u) = E(u') = 0{,}5772\ldots \text{ (Euler constant)}, \quad \text{varu} = \text{varu}' = \pi^2/6,$$

$$E(u^2) = E(u'^2) = \text{varu}+E^2(u) = \text{varu}'+E^2(u')$$
$$\text{and} \quad E(\sigma^2) = \text{var}\sigma+E^2(\sigma). \quad (5)$$

The scale and location parameters are then derived from the relations

$$\text{varx} = (\pi^2/6).\sigma^2, \quad \mu = E(x)-E(u).\sigma,$$

$$\text{var}x_M = (\pi^2/6).\sigma_M^2, \quad \mu_M = E(x_M)-E(u').\sigma_M. \quad (6)$$

3. Errors of estimation

Using relations (6) for the corresponding estimates σ^*, μ^*, σ_M^* and μ_M^*, differentiation of the same relations gives

$$d\sigma^* = 3(\pi^2\sigma)^{-1}.\text{dvarx} \quad \text{and} \quad d\mu^* = dE(x)-E(u).d\sigma,$$

from which follows

$$\text{var}\sigma^* = E(d\sigma^*)^2, \quad \text{var}\mu^* = E(d\mu^*)^2$$
$$\text{cov}(\sigma^*,\mu^*) = E(d\sigma^*.d\mu^*) \quad (7)$$

and similar relations for σ_M^* and μ_M^*.

Note that if we put

$$p = \sigma_M/\sigma \quad \text{and} \quad q = (\mu_M-\mu)/\sigma, \quad (8)$$

we have the identities

$$\sigma_M \equiv p.\sigma \quad \text{and} \quad \mu_M \equiv \mu+q\sigma, \quad (9)$$

which allow immediate determination of σ_M and μ_M from the estimates of σ and μ.

It follows that the knowledge of p and q (i.e. from a previous series of observations) leads to improved estimates for σ_M and μ_M.

4. The determination of extreme value distributions under right-censored samples.

In the case of the one sample problem, the set $x_1, x_2, \ldots x_N$, reduces to the N largest values of an observed random sample, for which Weissman (1978) has given the MLEs and the MVUEs for the parameters of the double exponential asymptotic extreme value distribution.

In the case of the multiple sample problem, even if for each sample, the parent population remains the same, the largest values being strongly correlated to the maximum value, the n sets of N values is no longer homogeneous. Neglecting this property and basing the estimation on the total sample leads to underestimated fractiles of the extreme distribution, as was experienced in Sneyers and Vandiepenbeeck (1983).

To clear the ambiguity raised by the possibility that the non homoge-

neity of the successive parent populations is involved in the non-homogeneity of the n sets, the homogeneity of the estimates given by Weissman's procedure applied to each of the n sets may be tested by an appropriate chi-square test.

On the other hand, the fact that the largest values near the extreme value have distributions differing few from the double exponential one (Gumbel, 1958 p.189) enables to apply the above described method based on the consideration of a compound double exponential distribution.

Moreover, this has the advantage that it is applicable even if the parent populations are not homogeneous.

5. An application to the annual heaviest windgust in Belgium.

In the case of short series of observations, it is especially important to know that improved estimations of the extreme value distribution is possible through modelling right-censored samples.

This possibility has been especially useful to solve the problem of the determination of the mean recurrence interval of the heavy storms which occurred in Belgium during January and February 1990 and wrought extended damages all over the country (Sneyers, 1990).

In fact, the complete change of the anemometric network in 1985 lead to the consequence that the available series of observation was reduced to the five years 1985 to 1989.

To investigate the efficiency of our method, a homogeneous series of observations of the ten largest windgusts observed at the central station of Brussels-Uccle for the period 1961-1983 was at disposal.

The statistical analysis of this series showed: non-homogeneity between the annual sets (Kruskal and Wallis test, 1952), chronological independence inside the annual series (Wald and Wolfowitz serial correlation test, 1943), fit of the double exponential distribution to the total sample and independence as well as randomess of the distributions of the scale and location parameters of the double exponential distributions adjusted to each of the sets.

These last distributions were found to be respectively normal and gamma for the location and the scale parameter.

Compared to the direct estimation of the extreme distribution from the observed annual extremes, our method was found to be respectively 1,42 and 4.0 times more efficient than the direct method for the location and the scale parameter when p and q are unknown, this factor raising to 2,24 and 5,5 when p and q are given (Sneyers and Vandiepenbeeck, 1991).

The result was that the obtained accuracy in this way is equivalent to the one found for the old network based on twelve years of observations (Sneyers, 1971).

Note that the knowledge of p and q allowed better estimates then the one given by Weissman's method, but that this is not the case without this knowledge.

Conclusion

The results obtained in the above given example demonstrate sufficiently the advantage of the developed method.

Accessorily, the contrast between the increase of efficiency for the respective estimates of the location and scale parameters shows, if still necessary, that the loss of information caused by the interruption of long homogeneous series of observations remains irreparable, especially in the case of the location parameter.

References

Buishand, T.A., 1989. Statistics of extremes in climatology. Statistica Neerlandica,43, 1, 1-30.

Gumbel, E.J., 1958. Statistics of Extremes. Columbia Univ. Press, New-

York, 376p.

Kruskal, W.H., and Wallis, W.A. 1952. Use of ranks in one criterion variance analysis. J. Amer. Statist. Ass., 47, 583-621.

Sneyers, R., 1971. Les Pointes Maximales du Vent en Belgique.Inst. R. Météor. de Belgique, Pub.A, N'73, 32p.

Sneyers, R. and Van Isacker, J., 1972. Sur l'ajustement de la loi de répartition de Fisher-Tippett de type I au moyen des estimateurs de Kimball. Rev. Statist., Infor., Rech. Opér., 12, 1, 36-43.

Sneyers, R. and Vandiepenbeeck, M., 1983. On the Use of Large Values for the Determination of Maximum Values. Arch. Met. Geoph. Biocl., Ser. B, 32, 279-286.

Sneyers, R., 1990. Rapport sur les pointes maximales du vent mesurées à Uccle et en Belgique et la prévention des dégâts de tempête. Inst. R. Météor. de Belgique, Internal report.

Sneyers, R. and Vandiepenbeeck, M., 1991. Note on the Use of Large Values for the Determination of Maximum Values. Applied Stochastic Models and Data Analysis. Proceedings of the Fifth Symposium on ASMDA, Granada, Spain, 23-26 April 1991, Eds: Gutiérrez, R. and Valderrama, M.J., World Scientific, Singapore, New Jersey, London, Hong-Kong., 568-573.

Wald, A. and Wolfowitz, J., 1943. An exact test for randomness in the non parametric case based on serial correlation. Ann.Math. Statist., 14, 378-388.

Weissman, I., 1978. Estimation of Parameters and Large Quantiles Based on the k Largest Observations. J. Amer. Statist. Ass., 73, 364, 812-815.

STUDY ON THE GENERALITY OF THE GAMMA PRECIPITATION DISTRIBUTION MODEL

Ding Yuguo

(Nanjing Institute of Meteorology, China)

ABSTRACT

It is theoretically demonstrated by using the well-known polynomial distribution and from the viewpoint of statistical physics that the most frequent distribution for precipitation frequency is bound to be the Gamma distribution. And verification of vast numbers of observational data have shown that the optimal model for precipitation frequency distribution is undoubtedly the Gamma distribution. Thus, a model in common use is proposed, which is applicable to the Gamma distribution of precipitation in an arbitrary time period and month at any place.

1. Introduction

The precipitation probability distribution patterns have been long studied by many scholars. Particularly, studies of the Gamma distribution for fitting precipitation have yielded positive results (Yao and Ding, 1991, Suzuki, 1980). Another family of distribution except the Gamma model is also investigated for describing precipitation. Nowadays researches in this field remain very active. At the four international meetings on statistical climatology held during the last ten years or so, academic papers on the distribution models have occupied an important position.

The temporal and spatial precipitation distribution is very complicated. Although some physical and dynamical processes can be interpreted from the contributing factors, it is difficult to describe satisfactorily their temporal and spatial distributions. Quite a few scholars have used the Gamma distribution for fitting precipitation and the parameter estimation has been discussed from different angles. Yet, there remains a question to be investigated as to the further theoretic demonstration of which probability distribution model

is the optimal one and what its generality is like.

In this study, the generality of the Gamma distribution is further demonstrated by using probability theory and from the viewpoint of statistical physics.

2. The Most Frequent Distribution of Precipitation Frequency

Suppose that the precipitation measured in a single process at a given place is bound to fall in one of the numerical intervals $[x_0, x_1), [x_1, x_2), \ldots, [x_K, x_{K+1})$, where $x_0 \geq 0.0$ mm, and for convenience, is written successively as $A_0, A_1, \ldots, A_K$. The probability specified to fall in the interval A_i $(i = 1, 2, \ldots, k)$ is $P_i = p(A_i)$, and thus it can be proved that the precipitation obtained from N independent measurements falls in the frequencies of all the A_i intervals, that is, $n_0, n_1, \ldots, n_K$, with their joint probability as

$$P(n_0, n_1, \cdots, n_K) = \frac{N!}{\prod_{i=0}^{K} n_i!} \prod_{i=0}^{K} P_i \qquad (1)$$

We hope to find a set of frequencies $(n_0, n_1, \ldots, n_K)$, which is the most possible numerical value for precipitation distribution, i.e., to find the frequency distribution with the maximum likelihood. This is equivalent to the determination of the maximum of the function $P(n_0, n_1, \ldots, n_K)$ given $p_0, p_1, \ldots, p_K$ under certain conditions.

Statistical physics shows that if an isolated system is in equilibrium, it will occur in each possible state with equal probability, and if the system does not occur in each possible state with equal probability, it will not be in equilibrium and then will vary with time until it eventually reaches each possible state with equal probability. From the viewpoint of climatological equilibrium, the probabilities p_i of any place falling in the intervals $A_0, A_1, \ldots, A_K$ are in fact equal.

By deduction and verification, we have the precipitation frequency distribution in the sense of the maximum likelihood as the exponential distribution (a special case of distribution)

$$f(x) = \lambda e^{-\lambda x} \qquad x > 0 \qquad (2)$$

As a single precipitation process at a given place is actually the basis for forming the climatic-statistical precipitation amount (e.g., daily, pental, dekad, monthly, seasonal and annual), it is not difficult to prove by using the property of distribution function that the greatest part of precipitation is the Gamma distribution (with the execption of annual precipitation amount).

3. Verification of Cases

In this study, patterns of the monthly precipitation probability distribution from more than 20 (secular) weather stations in eastern China have been verified. It

is found that among the various patterns the Gamma distribution is the best and none of the distributions, such as Weibull distribution, lognormal distribution, Kappa distribution and normal distribution, are as good. Table 1 gives the fitting results of five distribution models for June precipitation amount (1873-1982) at Shanghai. It is found in the table that in the fitting of the five distribution models, the smallest value of χ^2(ambiguity) is the Gamma distribution, only 0.9912, and the relative error is also small, only 0.141, obviously superior to the fitting results of other distributions.

Based on the theoretical demonstration and actual calculation, we have further reasons to design a general Gamma distribution model applicable to all natural time priods in a year. Thus, we have

$$f(x) = \frac{\beta^{\alpha^*}}{\Gamma(\alpha^*)} x^{\alpha^*-1} e^{-\beta x} \qquad (3)$$

where $\alpha^* = \sum \alpha_i$ is the parameter of shape and β the parameter of scale

Such results are of great significance to climatic impact assessment, water conservancy works and agricultural production.

Table 1. Comparison between the fitting results of June Precipitation Frequency Distribution (110 yrs)

Spacing (mm)	Frequency (measured)	Theoretical Frequency Distribution				
		Gamma	Log-normal	Weibull	Kappa	Normal
0-50	4	4.4	2.9	3.9	18.1	9.4
50-100	11	10.5	10.8	10.6	17.3	11.5
100-150	17	18.2	19.7	18.7	16.5	17.1
150-200	24	22.0	23.4	24.1	16.5	21.0
200-250	18	19.8	19.9	23.2	15.2	20.1
250-300	20	17.6	14.5	16.6	13.7	15.3
300-350	7	7.7	8.8	8.5	9.0	9.4
350-400	7	5.5	4.7	3.1	3.8	4.4
400-450	1	1.3	2.0	0.5	0.3	0.5
450-500	1	1.0	0.8	0.3	0.2	0.1
χ^2(ambiguity)		0.9912	3.8655	2.9909	20.6633	8.901
relative error		0.141	0.552	0.498	2.952	1.113

Further verification of the relationship of the Gamma precipitation distribution parameters α and β to the length of time period (the number of days) shows that α and β of the total precipitation amount in

n days of different months vary considerably with the season and generally have good secondary curvilinear or approximate linear relationship, while the Gamma distribution parameters of the total precipitation amount in n days at different stations have obvious harmonic features varying with the months (formulas ommitted).

4. Conclusions

From the above disscusion, we may draw some conclusions as follows:

(1) It can be demonsrated theoretically that the Gamma distribution is bound to be the most frequent of the precipitation probability distributions in any period of time.

(2) Verification of observational data indicates that the Gamma distribution is in fact the optimal distribution model for fitting climatological precipitation amount, monthly rainfall for example.

(3) As the precipitation amounts in different short time periods can compose the rainfall of the total time period, various general distribution models can be derived by using the eigenfunction in the Gamma distribution. This idea can also extend to the calculation of regional precipitation amount or the precipitation probability distribution at the neighboring weather station.

REFERENCES

Yao, C. S., 1984: Fundamentals of Climatic Statistics, The Science Press, 161, 516 (in Chinese).

Yao, C. S. and Y. G. Ding, 1990: Climatic Statistics, The Meteorology Press, 156-265 (in Chinese).

Suzuki, E., 1980: Statistical climatology, 1-20.

Zhang, Y. C. and Y. G. Ding, 1991: A general Gamma probability model for precipitation in various periods, Acta Meteo. Sinica, 49, 1, 80-84 (in Chinese).

Ding, Y. G. etc. 1991: Research of the monthly summer rainfall in China, Climatology Research-(Statistical Climatology), Meteo. Press (Beijing).

CHARACTERISTICS OF RAINFALL OVER THE EASTERN REGION OF TIBET PLATEAU IN SUMMER

Xiao-ping ZHONG

Chengdu Research Institut of Plateau Meteorology
Guang Hua Cun
Chengdu, Sichuan 610072
P. R. C.

ABSTRACT

Tibet Plateau, the highest place in the world, has drawn many meteorologists' attention because its special effects on atmosphere, especilly in the lee side, i.e., the eastern region of Tibet Plateau. This study is going to detect the characteristics of precipitation over this region from June through August according to hourly precipitation data by means of some new analysis techniques.

The first problem is how to distinguish the heavy rainfall, or convective rainfall. Due to the different strength of convection in different regions the criterion for convective precipitation used in this study is decided from the percentage frequency distribution of all non-zero observations in the area interested, instead of giving a certain value subjectively. This is a new way developed by the author. For most part of this region, the criterion is about 5mm/hr.

Secondly, diurnal cycle of precipitation will be stressed by using a harmonic analysis method corrected by the author. Results from some areas have shown that more than 70 percent of non-zero observations occurs from 20LST to 08LST of the second day. There is a few non-zero observations in the afternoon and the evening. The maximum frequency appears at about 04LST, and the minimum at 17LST. Different diurnal cycles are obtained by analysis for categories of non-zero observations with different rain intensity. Comparison is also made between the diurnal cycles of non-zero observations and rain volume produced by

them for different categories. There are, of course, some siginificant difference between them. Rain volume for a station in one hour is computed according to the rain intensity and the size of the area controlled by the station. Then put rain volumes of all stations together. Here another problem rises. How to decide the area controlled by one station more accurately? A new way developed by the author is used in this study. Its accuracy, in turn, the accuracy of rain volume, is much higher than conventional ways.

Thirdly, the importance of heavy rain is partialy demonstrated by some statistics. The area with heavy rain is about 20 percent of the entire rain area while the rain volume from heavy rain is more than 70 percent of the total rain volume. From the view of the time scale, the percentage of the times having heavy rain to the total time with all kinds of rain is also very small but the rainfall produced in heavy rain times occupies a large percentage of the total rainfall.

INTERANNUAL VARIABILITY IN THE TROPICAL PACIFIC AS REVEALED BY ROTATED EXTENDED EMPIRICAL ORTHOGONAL FUNCTION ANALYSIS

C. F. Ropelewski[1], M. Chelliah[2], T. Smith[1]
[1]Climate Analysis Center, NMC/NOAA, [2]RDC Corp. Greenbelt MD

Empirical Orthogonal Function (EOF) analysis has become one of the standard tools used, if not overused, in climate studies. While EOFs provide a convenient way to compact large data sets and also provide a means to describe large-scale stationary climate patterns they have inherent short comings for the study of evolving, or propagating phenomenon, Weare and Nasstrom (1982). Furthermore, the orthogonality constraints inherent in simple EOFs can force the analysis into unrealistic, or "unphysical" patterns, Richman (1981), especially for the higher modes if the signal to noise ratio is small.

Wyrtki (1975) suggested that the El Niño/Southern Oscillation (ENSO) can be described as a modulation, or amplification, of the mean annual cycle of sea surface temperature (SST). We are interested in studying the evolution of ENSO's from the framework suggested by Wyrtki but, since both ENSO and the annual cycle of SST show features which appear to propagate, conventional EOF analysis can be misleading. In addition, since ENSO involves both the ocean and the atmosphere we also need to perform an analysis of the <u>co-variability</u> between these components of the system. Finally, since we expect some components of the ENSO to show regionality , rotated EOF's are

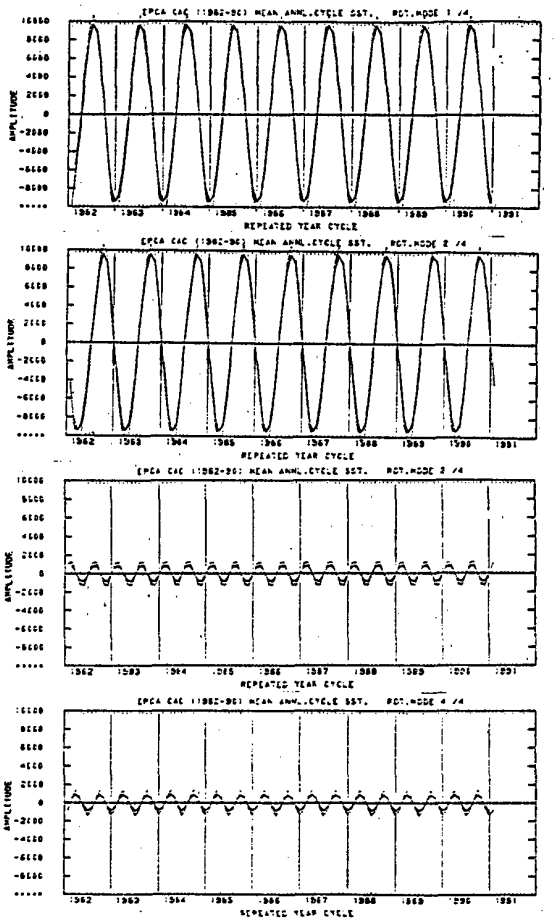


Fig. 1 Time series of the leading 4 Rotated Extended Empirical Orthogonal Functions (REEOFs) for the mean annual cycle of the equatorial Pacific sea surface temperature.

appropriate to the analysis. Thus, we developed a Joint Rotated Extended EOF (REEOF) analysis.

To illustrate the technique in the limited amount of space available here, a REEOF

analysis is presented for the mean annual cycle of SST in the equatorial Pacific based on satellite and ship data Reynolds (1988).

The leading REEOF loadings time series, Fig. 1, break into two pairs, an annual cycle pair, offset by 3 months from each other, and a semi-annual cycle pair, offset 1.5 months. Together these pairs represent 97% of the total variance in the mean annual cycle (86% and 11% respectively).

The spatial patterns for the first REEOF pair, Fig. 2, can be represented by a series of only six patterns since the first two REEOFs are identical with a 3 month loadings offset and the entire pattern repeats with a change in sign in six months. The REEOFs, not only emphasize the evolution of the well documented "cold tongue" in the Eastern Pacific, but also reveal a strong annual cycle in the equatorial SST Pacific from 140^{o}W to 160^{o}W. A strong annual SST cycle centered near 5^{o}S between 170^{o}E and 180^{o} is another interesting feature revealed by this analysis but not a common feature of most standard climatologies. Whether the indications of a strong annual cycle in these regions of the equatorial Pacific is a result of the relatively short analysis period (1982 to 1991) or because we are using a "blended" satellite and ship data set needs to be examined further. However, since we are using a spatially rotated EOF we are fairly confident that these features are not merely an artifact of the analysis technique. Furthermore, the portions of the Pacific discussed here are generally very poorly sampled in standard "ship only" analyses.

The second pair of REEOFs, corresponding to spatial modes 3 and 4, represent semi-annual variability in the SST, Fig 3. Although the REEOF analysis relaxes the ortogonality constraint in the spatial domain, orthogonality is still required in the time domain. Thus the time series of REEOF loadings are orthogonal. This property of the analysis provides an efficient and fairly "clean" way to separate the annual cycle from semi-annual variability.

REFERENCES

Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. Jour.of Clim., 75-86.

Richman, M. B., 1981: Obliquely rotated principal components: An improved meteorological map typing technique? Jour. of Appl. Meteor.,**20**, 1145-1159.

Weare, B.C., and J. S. Nasstrom, 1982: Examples of extended empirical orthogonal function analysis. Mon. Wea. Rev.,**110**, 481-485.

Wyrtki, K., 1975: El Niño-the dynamic response of the Equatorial Pacific Ocean to atmospheric forcing. Jour. of Phys. Ocean., **5**, 572-582.

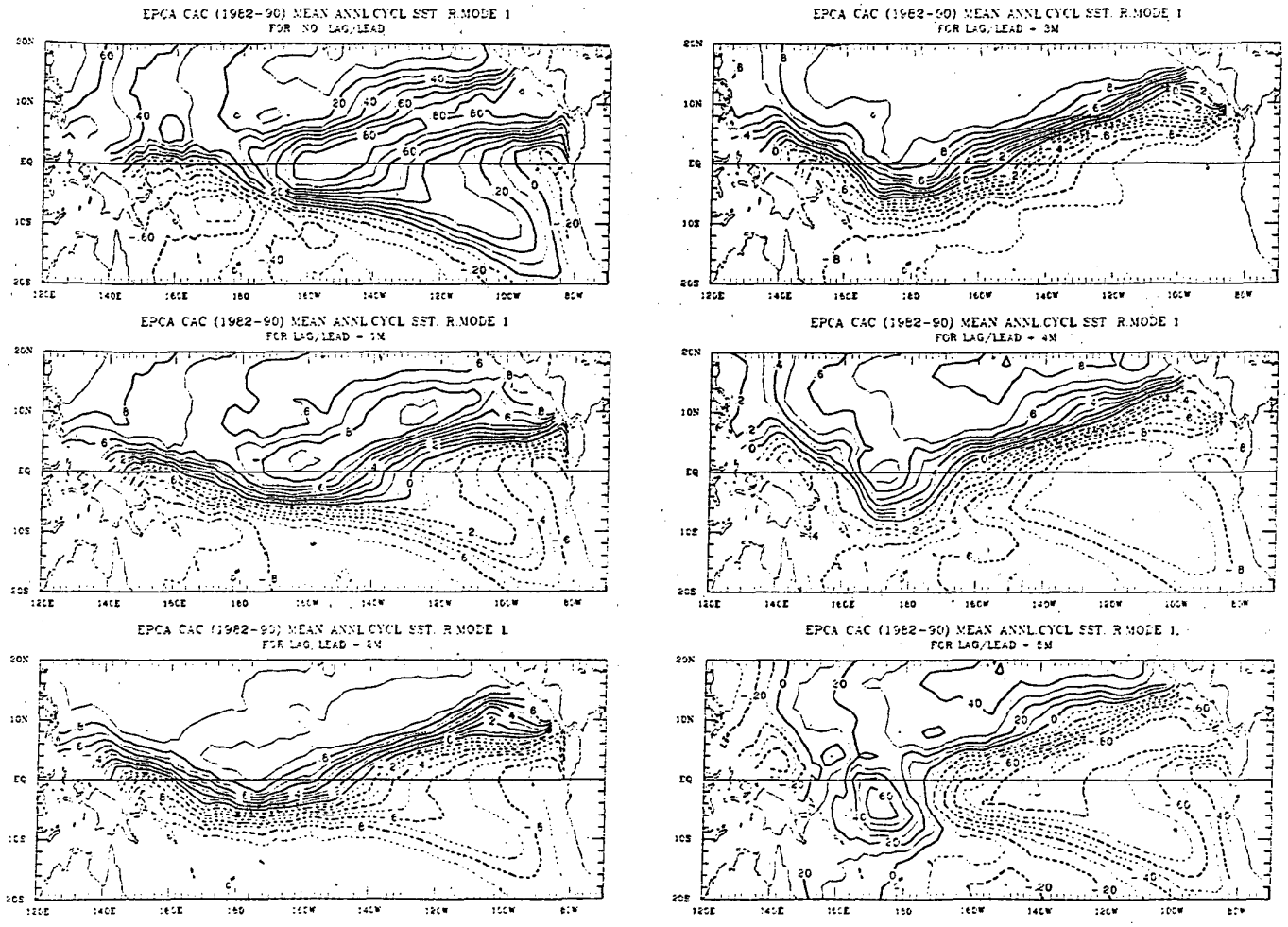


Fig. 2 Rotated extended EOF (REEOF) Mode # 1 for equatorial Pacific mean sea surface temperature, illustrated at 1 month lags. Lags 0 through 5 are shown. The pattern for a six month lag simply repeats the lag 0 pattern (upper left hand panel) with the signs reversed. The leading REEOF accounts for 43% of the total variance. REEOF Mode # 2 is identical but the patterns are shifted by 3 months. This second Mode also accounts for 43% of the total variance. The leading two patterns illustrate a "pure" mean annual cycle based on blended satellite and ship data (Reynolds, 1988) for the period 1982 to 1991.

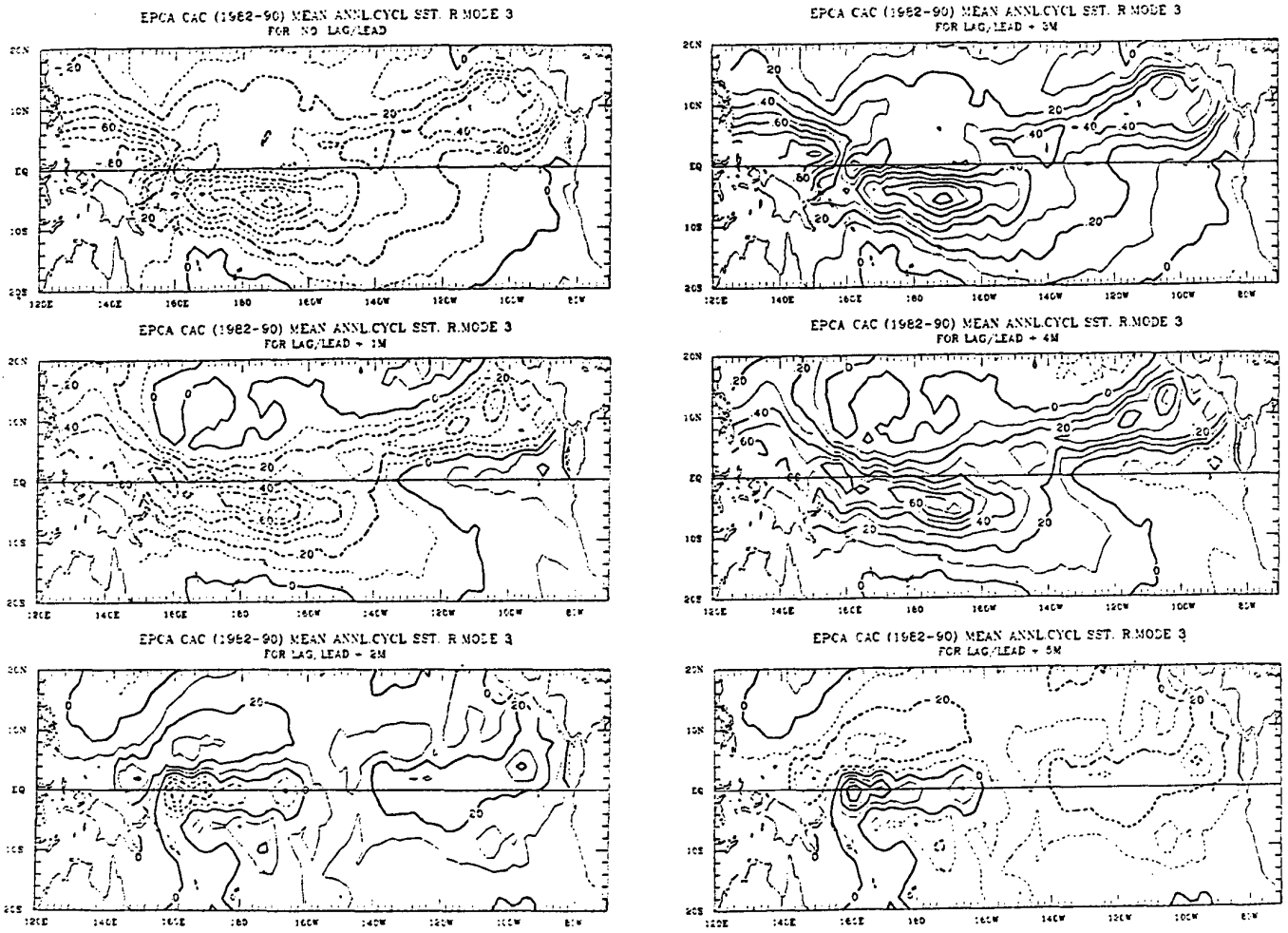


Fig. 3 Rotated extended EOF (REEOF) Mode # 3 for equatorial Pacific mean sea surface temperature, illustrated at 1 month lags. Lags 0 through 5 are shown. Note that the pattern for a three month lag simply repeats the lag 0 pattern (upper left hand panel) with the signs reversed. The Mode #3 REEOF accounts for 5.3% of the total variance. REEOF Mode # 4 is identical but the patterns are shifted by 1.5 months. Mode # 4 also accounts for 5.3% of the total variance. These two modes illustrate a "pure" mean semi-annual cycle based on blended satellite and ship data (Reynolds, 1988) for the period 1982 to 1991.

The Statistical Analysis of Streamflow Patterns in the U.S. in Relation to the Extreme Index Phases of the Southern Oscillation

John A. Dracup and Ercan Kahya

Civil Engineering Department, UCLA
Los Angeles, CA 90024

ABSTRACTS

The potential for long-term streamflow forecasts based on atmospheric circulation indices in the United States has been assessed. The extreme phases of the Southern Oscillation (SO) have been linked to fairly persistent classes of atmospheric anomalies (e.g., precipitation and temperature) over the low and midlatitudes at regional and global scale. This study examines the relationships between the extreme index phases of the Southern Oscillation and unimpaired stream discharges over the contiguous U.S. using 1009 stations. Of particular interest in this investigation is the identification of areas of land that appear to have consistent and strong El Niño-Southern Oscillation [ENSO]/La Niña-Southern Oscillation [LNSO] related streamflow signals. The method of analysis initially involves transformation of monthly streamflows to percentile ranks and construction of the 24-month warm/cold event composite that is based on the nine latest moderate/strong ENSO/LNSO episodes. The first harmonic extracted from a such composite is assumed to be the ENSO/LNSO signal appearing in streamflow anomaly. The amplitude and phase of harmonic indicate the strength of relationship and time of maximum anomaly within the ENSO/LNSO cycle, respectively. The statistical significance for the amplitude is assessed by Schuster's test for autocorrelated series. The vectorial display of these harmonics over a map enables us to subjectively identify some candidate regions that are the areal extents of ENSO/LNSO influence on streamflow. Then an aggregate composite through the percentiles based on log-normal frequency distribution

is constructed in the coherent region to detect the season of apparent signal. The index time series of spatially averaged streamflow percentiles of the season are plotted to confirm the consistency of the proposed relationships.

So far, the strong, coherent, and significant streamflow responses to hypothesized ENSO forcing have been found along the five South Atlantic Coastal [SAC] states and in the six Pacific Northwest [PNW] states. A persistent wet season coincided with the 12 months of ENSO years. The index time series of the region indicated the consistent relations (7 out 9 cases). The pronounced dry and wet periods consistent previous results of precipitation are detected in the SAC core area. The variation in streamflow in relation to ENSO occurred negatively during the (-) October to (0) May season and positively during the (+) January to (+) May season. Time series analysis for the SAC region revealed %100 confirmation for dry season and %78 for wet season. Once an El Nino event sets in, a long-term forecasting utility may be available for these core regions. The results of this analysis, confirming the previous precipitation studies, also demonstrates the midlatitude hydrologic response to tropical ENSO phenomena. The other four candidate regions are under the investigation to determine whether or not any ENSO-related signals exist.

Compositing based upon nine ENSO episodes for the California streams did not reveal satisfactory results, however compositing based upon five Type 1 ENSO events, classification scheme of Fu et al., showed an excellent correspondence. These preliminary findings motivate us to proceed with the same analysis for the Southwest candidate region.

The same analysis based on the La Nina phase of the SO is the second main goal of this study. In order to explain streamflow fluctuations in the regional aggregate composites, the analyses are done for temperature and precipitation stations in the core areas. Some statistical hypotheses testing, as well as interrelations among the core areas are considered through the nonparametric method of Mann-Whitney U and correlation analyses.

A CLIMATOLOGICAL STUDY OF NORTHERN HEMISPHERE WINTER CIRCULATION PATTERNS ASSOCIATED WITH DIFFERENT TYPES OF LARGE-SCALE FORCING

J.A.M. CORTE-REAL and C.C. DACAMARA
Department of Physics, University of Lisbon, Lisbon, Portugal

1. INTRODUCTION

Since the pioneering studies of Charney and Eliassen (1949) and Smagorinsky (1953), forcing has been recognized as a primary factor in the dynamics of General Circulation and Climate.

It has been shown (e.g. DaCamara et al., 1992) that blocking type circulations may be related to thermal forcing, associated with anomaly fields of SST in the Pacific and Atlantic.

The main purpose of the present study is to relate spatial distributions of large-scale forcing with characteristic circulation patterns.

2. SCHEME OF ANALYSIS AND DATASET

The vorticity equation has been used in the diagnosis of forcing effects (e.g. Saltzman, 1962) as well as in simulation of atmospheric responses to prescribed forcings (e.g. Egger 1978).

The forced barotropic two-dimensional quasi-geostrophic vorticity equation may be written

$$\frac{\partial}{\partial t}\nabla^2\Psi + J(\Psi, \nabla^2\Psi + f) = g \qquad (1)$$

$$\Psi = \frac{gz}{f_o}$$

where g is a forcing function related to diabatic and topographic effects (e.g. Mitchell and Dutton, 1981).

The application of the monthly mean operator < > to (1) leads to relation

$$\langle J(\Psi, \nabla^2\Psi + f)\rangle = \langle g\rangle . \qquad (2)$$

which may be used to estimate <g>, through evaluation of time-averaged vorticity advection fields.

Because of the complexity of <g> distributions, the analysis is performed through the use of adequate statistics obtained from a large ensemble of data.

The dataset consists of daily National Meteorological Center (NMC) grid analyses of 500 hPa height fields for 102 winter months (December, January and February), covering the period December 1955 - February 1989.

3. FORCING AND PATTERNS OF CIRCULATION

In the latitudinal band 20^0-60^0N, the grand-mean (over the 102 months) of <g>, denoted by {g}, shows an organized structure over the Oceans, with large areas of positive values (Fig. 1) associated with higher values of standard deviation (Fig. 2). It may be noted that regions of positive values of {g} centered around (40^0N, 150^0W) and (60^0W,40^0N) seem in fair agreement with northern winter locations of warm surface currents, associated with the Pacific and Atlantic gyres (Tolmazin, 1985). On the other hand, tongues of negative {g} values appear over regions of cold surface currents. Values of <g>, spatially averaged over oceanic areas may thus reflect large-scale effects of thermal forcing.

Accordingly, two areas over the Pacific and Atlantic were respectively defined as (18^0-40^0N, 150^0E-150^0W) and (18^0-40^0N, 65^0-5^0W). In these areas, spatial averages of <g>, denoted by g_{Pac} and g_{Atl}, were obtained for the 102 winter months. The winter months were then classified

TABLE 1. Means ($10^{-12}s^{-2}$), standard deviations ($10^{-12}s^{-2}$) and explained variances (%) of g_{Pac} and g_{Atl} for 4 categories of large-scale forcing.

	Pacific			Atlantic		
	Mean	St. Dev.	Exp. Vr.	Mean	St. Dev.	Exp. Vr.
+ -	52.29	9.56	10%	3.22	4.64	4%
- +	25.91	7.70	5%	23.24	10.36	15%
+ +	52.41	8.27	8%	23.13	6.40	8%
- -	28.42	9.89	11%	4.70	6.31	7%
Inter-Groups	----	----	66%	----	----	66%
Grand Mean	40.84	15.38	100%	13.00	11.87	100%

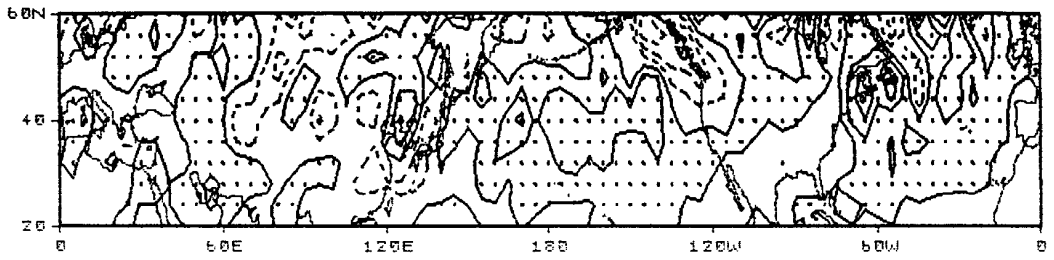


FIGURE 1. Grand-mean of forcing function g for the 102 winter months. Contour interval is $75x10^{-12}s^{-2}$. Regions with positive values are dotted.

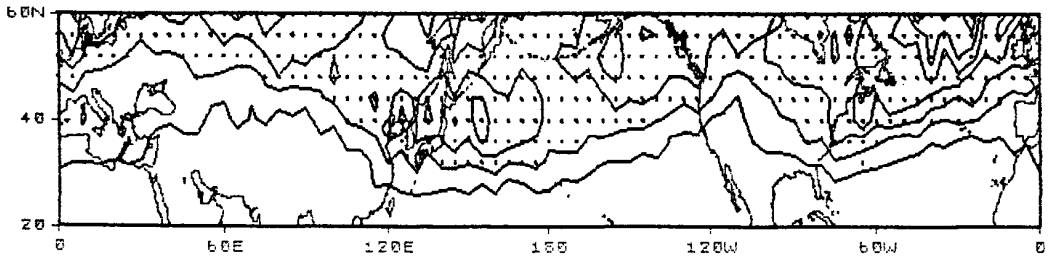


FIGURE 2. Standard deviation of forcing function g for the 102 winter months. Contour interval is $50x10^{-12}s^{-2}$. Regions with values greater than $150x10^{-12}s^{-2}$ are dotted.

into 4 groups, identified by + - (27 cases), - + (20 cases), + + (28 cases) and - - (27 cases), according to the signs of anomalies g_{Pac}-{g} and g_{Atl}-{g}. Table 1 presents means and standard deviations of g_{Pac} and g_{Atl}. Percentages of total variance explained by each group and by differences between groups are also included. It is apparent the 4 chosen groups are distinct among themselves as well as distinct from the grand means. This result indicates the 4 groups, whose spatial structure is shown in Fig. 3, may reflect different regimes of thermal forcing.

Composites of 500 hPa height field anomalies were then obtained for each group (Fig. 4). This figure clearly shows that circulation patterns associated to forcing regimes of opposite characteristics (i.e. + - vs. - + and + + vs. - -) have anomaly fields with spatial distributions of opposite sign. On the other hand, over the Oceans, south of 50°N, positive (negative) values of g-anomalies reflect on negative (positive) height anomalies.

It is worth mentioning height anomaly fields of the 4 identified types of winter circulation patterns closely resemble those based on different types of blocking (DaCamara et al., 1992). These large-scale patterns are related to different planetary wave activities, have energetics of their own and seem associated with typical patterns of SST anomalies. In fact, composites corresponding to groups - +, + - and + + have structures similar to those of Atlantic, Pacific and concurrent (Atlantic and Pacific) blocking type circulations, whereas those of - - are characteristic of high index circulation. Agreement between classifications based respectively on blocking and on large-scale forcing gives further indication that <g> fields in oceanic areas may adequately reflect thermal forcing effects.

4. CONCLUDING REMARKS

The systematic analysis adopted in this study, on the basis of forcing fields obtained from the barotropic quasi-geostrophic equation, may be associated with different types of characteristic circulation patterns. Given the observed persistence of large-scale thermal forcing, it is hoped the present research may have some relevance in studies on long-range forecasting of the northern hemisphere winter circulation.

REFERENCES

Charney, J.G. and A. Eliassen, 1949: A numerical method for predicting the perturbation of the middle-latitude westerlies. Tellus, 1, 38-54.

DaCamara, C.C., E.C. Kung, W.E. Baker, B.-C. Lee and J.A.M. Corte-Real, 1992: Long-term analysis of planetary wave activities and blocking circulation in the Northern Hemisphere winter. Beitr. Phys. Atmosph. (accepted for publication).

Egger, J., 1978: Dynamics of blocking highs. J. Atmos. Sci., 35, 1788-1801.

Mitchell, K.E. and J.A. Dutton, 1981: Bifurcations from stationary to periodic solutions in a low-order model of forced, dissipative barotropic flow. J. Atmos. Sci., 38, 690-716.

Saltzman, B., 1962: Empirical forcing functions for the large-scale mean disturbances in the atmosphere. Geofisica Pura e Applicata, 52, 173-188.

Smagorinsky, J., 1953: The dynamical influence of large-scale heat sources and sinks on the quasi-stationary mean motions of the atmosphere. Quart. J. Roy. Meteor. Soc., 79, 342-366.

Tolmazin, D., 1985: Elements of Dynamic Oceanography. Allen and Unwin, Winchester, MA, 181 pp.

ACKNOWLEDGMENTS

The authors are indebted to Mr. Ricardo Trigo for his technical assistance during the preparation of this paper. This research, was supported by the National Council for Scientific and Technological Research (JNICT) Grant PEAM/C/-CVC/12/91.

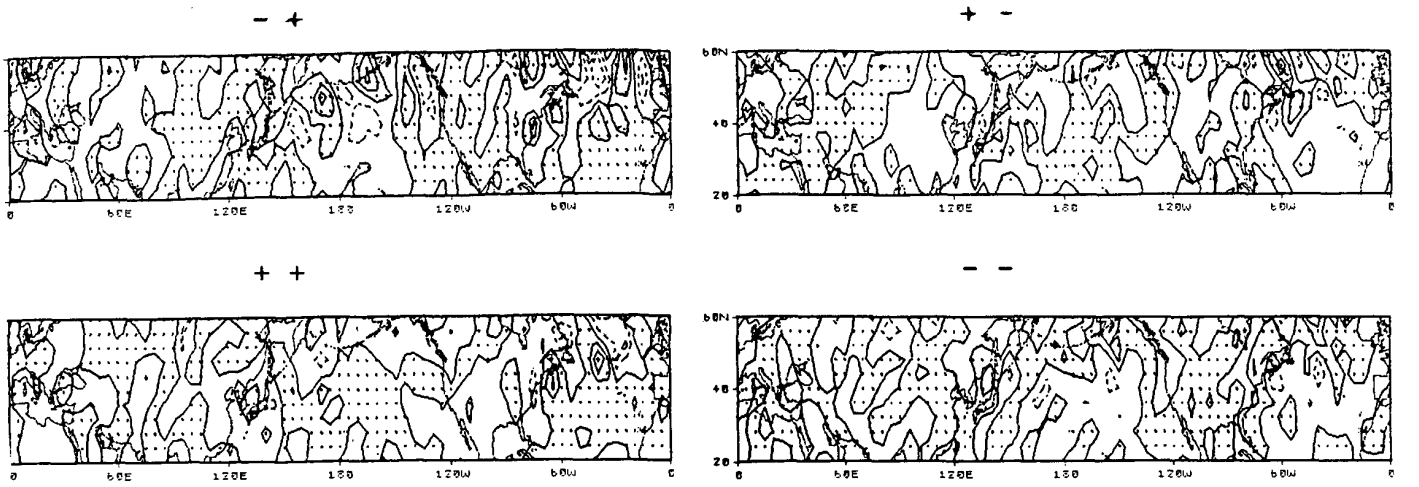


FIGURE 3. Composites of monthly anomaly fields of g for 4 categories of large-scale forcing. Contour interval is $50 \times 10^{-12} s^{-2}$. Regions with positive anomalies are dotted.

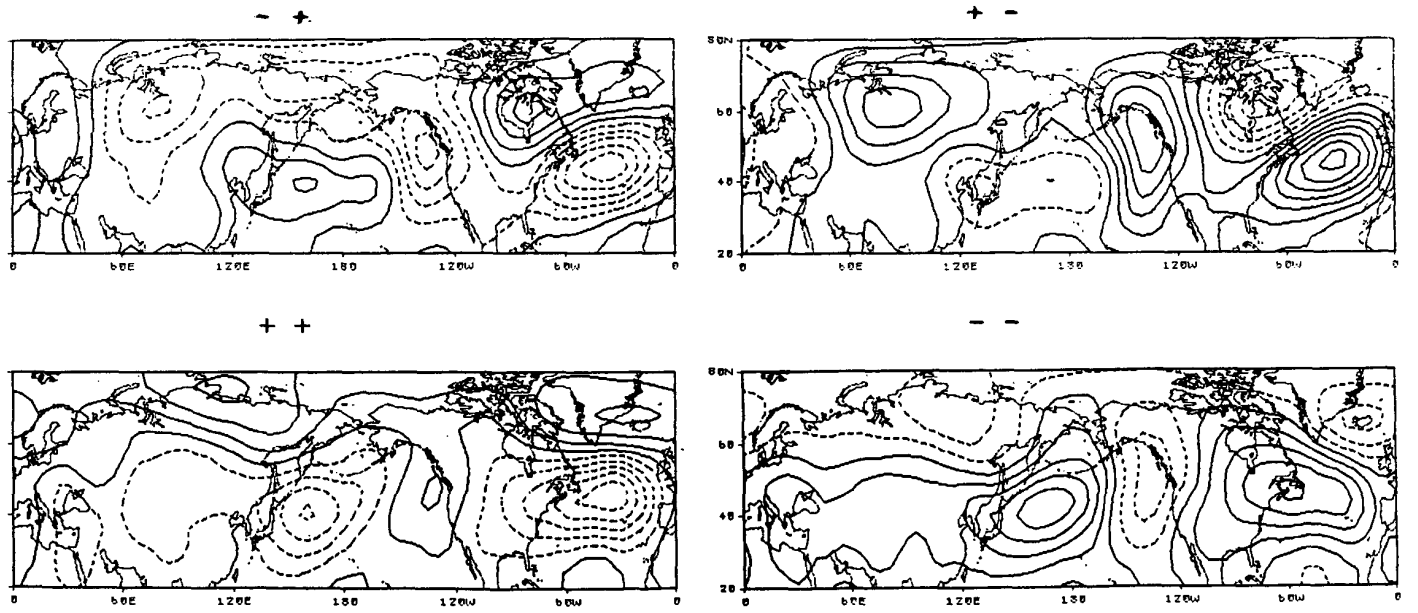


FIGURE 4. As in Figure 3 but for monthly anomaly fields of 500 hPa height. Contour interval is 10m. Dashed contours indicate negative anomalies.

MONTHLY ANTICYCLONICITY AND CYCLONICITY IN THE SOUTHERN HEMISPHERE; AVERAGE FOR MARCH AND SEPTEMBER.

R. M. Leighton

National Meteorological Centre
Bureau of Meteorology, Melbourne, Australia

1. INTRODUCTION

An excellent account of the early history of synoptic meteorology and climatology is given by Klein (1957) with references to the tracking and frequency of cyclones and anticyclones in the Northern Hemisphere. By the end of the nineteenth century many studies of cyclone frequency and tracks were made not only for the U.S. but for Europe, the Atlantic, India, the west Indies and the Far East. The first thorough investigation for the Northern Hemisphere was completed in 1894 containing the average frequency of cyclones during each of the 12 months and their principal tracks.

During the twentieth century notable work on various aspects of synoptic climatology in the Southern Hemisphere has been done by Berry et. al. (1952); Streten and Troup (1973); Streten (1980); Streten and Pike (1980); Guymer and Le Marshall (1980); Trenberth (1984); Le Marshall et. al. (1985); Trenberth and Mo (1985); Trenberth (1991) and Jones (1991).

For the Australian region the method of representation of the surface circulation for long periods (i.e. months, season, etc) by charts of anticyclonicity and cyclonicity was developed during 1951-53 in connection with the problem of extended and long-range forecasting in Australia. Later charts of 7 year averages (1946-1952) of monthly and seasonal anticyclonicity and cyclonicity and their description were published by Karelsky (1954). Charts of monthly and seasonal anticyclonicity and cyclonicity for the Australian region for 15 years (1946-1960) were published by Karelsky (1961). Leighton and Deslandes (1991) published monthly charts for 23 years (1965-1987) for the Australian region including the eastern Indian Ocean.

2. ANTICYCLONICITY AND CYCLONICITY

The term "anticyclonicity" ("cyclonicity") is defined as the time in hours during which anticyclone (cyclone) centres occupied a given 5° cell north of 55°S during a given period (i.e., week, month, season, etc.). South of 55°S cells are 5° latitude by 10° longitude.

Cyclonicity and anticyclonicity were computed from the plotted locations of cyclone (anticyclone) centres at 00 UTC and 12 UTC and, assuming a uniform speed of movement, the systems could be tracked. Single centres which were not part of a track were assigned a lifetime of 6 hours.

Cyclones whose centres were tracked fitted into one or more of the following classes; identifiable closed circulation with or without an associated front, thermal (heat) low, wave lows which developed into closed circulations, tropical lows and tropical cyclones.

Anticyclones in general covered much larger areas than cyclones. On occasions when anticyclones where shown to have a double centre both centres were considered when "budding" (an anticyclone centre forming in a ridge which is extending from and

established anticyclone) was the obvious cause. Centres in ridges were taken into account when there was an identifiable outflow or as a prelude to anticyclogenesis.

Manually drawn archive charts prepared by the National Meteorological Centre of the Australian Bureau of Meteorology were used to obtain means of cyclonicity, anticyclonicity, and immobility for March and September from 1973-1987 (1965-1987 in the Australian Region (90°E-180°E)).

3.CHARTS OF CYCLONE (ANTICYCLONE) IMMOBILITY

Immobility of a cyclone(anticyclone) centre is defined as the time taken by the cyclone (anticyclone) centre to traverse a 5° cell. The average immobility of cyclone (anticyclone) centres across a 5° cell was calculated by dividing the cyclonicity (anticyclonicity) of a 5° cell by the number of cyclone (anticyclone) centres which have occupied that square during the same period.

4. DISCUSSION.

Anticyclonicity

In the Indian Ocean the anticyclonicity chart for March shows a broad maximum lying across the ocean between 35°S and 40°S. By contrast the anticyclonicity for September lies across the ocean between 30°S and 35°S. Peaks in the maxima are comparable in both months near 100°E and near 60°E. These months clearly show the seasonal shift in latitude of the paths preferred by anticyclones although there are similarities in preference of longitudinal position.

Across the Australasian region the anticyclonicity chart for March has distinct maxima at higher latitudes than 35°S in the Bight and the Tasman Sea. Notable minima occur south of the west coast of Australia, across southeastern Australia and across New Zealand. Anticyclones located in the east Indian Ocean can extend a ridge under W.A. into the Bight where a new anticyclone centre forms in the ridge - a process known as "budding". Anticyclones are known to remain in the Bight before a similar budding process occurs across southeastern Australia and anticyclone centres strengthen in the Tasman Sea. New Zealand is also a notable region for budding. In September the anticyclonicity maxima are located between 30°S and 35°S at longitudes similar to that of March.

The Pacific Ocean has anticyclonicity maxima at high latitudes between 40°S and 50°S east of New Zealand both in March and September although the March values are much higher. Across the mid-Pacific the March maxima values lie between 35°S and 45°S whereas in September the maxima are between 30°S and 40°S. In the eastern Pacific the anticyclones tend to be located between 30°S and 35°S near 90°E in March and between 25°S and 35°S east of 90°E in September.

In the western Atlantic Ocean near 40°W the anticyclonicity maximum in March is between 35°S and 40°S and in September between 30°S and 35°S. In the eastern Atlantic Ocean near 0° the March maximum lies between 30°S and 35°S whereas the September maximum is between 25°S and 30°S. The March maxima values are notably higher than the September values similar to the other oceans.

Cyclonicity

Across the Indian Ocean at high latitudes both months show cyclonicity maxima at the same locations with the largest maxima for the hemisphere between 60°S and 65°S south of Western Australia.

Cyclonicity maxima cover the Tasman Sea and much of the central Pacific Ocean for both

months. In September maxima values are particularly high across the Pacific between 40°S and 50°S. A significant minimum is highlighted between 50°S and 60°S before the maxima occur in the sub Antarctic pressure trough near 70°S to 75°S.

Cyclones which form in the south west Pacific tend to move northeastwards into the mid-Pacific. These cyclones and the low centres which develop on the South Pacific Convergence Zone account for much of the cyclonicity maxima in the central Pacific. A col region or a narrow latitudinal ridge located between 50°S and 60°S usually separates the cyclones from the large southern depressions located in the sub-Antarctic pressure trough. If high latitude ridging cradling lower latitude cyclonic cells identifies a blocking regime then the greater part of the South Pacific could be considered a potential region for the occurrence of blocking patterns.

In the southwestern Atlantic Ocean cyclonicity maxima in both months are associated with cyclogenesis in the frontal zone off South America particularly in September and developments in easterlies over or near the South American continent.

In both months the cyclonicity maxima across northwestern Australia and northern Argentina are related to heat lows.The Australian heat low is more evident than the corresponding heat low in South America because the location in Argentina is over relatively high and dry terrain east of the Andes where the reduction of station level pressure to mean sea level pressure is a problematic procedure.

In March the section of maxima across tropical waters in the Indian Ocean, the Australasian region and the western Pacific indicate regions for tropical lows and cyclones.

5. CONCLUSION

(Anti)cyclonicity averages supplement conventional climatologies by showing regions of slow and rapid movement of surface centres and the location of potential blocking areas. As such they can be used as a guide for forecasting. The averages could be useful for verifying the realism of Regional and Global Weather Prediction and the basic figures, being on a five degree grid, could be used as a data base in models. Anomalies can be calculated for ongoing current months and these would help in defining severe weather occurrences or possible climate change.

Acknowledgements

The author wishes to thank Roger Deslandes for assistance in the Australian Region, Terry Skinner for comments on the work and on the manuscript and Cheryl Hatcher for typing the manuscript.

References

Berry, F.A., Owens, G.V. and Wilson, M.P. 1952 'Artic Weather Maps', U.S. Navy Bureau of Aeronautics, Project AROWA, 102 pp.

Guymer, L.B. and Le Marshall, J.F. 1980. 'Impact of FGGE buoy data on Southern Hemisphere Analyses', Aust. Meteorol. Mag 28, 19-42

Jones, P.D. 1991. 'Southern Hemisphere Sea-Level Pressure Data: An Analysis and Reconstruction back to 1951 and 1911', J. Climatol. 11 585-607

Karelsky, S. 1954. 'Surface circulation in the Australasian region', Bureau of Meteorology Study No. 3, AGPS, Canberra 45 pp.

Karelsky, S. 1961. 'Monthly and seasonal anticyclonicity and cyclonicity in the Australian region - 15 year average (1946-1960)', Bureau of Meteorology Study No. 13, AGPS, Canberra, 11 pp.

Klein, W.K. 1957. 'Principal tracks and mean frequencies of cyclones and anticyclones in the Northern Hemisphere', Research Paper No. 40, U.S. Weather Bureau, Washington D.C.

Leighton, R.M. and Deslandes, R. 1991, 'Monthly Anticyclonicity and Cyclonicity in the Australasian Region: Averages for January, April, July and October', Aust. Met. Mag, 39, 149-154.

Leighton, R.M. and Deslandes, R. 1991. 'Monthly Anticyclonicity and Cyclonicity in the Australian Region: 23 year (1965-1987) Averages', Bureau of Meteorology Technical Report No. 64, 29 pp.

Le Marshall, J.R., Kelly, G.A.M. and Karoly, D.J. 1985. 'An atmospheric climatology of the Southern Hemisphere based on the ten years of daily numerical analyses (1972-1982): I overview', Aust. Meteor. Mag., 33, 65-85.

Streten, N.A. 1980. 'Some synoptic indices of the Southern Hemisphere mean sea level circulation 1972-1977' Mon. Wea. Rev., 108, 18-36.

Streten, N.A. and Pike, D.J. 1980. 'Indices of the mean monthly surface circulation over the Southern Hemisphere during FGGE', Aust. Meteorol. Mag., 28, 201-215.

Streten, N.A. and Troup, A.J. 1973. 'A synoptic Climatology of satellite observed cloud vortices over the Southern Hemisphere', Quart. J. R. Met. Soc. 99, 56-72.

Trenberth, K.E. 1984. 'Interannual variability of the Southern Hemisphere circulation: representativeness of the year of the Global Weather Experiment', Mon. Wea. Rev., 112, 108-123.

Trenberth, K.E. 1991. 'Storm tracks in the Southern Hemisphere', J. Atmos. Sci. 48, 2159-2178.

Trenberth, K.E. and Mo, K.C. 1985. 'Blocking in the Southern Hemisphere', Mon. Wea. Rev. 113, 3-21.

Interannual Variability of Tropical Cyclone Movement over the Western North Pacific

JOHNNY C. L. CHAN
Department of Applied Science, City Polytechnic of Hong Kong

1. Introduction

The movement of a tropical cyclone (TC) is to a large extent governed by the environmental flow surrounding the cyclone (Chan and Gray 1982). This flow is in turn influenced by circulations of planetary scale which have interannual variations. Therefore, the movement of TCs can be expected to have similar variations. A knowledge of such variations are important in the seasonal prediction of the occurrence of TCs in a particular region.

In a recent study, Chan (1990, 1991) found that annual tropical cyclone occurrence over regions of the western North Pacific (WNP) can be predicted by patterns of sea-surface temperature (SST) and the 850- and 500-hPa zonal winds. These parameters are related to the genesis and intensification processes as well as the "steering" flow. The present study represents an extension of this previous research by analyzing the interannual variations of the direction of TC movement over the WNP. These variations are correlated with the 850- and 500-hPa zonal winds. Correlations between the direction of TC movement and the SST have also been studied. However, no significant correlation can be found.

The data and methodology used in this study are described in Section 2. Regions of largest variability in the direction of TC movement are also identified. Because of space limitations, only results in one region will be shown. Other results will be presented at the Conference. Correlations between these variations in this region and the principal components of the 850-hPa zonal winds are given in Section 3 together with plausible physical explanations of these correlations. The results are summarized in Section 4.

2. Data and methodology

a. Data

Tropical cyclones over the WNP between 1947 and 1988 form the basic data set for this study. Best tracks of these cyclones are obtained from the Royal Observatory, Hong Kong. These tracks for years prior to 1984 were analyzed only in the region west of 160°E. Therefore, this study is only limited to cyclones west of this longitude. This limitation, however, should not restrict the usefulness of the study as most of the TCs over the WNP form west of 160°E (Xue and Neumann 1984).

To correlate TC movement in the WNP with the large-scale environment in the region, the 850- and 500-hPa zonal wind components are analyzed. Data for these parameters are obtained from the US Climate Analysis Center in a gridded format. The years of coverage is 1974-89 for the 850-hPa and 1970-89 for the 500-hPa zonal winds. The horizontal resolution of each parameter is 5° longitude in the zonal direction and between 3.5° to 5° latitude in the meridional direction.

The fields of each parameter is represented by a set of empirical orthogonal functions (EOFs) using the method described by Shaffer and Elsberry (1982). It is found that the first three EOFs explain 62% and 71% of the variance of the 850- and 500-hPa zonal winds respectively. Furthermore, the sampling errors of the first three eigenvalues in each field calculated using the method of North *et al.* (1982) are much smaller than the spacing between neighbouring eigenvalues. These sample EOFs should therefore be representative of the actual EOFs.

The principal components (PCs) associated with these EOFs are then used to correlate with the annual number of TCs moving within a particular direction range in a pre-defined region. For the results to have predictive value, only the PCs between the months of November of the previous year to April of the current year are included in the regression analysis.

b. Grouping of TCs based on the direction of movement

In Chan's (1990) study of TC activity, only the number of TCs which occurred in each 5° x 5° grid box needs to be counted. However, within each grid box, the TCs can move in a variety of directions. Therefore, it is necessary to group them into direction categories. The interannual variations in the number of TCs *with a given range of direction of movement* in a particular grid box or region can then be studied.

Four direction categories are defined based on the Cardinal directions (Table 1). These represent the most frequent directions of TC movement in the WNP. Other movements are ignored as they are generally associated with erratic tracks or TCs close to becoming extra-tropical. The direction of movement is calculated using the ± 6h best-track positions.

TABLE 1. Direction categories defined in this study.

Direction Category	Range (Degrees)
westward	270 ± 22.5°
northwestward	315 ± 22.5°
northward	337.5 - 360°, 0 - 22.5°
northeastward	45 ± 22.5°

c. Identifying regions of largest variability

To identify regions which have the largest variability in each of the direction categories defined in Table 1, the following procedure is used.

For each 5° x 5° longitude-latitude grid box, the direction of a particular TC in a given year when it first enters the box is calculated. The number of cases for that year in the particular direction category to which this direction belongs is then increased by one. During the transit of the TC through the grid box, if its direction remains in the same category, the number of cases in that category is not increased. However, if the TC changes direction within the grid box such that the new direction belongs to another category, the number of cases for the year in this new category is then increased by one. This procedure therefore counts the *number* of TCs having the same direction (range) within each box in each year. The standard deviation of this number for all the years in the data set is also calculated.

In the *westward* category, the axis of maximum occurrence, as expected, is in the tropics (Fig. 1) with the centre around (15°N, 125°E). The area of large interannual variation, however, occurs generally south of this axis (Fig. 2). At these low latitudes, it might be expected that most of the TCs would be travelling westwards. The existence of such a high value of standard deviation σ suggests that in some years, even TCs in these locations do not necessarily have a westward motion. Since in this region of largest variability, westward TCs are likely to move into the South China Sea, a knowledge of the conditions leading to this variability should help in the seasonal prediction of the number of TCs entering the South China Sea.

Based on Fig. 2, a region of "maximum variability" bounded by (7.5°N, 117.5°E) and (17.5°N, 142.5°E) is defined. For each year, the number of TCs having a westward movement (as defined in Table 1) in each 5° x 5° grid box in this region is then totalled. This gives the number of occurrences of westward TCs within the region.

In each of the other three categories in Table 1, similar regions have been identified.

These results will be presented at the Conference.

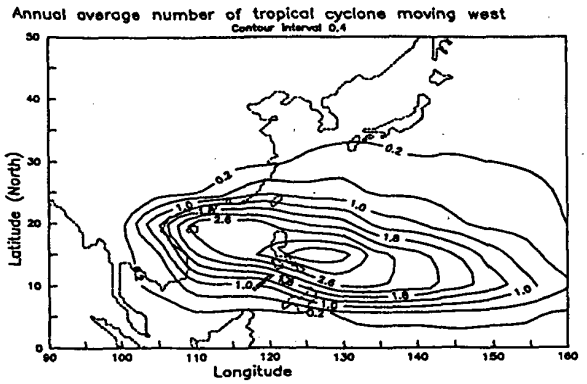


Fig. 1. Annual average number of westward-moving TCs in every 5° x 5° latitude/ longitude box.

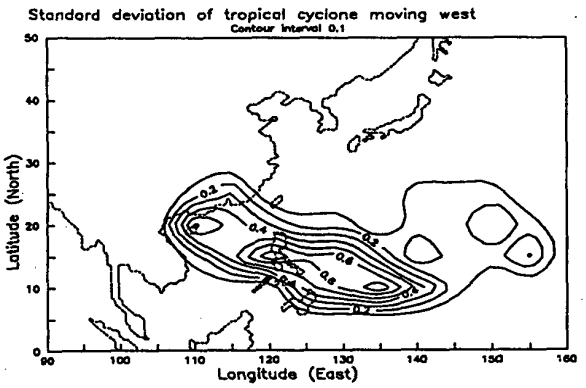


Fig. 2. Standard deviation of the averages in Fig. 1.

3. Correlation with 850-hPa zonal winds

A linear stepwise multiple regression analysis is made between the anomalies in the number of occurrences of westward TCs and the first three PCs of the 850-hPa zonal winds for the months of November of the previous year to April of the current year. This gives 18 potential predictors. Assuming only 4 or 5 will be chosen, the F-to-enter value is set at 5.0. Because only 14 years of wind data are available, the regression analysis is performed using PCs of one month at a time. All the selected predictors are then combined to obtain the final set of predictors.

Two variables are selected by the regression procedure : first principal component (labelled as PC1) of March and PC2 of January. These two predictors together explain a 64% of the total variance of the number of occurrence of westward TCs in the region defined in the last section.

A reconstruction of the pattern of the first EOF for March in years with a large number of occurrence of westward TCs in the region east of the Philippines shows particularly strong easterly winds in the tropics (not shown). However, the easterly winds are weak for years with a small number of occurrence. Thus, the results obtained from the regression analysis appear to be reasonable. The physical interpretation of the January EOF2 is rather difficult as the principal component associated with this EOF correlates with the residual of the variance not explained by PC1 of March.

Nevertheless, predictions are made from these two principal components using the entire dependent sample as well as the jackknife method. The results are shown in Table 2.

It can be seen from Table 2 that the predictions from both the dependent sample and using the jackknife method give similar results although the former, as expected, has a smaller root-mean-square (RMS) error. Other than a few cases, the errors are within one standard deviation of the observed number of occurrences. The RMS errors are also quite small. This result suggests that by using the 850-hPa zonal winds, the number of occurrences of westward TCs to the east of the Philippines can be rather accurately predicted.

4. Summary

Because of space limitations, only one result can be presented. Nevertheless, the correlation between the principal components of the 850-hPa zonal winds and the number of occur-

rences of westward TCs in the region east of the Philippines suggest that the interannual variations of TC movement can indeed be related to the large-scale flow patterns. Furthermore, such relationships can have some seasonal predictive value. Other results will be presented at the Conference to illustrate this point further.

TABLE 2. Prediction of the annual number of occurrences of westward TCs in the region bounded by (7.5°N, 117.5°E) and (17.5°N, 142.5°E) using PC1 in March and PC2 in January of the 850-hPa zonal winds. The "independent" sample result is obtained using the jackknife method. The error of prediction is in parentheses.

		Predicted	
Year	Observed	Dependent Sample	Independent Sample
1975	15	16(+1)	18(+3)
1976	20	22(+2)	22(+2)
1977	21	23(+2)	24(+3)
1978	31	25(-6)	25(-6)
1979	32	33(+1)	33(+1)
1980	32	26(-6)	25(-7)
1981	36	33(-3)	32(-6)
1982	34	29(-5)	29(-5)
1983	24	26(+2)	27(+3)
1984	21	27(+6)	32(+11)
1985	25	30(+5)	30(+5)
1986	35	31(-4)	31(-4)
1987	39	40(+1)	40(+1)
1988	33	36(+3)	38(+5)
Mean	28		
Standard deviation	7.3	RMS error	3.8 5.1

Acknowledgements. The author would like to thank the Royal Observatory Hong Kong for providing the best-track data and the Climate Analysis Center of the United States for providing the data on the 850- and 500-hPa zonal winds. The programming and data analysis efforts of Mr. S. H. Lam are also gratefully acknowledged. Ms. Rolin M. N. Ng expertly helped in the manuscript preparation.

This research is supported by the Indicated Research Grant No. 700039 of the City Polytechnic of Hong Kong and the Earmarked Grant No. 904024 of the Universities and Polytechnics Grants Committee of Hong Kong.

References

Chan, J. C. L., 1990: Regional interannual variability of tropical cyclones in the western North Pacific. *Dept. of Appl. Sci. Res. Rep.* No. AP-90-06, City Poly. of Hong Kong, 40pp.

Chan, J. C. L., 1991: Prediction of annual tropical cyclone activity in the western North Pacific. *Dept. of Appl. Sci. Res. Rep.* No. AP-91-05, City Poly. of Hong Kong, 24pp.

Chan, J. C. L., and W. M. Gray, 1982: Tropical cyclone movement and surrounding flow relationships. *Mon. Wea. Rev.*, **110,** 1354-1374.

North, G. R., 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.

Shaffer, A. R. and R. L. Elsberry, 1982: A statistical climatological tropical cyclone track prediction technique using an EOF representation of the synoptic forcing. *Mon. Wea. Rev.*, **110,** 1945-1954.

Xue, Z. and C. J. Neumann, 1984: Frequency and motion of western north Pacific tropical cyclones. *NOAA Tech. Memo.* NWSNHC 23, National Hurricane Center, Miami, 89pp.

Fractal statistics of large-scale atmospheric circulation.

R.V.Bekryaev (The Arctic and Antarctic Research Institute, 38, Bering st., St.Petersburg, 199397, Russia)
telefax: (812)-352-2688.

1. Introduction.

Time series characterizing the large-scale atmospheric dynamics display fractal properties very often. Apparently the observed self-similar structure in the time sequences of the meteorological characteristics is of dual nature. The presence of fractality may be determined by Kantor properties of an attracting set of an ideal circulation atmospheric model. On the other hand the time series self-affinity which is not directly connected with the deterministic chaos appearing in the system can be caused by the persistent character of the processes taking place. The determing of the fractality causes is of importance as far as diagnosis and prognosis of climatic oscillations are concerned.

The fractal structure of weather and climate attractors was considered in the studies of Fraedrich K., 1986, Essex C. et al, 1987 and some others. The estimate of the attracted set dimension m was carried out using Grassberger-Procaccia procedure (PGP) (1983). This computational procedure permits to estimate the attractor dimension using the time series just of one of the phase space coordinates. Let's consider a mode of some initial dynamical system; $x_1, x_2, \ldots x_n$ being its successive values taken at equal time interval t. We can compose a dynamical system of the dimension m having assumed that $Y_i^{(m)} = \{ x_i, x_{i+1}, \ldots, x_{i+m-1} \}$ for the i- th value of the $Y^{(m)}$ vector of the point position in a new phase space. To estimate the attractor fractal dimension we define the correlation index ν as

$$\nu = \lim_{\varepsilon \to 0} \frac{\ln [C_m(\varepsilon)]}{\ln (\varepsilon)} , \qquad (1)$$

where $C_m(\varepsilon)$ is the correlation integral,

$$C_m(\varepsilon) = \lim_{N \to \infty} \sum_{i,j=1}^{N} \frac{1}{N^2} \mathrm{Hev} \left[\varepsilon - | Y_i^{(m)} - Y_j^{(m)} | \right] ; \quad (2)$$

$Hev(x) = \begin{cases} 1 \text{ with } x \geq 1, \\ 0 \text{ with } x < 1; \end{cases}$, N is the number of samples.

2. An effective procedure for the correlation integral computation.

To determine the correlation index ν it is possible to use the formulae (1),(2) immediately. However in this case the amount of calculations will be too large (proportional to $N^2 m$). The equation for the correlation integral $C_m(\varepsilon)$ should be modified in order to obtain more effective computational procedure.

Taking into account the metric properties in the Euclidean space $\mathbb{R}^n$ the equation (2) can be written as

$$C_m(\varepsilon) = \lim_{N\to\infty} \frac{1}{N^2} \left[N + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} Hev \left[\varepsilon - | Y_i - Y_j | \right] \right] . \qquad (3)$$

Let $j = i + k$, then

$$C_m(\varepsilon) = \lim_{N\to\infty} \frac{1}{N^2} \left[N + 2 \sum_{i=1}^{N-1} \sum_{k=1}^{N-i} Hev \left[\varepsilon - | Y_i - Y_{i+k} | \right] \right] . \qquad (4)$$

For subsequent transformation the following lemma is necessary:

LEMMA: Let the vector Y_i be determined for $\forall$ i =1,2,...N , $F(x)$ is an arbitrary function; then if $n \leq N - 1$, the equality

$$\sum_{i=1}^{n} \sum_{k=1}^{n-i+1} F\left[Y_i - Y_{j+k} \right] = \sum_{k=1}^{n} \sum_{i=1}^{n-k+1} F\left[Y_i - Y_{i+k} \right] , \qquad (5)$$

is true, and a one-to-one correspondence between the elements of items sets of the right and left parts takes place.

The proof of the lemma is easily carried out by the mathematical induction method.

The correlation (5) permits to rewrite the equation (4) as

$$C_m(\varepsilon) = \lim_{N\to\infty} \frac{1}{N^2} \left[N + 2 \sum_{k=1}^{N-1} \sum_{i=1}^{N-k} Hev \left[\varepsilon - | Y_i - Y_{i+k} | \right] \right] . \qquad (6)$$

If the metric is given as follows

$$| Y_i^{(m)} - Y_l^{(m)} | = \left[\sum_{j=0}^{m-1} \left(X_{i+j} - X_{l+j} \right)^2 \right]^{0.5} . \qquad (7)$$

the equation (6) may be transformed in to

$$C_m(\varepsilon) = \lim_{N \to \infty} \frac{1}{N^2} \left[N + 2 \sum_{k=1}^{N-1} \sum_{i=1}^{N-k} \mathrm{Hev} \left(\varepsilon^2 - W_{k,i} \right) \right] , \qquad (8)$$

where

$$W_{k,i} = | Y_i^{(m)} - Y_{i+k}^{(m)} |^2 = \sum_{j=0}^{m-1} \left[X_{i+j} - X_{i+k+j} \right]^2 . \qquad (9)$$

Obviously, for $\forall$ k with $i \geq 2$ it is quite true that

$$W_{k,i} = W_{k,i-1} - \left(X_{i-1} - X_{i+k-1} \right)^2 + \left(X_{i+m-1} - X_{i+k+m-1} \right)^2 . \qquad (10)$$

The equations (8)-(10) suggested for the correlation integral computation are much more effective than the initial formula (2). The number of calculation operations does not depend any more on m and is proportional to $N^2/2$.

It is also possible to obtain the analogous formulae for the following metric

$$| Y_i^{(m)} - Y_l^{(m)} | = \sum_{j=0}^{m-1} | X_{i+j} - X_{l+j} | . \qquad (11)$$

3.On the identification of strange attractors using the method of phase randomization.

The Grassberger-Procaccia algorithm is often used to assess fractal dimension of some attractors even if an a priori knowledge of dynamical nature of relevant processes is absent. In case of stochastic processes this approach may lead to ambiguous or incorrect conclusions (Osborne & Provenzale, 1989). To overcome the above uncertainty, Osborne and Provenzale (1989) proposed the method of phase randomization (MFR) aimed to distinguish between deterministic and stochastic processes. Nevertheless, we suggested the MFR to not give an indication to the presence of low-dimensional chaos when the process is nonstationary. Atmospheric and oceanic processes are influenced by external

periodic forcing (e.g. seasonal), which causes periodic response of both - mean values of main climate characteristics and their dispersions. Let's consider a stochastic process $Y(t)=a(t)+X(t)$, where $a(t)$- is a periodic function and $X(t)= N(0,\sigma^2(t))$ is Gaussian white noise. Lets' define a cycle in time $i = \overline{1,4}$, putting $a_i = a_1$; $\sigma_i = \sigma_1$ if $i=1,2$, and $a_i = a_2$; $\sigma_i = \sigma_2$ if $i=3,4$.

In case $N \longrightarrow \infty$ the correlation integral $C_m(\varepsilon)$ tends to the distribution function $F(G)$, where $G = \sum_{l=0}^{m-1} H^2_{i+l,\,j+l}$, $H_{i,j} = X_i - X_j$, m- embedding dimension, $G=\varepsilon^2$ (Bekryaev, 1991).

To adapt the GP algorithm to the nonstationary process under study, let's consider distances only between vectors marked off from the same "months" and assume that $m=4$.

One can see that distributions of G for initial (G_*) and randomized (G_r) processes are different.

$$P(G_*) = \frac{1}{4(\sigma_1^2 - \sigma_2^2)}\left[\exp\left(\frac{-G_*}{4\sigma_1^2}\right) - \exp\left(\frac{-G_*}{4\sigma_2^2}\right)\right], \qquad (12)$$

$$P(G_r) = \frac{1}{4\sigma_r^4}\exp\left(\frac{-G_r}{2\sigma_r^2}\right) G_r, \qquad (13)$$

Here $\Gamma(\alpha)$ is the Euler gamma-function, and $\sigma_r^2 = (\sigma_1^2 + \sigma_2^2)$.

Thus, the difference between initial and randomized series may be caused by a nonstationary character of a process and not be connected with strange attractor existence. The trends also appear to lead to such differences.

Bekryaev R.V. Some features of a climatic attractor's structure. Izv.Acad.Sci. USSR, Atmos.Ocean.Phys., 1991, v.27, pp.1130-1139, (in Russian).

Grassberger P., Procaccia J. Characterization of Strange Attractors. - Phis.Rev.Lett., 1983, v.50, N 5, pp.346-349.

Essex C., Lookman T.,Nerenberg M.A.H. The climate attractor over short timescales. -Nature, 1987, v.326, pp.64-66.

Fraedrich K. Estimating the dimensions of weather and climate attractors. -J.Atmos.Sci., 1986, v.43, N 5, pp.419-432.

Osborne A.R., Provenzale A. Finite correlation dimension for stochastic systems with power-law spectra. - Physica, 1989, v.35D, N3, pp.357-381.

DEVELOPMENT OF STATISTICAL MODELS OF INTERNAL CLIMATE VARIABILITY

G.V.Alekseyev. The Arctic and Antarctic Research Institute
199226, 38 Bering Str., St.Petersburg, USSR

Statistical models are an integral part of climatology, however, the problem of adequate statistical description of the observed variability of climate characteristics is still topical. This problem is receiving marked attention due to modern theoretical and experimental data and complicated structure of internal non-linear variations in the climatic system [E.Lorenz, 1979].

Two stages may be distinguished in the development of statistical climatology, corresponding to a wide field of application of various fundamental statistical models, adopted from mathematical statistics and theory of stationary random processes. Different understanding of internal (weather) variability in climatic system, different climate definitions and operators for the estimation of climate characteristics correspond to each of the stages. Therefore the estimates of climate characteristics and interpretation of their changes depend upon the operator qualities as well as on the hypothesis for the structure of internal variability in the system. A heavy weather noise in climatic series is attributed to the discrepancy between the estimates and incomplete adequacy of fundamental models and real weather variability.

The third modern stage in the development of statistical models of climate variability is associated with the fundamental models, taking into account the features of internal fluctuations in a non-linear and very unstable system [G.Alekseyev,1991]. By means of this model, including individual abnormal weather structures into a stochastic process, the formation of low-frequency, narrow-band and violet components of weather noise in climatic data series are investigated. Estimates of weather noise contribution depending on season and region have been obtained. Displays of low-frequency weather noise in global climate variations have been revealed.

REFERENCES

1. E.N.Lorentz. Forced and free variation of weather and climate. Journ. of Atm Sci.,1979,vol.36, N 8, p. 1367-76.
2. G.V.Alekseyev,P.N.Svjascennikov. Natural variability of climate characteristics in the northern polar region and northern hemisphere. L.,Gidrometeoizdat, 1991, 159 p.

Spatial and Temporal Variations in Fire Climate

Bernard N. Meisner
Forest Fire Laboratory
Pacific Southwest Research Station
USDA Forest Service
4955 Canyon Crest Drive
Riverside, CA 92507-6099
U.S.A.

1. Introduction

The research unit studying Meteorology for Fire Severity Forecasting at the Pacific Southwest Research Station of the USDA Forest Service has been developing techniques for making monthly and seasonal predictions of potential fire severity for the contiguous United States and Alaska (Fujioka, et al. 1991). Part of this research involves identifying the magnitude, duration and spatial variations in those weather elements (temperature, precipitation, humidity and wind) that are directly related to fire severity. We are also investigating whether any significant secular trends in fire potential are occurring.

2. Methods

The Historical Climatology Network (HCN) (Karl et al. 1990) serial temperature and precipitation data provide an unique opportunity to study spatial, as well as temporal, variations in fire climate for the contiguous United States. The HCN data have been adjusted for any changes in station location, instrumentation, or time of observation (Karl et al. 1990). Data for about 1200 HCN stations over the period 1920-1987 formed the basis for this study (Fig. 1). Fire climate was estimated using the joint distribution of monthly mean maximum temperature (T) and total precipitation (P).

The T and P data were first separately transformed into cumulative percentiles for each station and month. Since the combination of above normal T and below normal P result in high fire potential, the data were sorted so the largest percentiles were assigned to the *highest* temperatures and *lowest* precipitation totals. The data were then interpolated to a latitude and longitude grid using a technique developed by Akima (1978). This technique fits a smooth surface to the data with local fifth degree polynomials such that the surface exactly fits all the observations.

In order to determine whether there has been a secular trend in the areal fraction of the country experiencing climate conducive to high fire danger, a time series of the number of 0.5 degree latitude by 0.5 degree longitude grid cells having concurrent warm and dry departures from normal during each summer (June-August) was computed (Fig. 2).

A principal component (PC) analysis of the correlation matrix of a 2.5 x 2.5 degree gridded joint distribution of T and P percentiles was used to reveal the spatial modes of the large-scale variations in fire climate for the United States. The data were analyzed using the typical S-mode decomposition (Cattell, 1952). The 302 grid cells (151 for T and 151 for P) represented the variables and each of the 816 months represented one observation.

3. Results

a. Secular Trends in Fire Climate

The time series of the number of grid cells experiencing concurrent warm and dry weather (Fig. 2) indicated no appreciable secular trend in fire climate. An average of 32% of the country experienced both above median T and below median P in any given summer during the study period. With the exceptions of 1980 and 1982, the fire climate of the most recent decade was near normal. The areal extent of warm and dry weather during the summer of 1980 (45%) was greater than any year since 1936, while that during 1982 (16%) was the fourth lowest on record. A similar analysis using other thresholds, such as the upper quartile of T and lower quartile of P, or the upper decile of T and lower decile of P, gave similar results (Fig. 2).

b. Principal Component Analysis of Fire Climate

The first 12 eigenvalues of the correlation matrix of the joint distribution of T and P percentiles on a 2.5 x 2.5 degree grid were significant (North et al. 1982) and explained 56% of the variance.

Plotted maps of the loadings on the unrotated PC's (not shown) revealed the typical domain shape dependence reported by Buell (1979) and Richman (1986). In order to accurately identify the spatial modes of joint T and P variation, the first 12 PC's were obliquely rotated using the Promax procedure (K=2) recommended by Richman (1986).

Maps of rotated PC loadings revealed that there are large contiguous regions of country that behave similarly with regard to T or P. Space limitations allow for only one example of each of the three predominant loading patterns to be shown here (Fig. 3). Four of the PC's (3, 4, 5 and 12) had large loadings from both the T and P percentile grids (e.g. Fig. 3b). Each had regions with positive loadings by both T and P, indicating concurrent warm and dry, or cool and wet, weather. Three other PC's (1, 2 and 11) had large T loadings only (e.g. Fig. 3a), while nearly half of the PC's had large P loadings only (e.g. Fig. 3c). Only six of the 24 maps contained both large positive and large negative loadings indicating that, for example, one region of the country would be dry, while another would simultaneously be wet.

4. Discussion

a. Secular Trends in Fire Climate

The time series of areal coverage by concurrent warm and dry weather revealed no long term trend in fire climate. This result agrees with that of Karl et al. (1991) who noted that most of the temperature increase over the past century has occurred in the minimum temperatures, rather than in the maxima that were used in this analysis. Fire potential is more closely related to afternoon maximum, than morning minimum teperature. Should any future warming be similar to that which has been observed in this century, with most of the increase occurring in the minimum temperatures, the effect on future fire potential would be less than if the maximum temperatures increase.

b. Principal Component Analysis of Fire Climate

Since only four of the PC's in this analysis had both large T and P loadings, many of the large-scale variations in T and P would not result in high fire potential. The regions of the country that do have large-scale, joint positive loadings, as indicated by the PC analysis, are 1) the extreme southeast (third PC; 12% explained variance), 2) Texas and Oklahoma (fourth PC; 10%, Fig. 3b), 3) the central Great Plains (fifth PC; 8%), and 4) the central Rocky Mountains (twelfth PC; 5%). Because large-scale variations are the most predictable, these regions may have the most promise for successful prediction of monthly and seasonal fire potential.

In this study, T and P were given equal weight in determining fire potential. Future research will seek to determine a more representative interrelationship among these parameters and fire potential. The temporal evolution of the fire climate patterns will also be explored using extended empirical orthogonal analysis.

REFERENCES

Akima, H., 1978: A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points. ACM Trans. Math. Software, 4, 148-159.

Buell, C. E., 1979: On the physical interpretation of empirical orthogonal function. Preprints, Sixth Conf. on Prob. and Stats. in Atmospheric Sci., Banff, Alta. Amer. Meteor. Soc., 112-117.

Cattell, R. B., 1952; Factor Analysis; An Introduction and Manual for the Psychologist and Social Scientist. Harper and Row, New York, NY.

Fujioka, F. M., B. N. Meisner, M. H. McCutchan and J. W. Benoit, 1991: A process for forecasting monthly fire weather. In Proc. 11th Conf. Fire and Forest Meteor., Missoula, MT, April 16-19, 1991. Soc. Am. For., Bethesda, MD. 526-535.

Karl, T. R., C.N. Williams, Jr., F. T. Quinlan, and T. A. Boden.,1990. United States Historical Climatology Network (HCN) serial T and P data. Oak Ridge National Laboratory Environmental Sciences Division Publication No. 3404.

Karl, T. R., G. Kukla, V. N. Razuvayev and M. J. Changery, 1991: Global warming - evidence for asymmetric diurnal T change. Geophys. Res. Letters, 18, 2253-2256.

North, G. R., 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.

Richman, M. B., 1986: Rotation of principal components. J. Climatology, 6, 293-335.

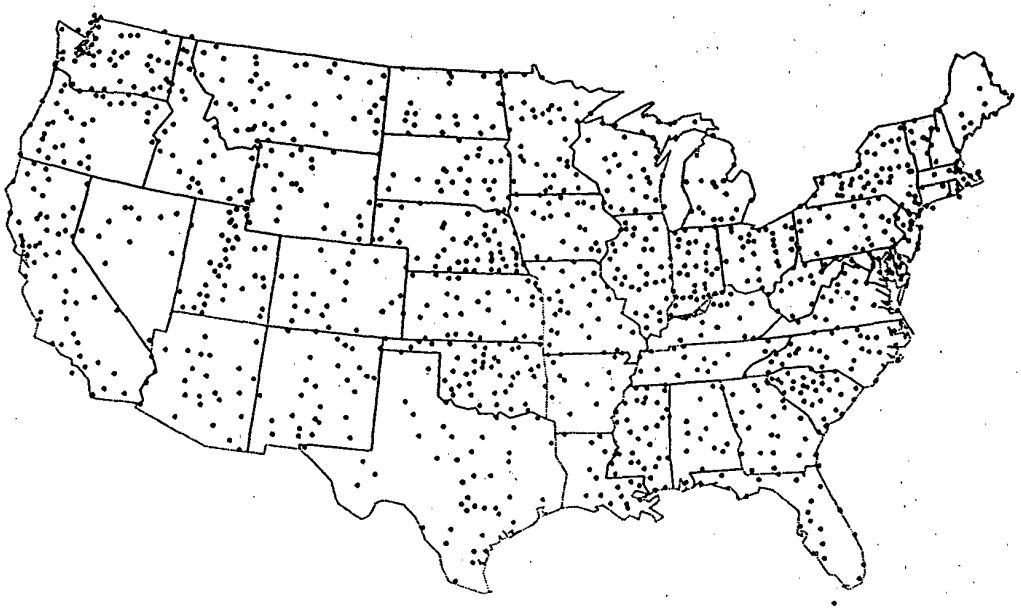


Figure 1. Locations of Historical Climatology Network stations used in this study.

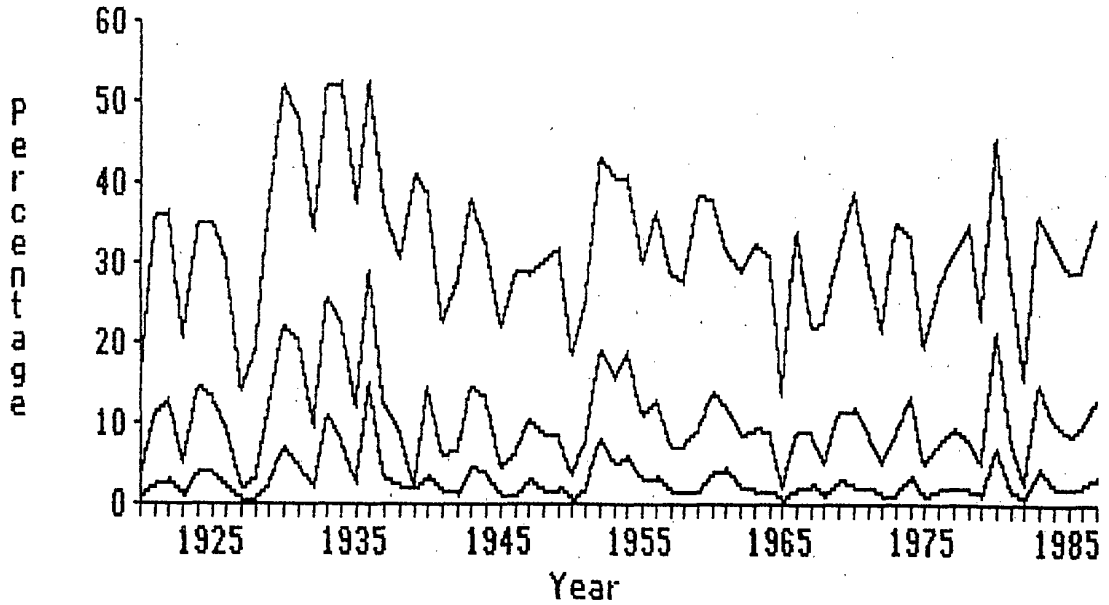


Figure 2. Time series of the percentage of grid cells covering the contiguous United States having concurrent warm and dry departures from normal during summer (June-August). Thresholds used were (top line) above median temperature and below median precipitation, (middle line) above upper quartile of temperature and below lower quartile of precipitation, and (bottom line) above upper decile of temperature and below lower decile of precipitation.

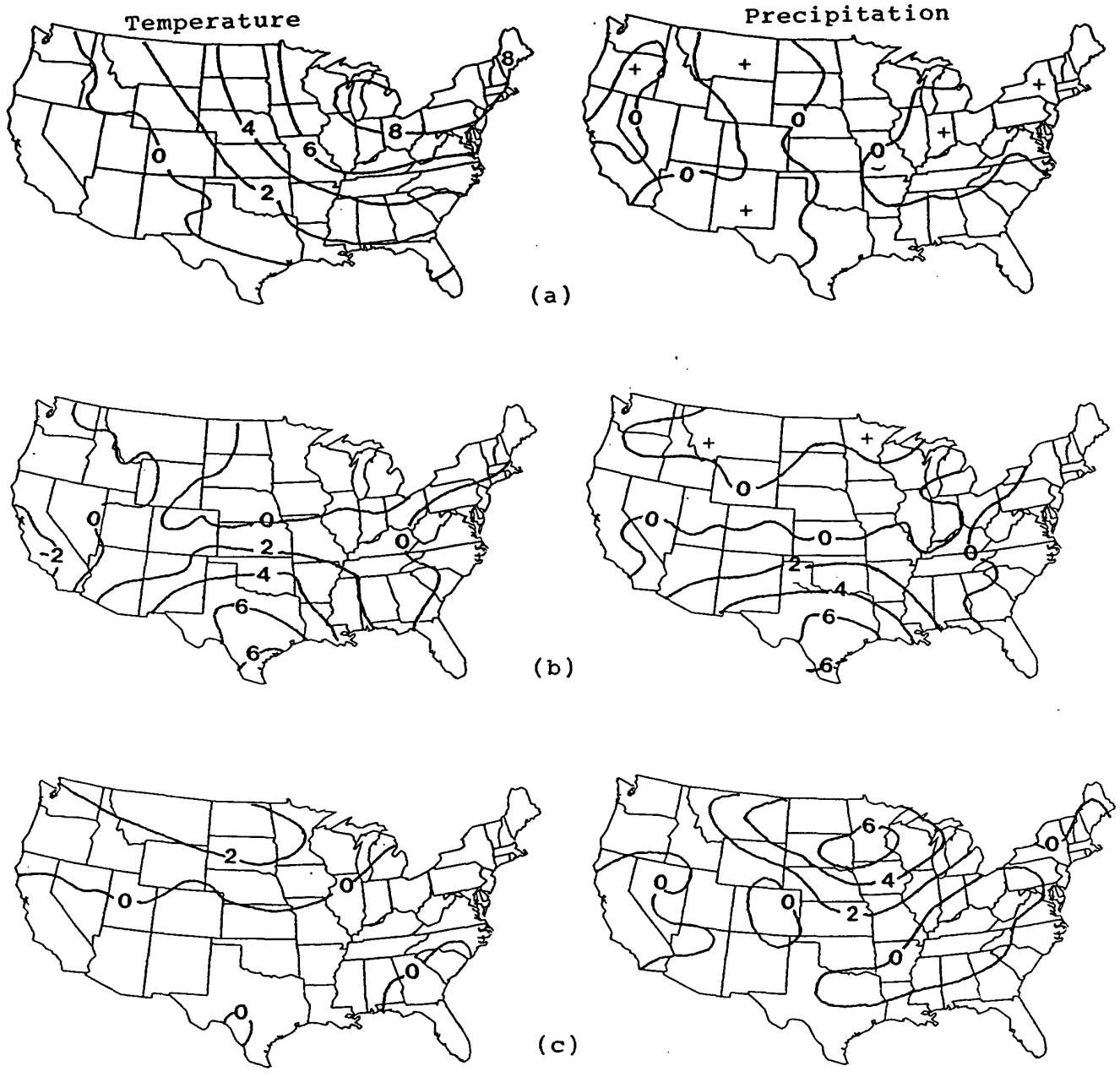


Figure 3. Temperature and precipitation loadings (x 10) on (a) first, (b) fourth, and (c) sixth obliquely rotated principal components of the joint distribution of temperature and precipitation monthly cumulative percentiles.

Seasonal Patterns of Monthly Rainfall and its Association with Particular Situation

Olga C. Penalba*
Institute of Atmospheric Physics, The University of Arizona

Walter M. Vargas*
Departamento de Ciencias de la Atmósfera, Universidad de Buenos Aires, Argentina

ABSTRACT

This paper examines temporal and spatial annual variability of precipitation for 10 stations in the central-east of Argentina, and its relationship with special situation as the El Niño events and maximum and minimum of precipitation. Spectral analysis of monthly series of rainfall is suggested as a method of regional classification of rainfall regimes. By mapping the variance associated with specific rainfall oscillations, it is shown that the first two harmonics are the most important contributors to the rainfall climatology. The different annual marches of rainfall are represented by the wave that contributes most to the total variance. The annual and semiannual cycles are more frequently in continental stations and in coast stations, it is more important short-period fluctuations. The probability that every station has the same harmonic is studied. Only one year out of 60 has this characteristic and is associated with El Niño. The annual cycle, for the driest years, tends to be dominant over the whole region and the spatial distribution of the predominant fluctuation for the wettest years shows distinctive regional groupings. The El Niño events tend to be associated with the wettest years.

1. Introduction

For the development of climatic models, it is very important to know the precipitation pattern and distribution and how to predict them. Therefore, it is necessary to study its spatial and temporal structures for a long record.

The study region, located in the Argentine humid pampa, represents one of the regions in the world with greatest agricultural and cattle possibilities. These activities depend on the regional precipitation variability, among other properties of this parameter. The region is flat, with annual rainfall varying from 600 to 1000 mm and monthly precipitation varying from 50 to 150 mm; both decreasing towards the southwest direction, without significant singularities due to geographic components (Hoffmann, 1975).

This research examines temporal and spatial oscillations of the annual march of rainfall. Spatial patterns, their frequencies and the relationship between these patterns and special situation as El Niño, maximum and minimum of precipitation, are studied.

Several authors have analyzed rainfall variability in different regions of the world, particularly in the Southern Hemisphere. Pittock (1980) studied the spatial patterns of year-to-year variations from the mean precipitation patterns and their relationship to variations in the atmospheric circulation. Caviedes (1981) used harmonic analysis to study how rain occurrence in one region in South America corresponds with rainfall or lack of it in another region. Krepper et al. (1989) looked at time and space variability of precipitation in central-east Argentina using both empirical orthogonal functions and spectral techniques. Penalba and Vargas (1992) studied the regional homogeneity of precipitation fields in the same region of this research, analyzing different statistical properties.

Different analytical techniques have been used to quantify and simplify the description of meteorological events. Principal component techniques have been frequently used in the analysis of precipitation patterns. A more objective approach to describing precipitation seasonality is harmonic analysis. This technique has been used by several authors as an objective way to study the temporal and spatial variations of precipitation. One of the first to use it to depict the patterns of precipitation over the United States were Horn and Bryson (1960). Subsequent applications

in rainfall over South Africa (Mc Gee and Hastenrath, 1966; Keen and Tyson, 1973), Australia (Fitzpatrick, 1964), different regions over United States (Scott and Shulman, 1979; Finkelstein and Truppi, 1991; Kirkyla and Hameed, 1989), South America (Caviedes, 1981) and on a global scale (Hsu and Wallace, 1976) attest to the usefulness of the technique.

2. Data

The available data consists of monthly rainfall totals for the 60-year record (1907–1966) at 10 measuring stations in the humid region of Argentina (locations given in Figure 1). The minimum time period required was 40 years, to assure stability in statistical estimations (Vargas and Penalba, 1985). Among the surface observation system of the National Meteorological Service stations having a complete data records were selected.

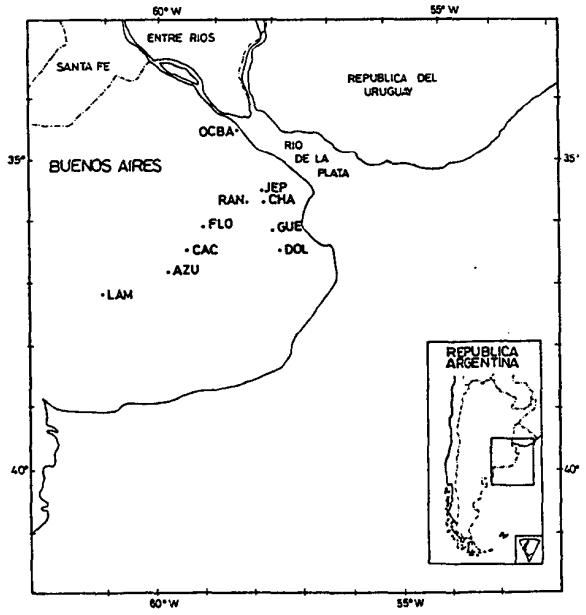


Fig.1. Location of the rainfall stations in the province of Buenos Aires.

3. Methodology and Results

Harmonic analysis of monthly rainfall series is suggested as a method of regional classification of rainfall regimen. The mathematical development follows a text by Panofsky and Brier (1963).

Monthly precipitation data for individual years are averaged to obtain monthly means for the 12 months of the year. These monthly means are subjected to harmonic analysis, which yield the amplitudes, phases and percentages of variance associated with the six harmonics. The results show that the annual and semiannual cycles are the most important contributors to the observed precipitation (almost 90% total variance). Variance associated with annual and semiannual oscillations of mean monthly rainfalls are plotted (Figures 2 and 3). The 12-month wave shows that areas of highest concentrations of explained variance (greater than 70%) occur in the continental stations. The 6-month rainfall oscillation shows more prominence along the coast (variances explained 30% to 40%). Maps of the other waves (harmonics 3 to 6) show no obvious patterns.

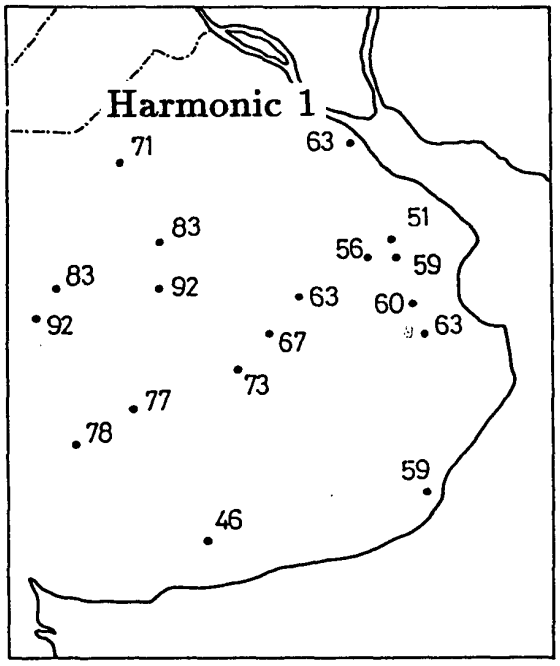


Fig.2. Variance associated with annual oscillation of mean monthly rainfalls.

The knowledge of the rainfall variation from year to year is essential in solving problems dealing with irrigation, crop yield, drainage, water supply, etc. In order to study this variation, the amount of variance accounted for by the different waves, for each year and station, is calculated. The different annual marches of rainfall are represented by

the wave that contribute most to the total variance. The results show that the six possible harmonics appear at least one year, showing a great variability of the monthly rainfall in the study region (Table I). The annual cycle is more frequently in continental stations and the six-month oscillation is more important in coast stations. In spite of the seasonal cycle constitutes an important temporal variation of the atmosphere, this fluctuation is not predominant in the region.

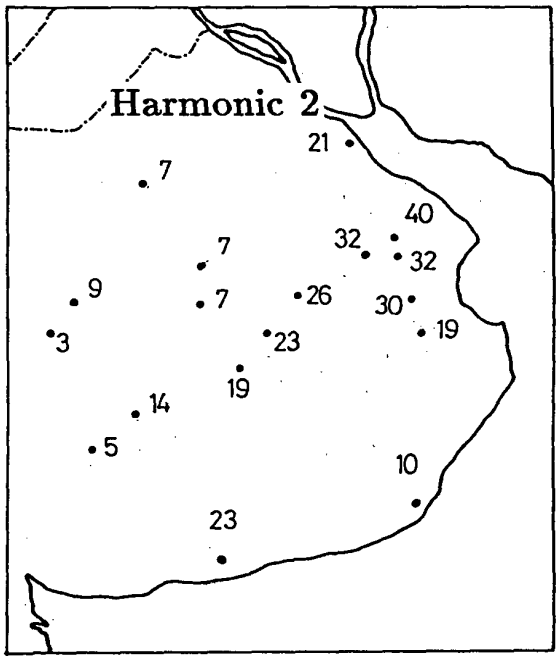


Fig.3. Variance associated with semiannual oscillation of mean monthly rainfalls.

The probability that every station have the same harmonic is studied. Only one year out of 60 has this characteristic and it is associated with El Niño. Due to this result, the probability that two stations have the same harmonic is calculated. Two reference stations (one in the south and the other in the north) are used to establish whether the result depends on the direction of the vector connecting the stations. The different harmonic distributions are plotted for each reference station. The probability is low and no consistent differences between reference stations are found.

The question of whether or not the seasonality of precipitation is influenced by variations in the annual amount of precipitation and/or the El Niño events is investigated. The twelve driest and wettest years and the years with the El Niño episodes from the records, for each station are selected, averaged separately and calculated harmonic analysis. The most important harmonic is plotted for these groups of years. These figures are compared between themselves. The annual cycle, for both the driest years and the El Niño events, tends to be dominant over the whole region while the spatial distribution of the predominant fluctuation for the wettest years shows distinctive regional groupings. To study if these groups (driest, wettest and El Niño) have particular patterns, the most important harmonic is plotted for each year and each group, separately. These figures are compared among themselves. No patterns emerged and no consistent differences are apparent. The probability that the driest and wettest years were the El Niño episode is calculated for each station. These events tend to be associated with the wettest years (50 to 70%).

4. References

Caviedes, C., 1981: Rainfall in South America. Seasonal Trends and Spatial Correlations, *Erdkunde*, Born, **35**, 107-118.

Finkelstein P. L. and L. E. Truppi, 1991: Spatial Distribution of Precipitation Seasonality in the United States, *Jour. Climate*, **4**, 373- 385.

Fitzpatrick, E. A., 1964: Seasonal Distribution of Rainfall in Australia Analyzed by Fourier Methods, *Arch. Met. Geoph. Biokl., B*, **13** , 270-285.

Hoffmann, J. A., 1975: Atlas Climático de América del Sur, *OMM* .

Horn, L. H., and R. A. Bryson, 1960: Harmonic Analysis of the Annual March of Precipitation over the United Stated of America, *Ann. Assoc. Am. Geogr.*, **50**, 157-171.

Hsu C. P., and J. M. Wallace, 1976: The Global Distribution of the Annual and Semiannual Cycles in Precipitation, *Mon. Wea. Rev.*, **104**, 9, 1093-1101.

Keen, C. S. and P. D. Tyson, 1973: Seasonality of South African Rainfall: a note on its Regional Delimitation using Spectral

Table I

The Relative Frequencies of the wave (group) that contribute most to the total variance for the diferrent annual marchs of rainfall

GROUP	OCBA	JEP	RAN	CHA	FLO	GUE	CAC	DOL	AZU	LAM
1	26.6	20.0	23.3	16.7	25.0	13.3	25.0	23.3	33.3	43.3
2	18.3	28.3	21.7	23.3	31.7	33.3	25.0	30.0	30.0	25.0
3	1.7	11.7	10.0	10.0	10.0	10.0	10.0	10.0	6.7	6.6
4	25.0	20.0	16.7	18.3	18.3	16.7	21.7	16.7	18.3	5.0
5	26.6	16.7	26.6	25.0	10.0	21.7	13.3	18.3	1.0	15.0
6	1.7	3.3	1.7	6.7	5.0	5.0	5.0	1.7	1.7	5.0

Analysis, *Arch. Met. Geoph. Biokl., B*, **21**, 207-214.

Kirkyla, K. I. and S. Hameed, 1989: Harmonic Analysis of the seasonal cycle in precipitation over the United States: A comparison between observations and a general circulation model, *J. Climate*, **2**, 1463-1475.

Krepper, C. M., B. V. Scian and J. O. Pierini, 1989: Time and Space Variability of Rainfall in Central-East Argentina, *J. Climate*, **2**, 39-47.

Mc Gee, O. S. and S. Hastenrath, 1966: Harmonic Analysis of Rainfall over South Africa, *Notos*, **15**, 79-88.

Panofsky, H. A., and G. W. Brier, 1963: *Some Applications of Statistics to Meteorology*, The Pennsylvania State University Park.

Penalba O. C. and Vargas W. M.,1992: Study of homogeneity of precipitation in a humid region in the province of Buenos Aires. *In press. Arch. Met. Geoph. Biokl., A*

Pittock, A. B., 1980: Patterns of Climatic Variation in Argentina and Chile - I, Precipitation, 1931-60, *Mon. Wea. Rew.*, **108**, 1347-1361.

Scott ,C. M. and M. D. Shulman, 1979: An Areal and Temporal Analysis of Precipitation in the Northeastern United States, *Jour. App. Met.*, **18**, 627-633.

Vargas, W. M., y O. C. Penalba, 1985: Análisis de la estabilidad de estimaciones estadísticas en series climáticas de Buenos Aires, *Meteorológica*, **15**, N^0 2.

* *Corresponding authors address:* Lic. Olga C. Penalba, Institute of Atmospheric Physics, Building #81, Tucson, AZ 85721. Dr. Walter M. Vargas, Departamento de Ciencias de la Atmósfera, F.C.E.yN., (1428), Ciudad Universitaria, Buenos Aires, Argentina.

The Frequency Distribution of Synoptic Weather Patterns Affecting South Carolina (1930 through 1990)

Scott F. Sidlow[1], David J. Smith[1], David Changnon[1], Clay Lawson[1]

ABSTRACT

Climatology can be defined as the statistics branch of meteorology. Without a better understanding of the frequency of synoptic scale weather events, the statistics in and of themselves are not as useful as they could be. A frequency analysis of synoptic scale weather patterns can yield information about climate change and its causes and impacts.

Therefore, we have performed a long-range time-series analysis of the various synoptic-scale weather patterns that have dominated the weather type in the region in and around South Carolina using 12Z daily weather surface maps from December 1930 through November 1990. Centers of cyclones and anticyclones outside the region, but having frontal activity associated with them within the region, were counted as being the predominant surface feature. If two or more features were present on any one surface weather map, the surface flow that most dominated the situation over South Carolina determined the category of that day's weather.

The synoptic scale features were divided into four categories:

I. Cyclones
   - A. Gulf
   - B. Atlantic Coast
   - C. Ohio Valley
   - D. Gulf Tropical
   - E. Atlantic Tropical

II. Non-cyclone or non-anticyclone
   - A. Surface Troughs
   - B. Stationary Fronts

III. Anticyclones
   - A. Bermuda
   - B. Continental
   - C. Cold-Air Wedge

IV. Combinations of the above

All patterns are defined by the location of the center of circulation within a specific quadrangle of at least 10° latitude and 10° longitude. Each synoptic pattern is assigned a numerical value. The average frequency of each pattern by month, season and year will be described. A time-series will be drawn to show the variability of each synoptic pattern.

Autocorrelations were calculated using the numerical values associated with each weather type. Using these autocorrelations the persistence of each weather type is calculated by month and season. Also, the probabilities of one weather type following another (or itself) are calculated. Combined with historical surface observations, this dataset will be used to determine the probability of extreme weather events given certain synoptic conditions, in future analyses.

The results of this research will provide a climatology of synoptic patterns that affect South Carolina and provide insight into the persistence of certain weather conditions during various times of the year. Therefore the data presented herein will be of use not only to researchers, but also forecasters and planners.

---

[1] South Carolina Water Resources Commission, 1201 Main Street, Suite 1100, Columbia, South Carolina 29201 USA. Phone (803) 737-0800.

Time Scale Differences in the Spatial Variability of Precipitation on a Mountainous Watershed: A Rotated Principal Components Approach

Gregory L. Johnson and Clayton L. Hanson
USDA-ARS, N.W. Watershed Research Center
800 Park Blvd.
Boise, Idaho 83712-7716

1. Introduction

The spatial distribution of precipitation over a watershed often has important implications for runoff, reservoir recharge and water management. Point measurements of precipitation, usually at low elevations, typically serve as indicators for an entire watershed using some prescribed relationships between the indicator points and the watershed average. Such methods fail to account for spatial variability, which can be significantly different under different sets of meteorological conditions.

There were two objectives of this initial study. First, a statistical means of characterizing homogeneous precipitation regions over a watershed with a dense network of stations and a large gradient in annual precipitation was sought. Second, an examination of differences in region definitions over various time scales and seasons was pursued.

2. Data

The USDA - Agricultural Research Service (ARS) maintains several well instrumented watersheds across the U.S. to study hydrologic processes over a wide range of climates and ecosystems. Precipitation data from the mountainous Reynolds Creek Watershed in southwestern Idaho were used in this study. The watershed covers 234 km^2 and elevation ranges from 1097 m at the outlet (north end) to 2195 m over highest peaks in the southwest. The western boundary ridge has a maximum elevation of 1830 m and the eastern range is approximately 1525 m. Average annual precipitation ranges from 250 mm over lowest elevations of the northeast sector to more than 1100 mm over ridges in the southwest.

From 1968 to 1975 complete, daily precipitation records from 46 locations on the watershed were available. A dual-gauge network consisting of one shielded and one unshielded gauge at each site was established during this period, of which 13 are still in continuous use today. A logarithmic relationship between shielded and unshielded catch was applied to all of the daily data (Hamon 1973). This conversion was found necessary because of severe undercatch by single, unshielded gauges in snowstorms. Approximately 75% of the annual precipitation at high elevation sites falls as snow.

3. Methodology and Analysis

PCA and Rotation

Principal components analysis (PCA) has proven to be a reliable method by which multivariate data can be expressed more simply (Tabachnick and Fidell, 1989). Relationships among variables allows many variables to be expressed as a few orthogonal components which explains nearly as much of the total variance in the dataset as in the original set of variables.

In this study, daily and monthly precipitation totals for each of the 46 stations served as the input data matrix. Data were first analyzed over one of two six-month seasons — winter (November - April) or summer (May - October). Other data transformations aimed at reducing positive skewness were examined, including square root and logarithmic, but none improved the data structure more than the normalization procedure. This may be due to the high percentage of days with small precipitation amounts, especially at lower elevations. Days with no precipitation recorded at any of the 46 stations were deleted prior to performing PCA. This reduced the daily datasets by approximately 20% in winter and nearly 50% in the dry summer periods. There were no months during this 8 year period in which all 46 sites recorded zero precipitation. A correlation matrix, R, was calculated over the network for each of the four periods examined (monthly in winter, monthly in summer, daily in winter, daily in summer). From these correlation matrices eigenvectors and eigenvalues were calculated in the standard fashion: $L = V'RV$, where the main diagonal of L contains the eigenvalues and the columns of V are the eigenvectors. These columns contain station coefficients, which when multiplied by the square root of the appropriate eigenvalues yield principal component station loadings (or, simply component loadings). Loadings in this case are really the correlation between the time series of the PC scores and each station's time series (Richman and Lamb, 1985).

Both station coefficients and component loadings were spatially mapped for the four cases considered, but the isoline analyses lacked definition. This issue has been addressed before, and rotation of the PC's is a workable solution to provide larger gradients (Ashbaugh, et al., 1984; Richman, 1986; Eder et al., 1987). Both classes of rotations, orthogonal and oblique, were performed on these components. Oblique rotations failed to reveal as many clearly defined patterns in the precipitation field as the orthogonal rotations, perhaps due to the lack of orthogonal constraints placed on the vectors under oblique rotation. For this reason, orthogonal rotation was performed, specifying the Varimax criterion.

There are two common approaches used to determine the number of components to retain for rotation. One is to retain all components with eigenvalues ≥ 1.0 and the other is called "Rule N", in which eigenvalues are compared to a distribution of uncorrelated gaussian variables (Preisendorfer and Barnett, 1977; Overland and Preisendorfer, 1982). Results of both of these methods for the four cases are shown in Table 1.

Table 1. PCA analysis and retention of components tests for four cases examined.

	Test		% Total Variance Explained by:	
Case	Eigenvalue ≥1	Rule N	1st Component	All
Winter, daily (1214 obs.)	3	2	78.0	86.0
Winter, monthly (48 obs.)	4	1	81.1	91.1
Summer, daily (835 obs.)	3	2	79.2	87.8
Summer, monthly (48 obs.)	4	1	91.1	96.9

Rule N seemed overly conservative and defined too few components while the eigenvalue ≥1 rule produced a sufficient number of distinct, explainable patterns, so it was selected. The first component in all four of the cases was dominant (Table 1), indicating precipitation producing mechanisms which were watershed-wide. The second, third and fourth components, though, were important for processes operating on smaller scales (sub-watershed runoff, windward vs. leeward snow accumulation, etc.). For this reason, it was necessary to retain enough components to capture these smaller but important processes. By rotating the three or four retained components the variance is divided more evenly, and distinct features are more observable in spatial maps (Figs. 1 and 2).

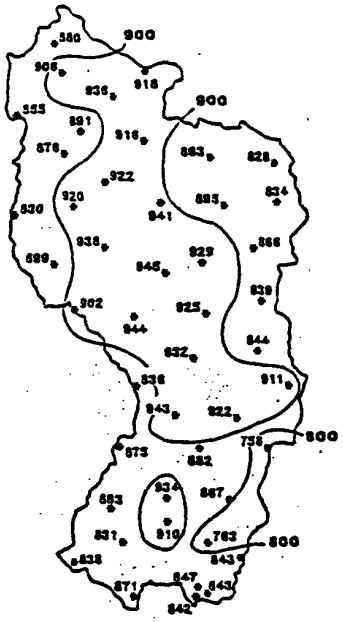


Fig. 1. First principal component station loadings (unrotated); daily, winter case.

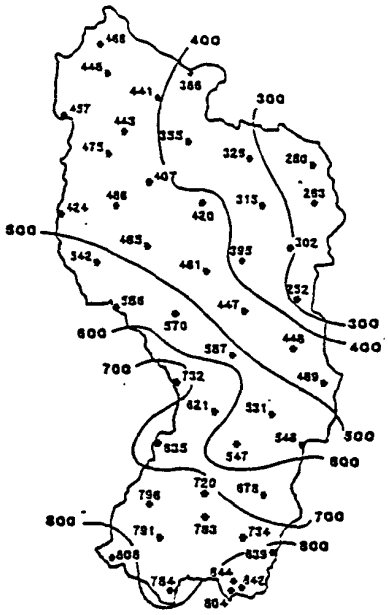


Fig. 2. First principal component station loadings (Varimax rotation); daily, winter case.

Homogeneous Regions

Once a rotated PCA was performed and maps similar to Fig. 1 were produced for each case, homogeneous precipitation regions were defined in a manner set forth by Eder et al. (1987). The component loading with the highest value at each station was plotted and stations were grouped according to these loadings. Homogeneous regions defined by this procedure are shown in Figs. 3-6.

In all cases, the southern highlands and the northeastern lowlands defined two distinct precipitation regions. A northwestern region is fairly well defined in most cases, and seems to be related to east to southeast flow impinging upon these mountains.

Some important differences between daily and monthly maps include valley vs. mountain delineation in winter and the configuration of the northwestern region in summer. Seasonal distinctions include north-south oriented regions in summer vs. east-west oriented regions in winter. This was especially apparent in the monthly, winter case (Fig. 4) where predominant west to southwest flow deposited significant orographic precipitation. A very interesting and distinct region (component 4) appeared in this case composed of just three stations located in the top of the western ridge. It has been observed that although these stations are high in elevation and very much exposed to the wind, precipitation tends to fall just immediately downwind. Another station at approximately the same elevation on this side of the watershed fell within zone 2 (far northwestern section of this component in Fig. 4), most likely because of gauge placement several hundred meters on the lee side of the western ridge. In general, winter regions seemed to be more closely aligned with some of the particular features of the topography (elevation, slope, aspect, windward vs. leewind, etc.) while summer regions were more difficult to

Homogeneous precipitation regions with highest component at each station

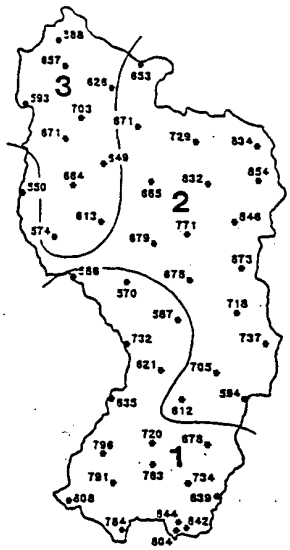


Fig. 3. Daily, winter case.

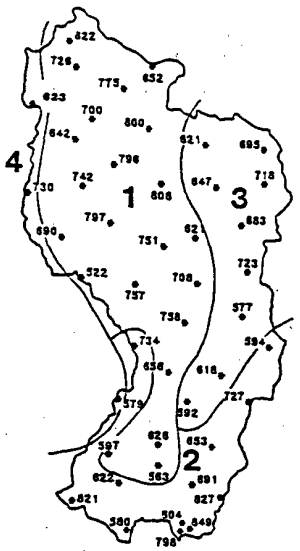


Fig. 4. Monthly, winter case.

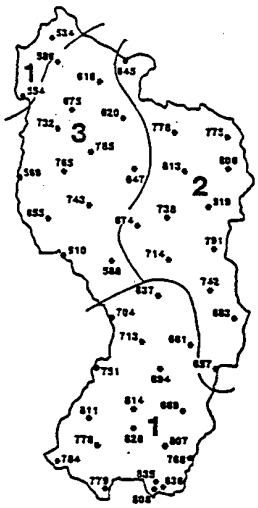


Fig. 5. Daily, summer case.

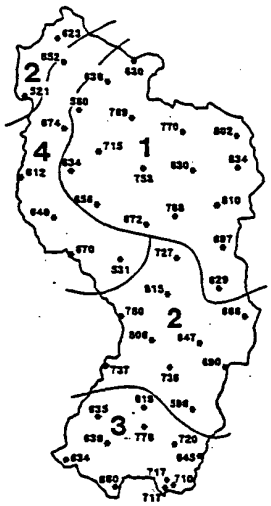


Fig. 6. Monthly, summer case.

explain - perhaps owing to the more convective and localized nature of precipitation during that season.

PC Scores

Principal component scores for each case reveal the time series for the normalized precipitation over each region. The PC score is derived by:

$$(PC\ Score)_{ik} = \sum_{j} O_{ij} L_{jk}$$

(Eder et al., 1987) where O_{ij} is the daily or monthly normalized precipitation for day or month i at station j, and L_{jk} is the loading of station j on component k. In this case, stations were classified according to regions as described above. A separate PCA was then performed only on those stations in each region, and the first rotated PC was used for derivation of factor loadings, L. PC scores were thus derived for each region over the time period of interest.

The four monthly PC scores for winter and summer are plotted in Fig. 7. It should be noted that the PC scores relate to the regions delineated in Figs. 4 and 6, so that PC scores 1-4 from November through April each year relate to regions 1-4 in Fig. 4 and PC scores 1-4 from May to October relate to regions 1-4 in Fig. 6. In most months precipitation was in phase over all regions. January of 1969 was one of the wettest months recorded on the watershed and all PC scores had large positive values. The wet summer of 1972 was wet over all regions, as well. In contrast, region 3 (Fig. 3) was anomalously dry in February of 1970 when other regions were only mildly so, or even slightly wet. December of 1971 was very dry in region 2, while regions 1 and 3 were somewhat wet. No observable trend was noted over the eight year period in any region.

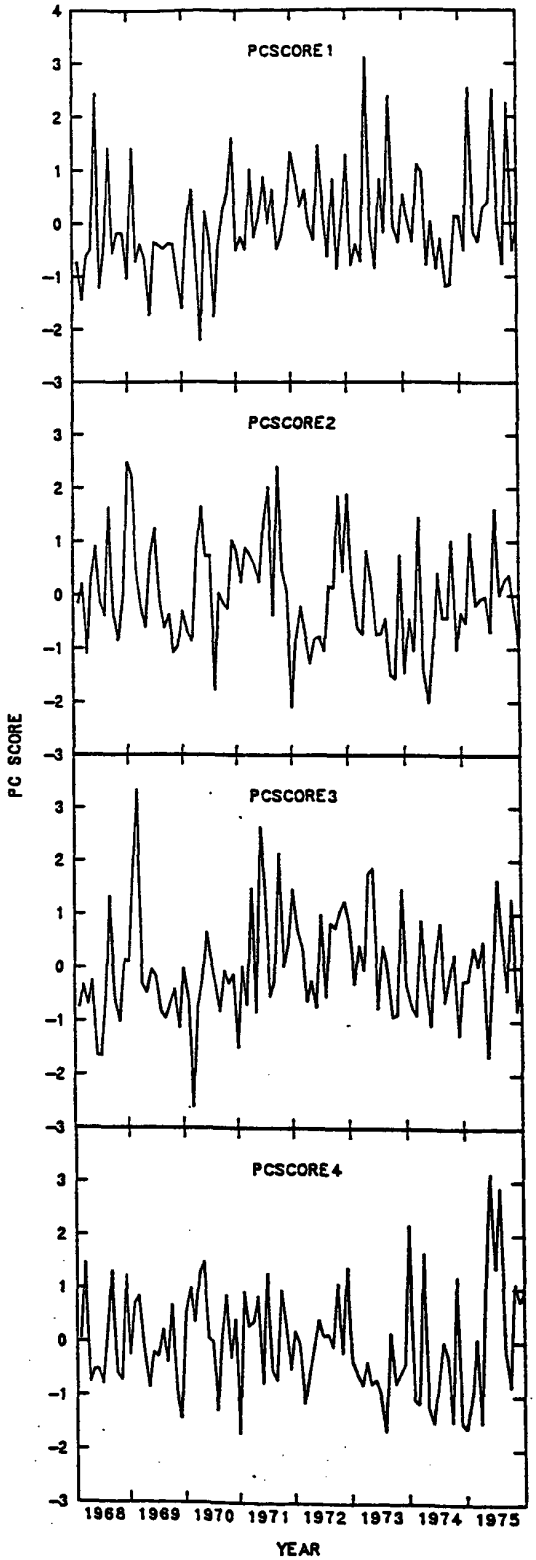


Fig. 7. Monthly principal component scores.

4. Conclusions and Future Work

Rotated principal components analysis has been used successfully to determine homogeneous precipitation regions on a small watershed and to examine differences in precipitation patterns between seasons and between daily and monthly time periods within seasons. Additional work will involve a direct gradient analysis of PC scores with atmospheric dynamics and flow, especially as they are influenced by topography. Upper air data and geographic information systems will be utilized to more thoroughly investigate these relationships.

5. Acknowledgements

Many thanks to Susan Stillings for computing assistance and for preparation of figures.

6. References

Ashbaugh, L.L., L.O. Myrup and R.G. Flocchini, 1984: A principal component analysis of sulfur concentrations in the western United States. Atmos. Environ., 18, 783-791.

Eder, B.K., J.M. Davis and J.F. Monahan, 1987: Spatial and temporal analysis of the palmer drought severity index over the South-eastern United States. J. Climtol., 7, 31-56.

Hamon, R.W., 1973: Computing actual precipitation. Proceedings., WMO–International association of hydrological sciences, Geilo, Norway, 1, 159-174.

Overland, J.E. and R.W. Preisendorfer, 1982: A significance test for principal components applied to a cyclone climatology. Mon. Wea. Rev., 110, 1-4.

Preisendorfer, R.W. and T.P. Barnett, 1977: Significance tests for empirical orthogonal functions. Preprints, Fifth Conf. on Prob. and Stats. in Atmos. Sciences, Amer. Meteor. Soc., 169-172.

Richman, M.B., 1986: Rotation of principal components. J.Climatol., 6, 293-335.

Richman, M.B. and P.J. Lamb, 1985: Climatic pattern analysis of three- and seven-day summer rainfall in the central United States: Some methodological considerations and a regionalization. J. Clim. Appl. Meteor., 24, 1325-1343.

Tabachnick, B.G. and L.S. Fidell, 1989: Using multivariate statistics. HarperCollins, 746 pp.

Signatures of a Universal Spectrum for Atmospheric Interannual Variability in COADS Pressure Time Series over the Indian Ocean Region

A.M. SELVAM, M.K. KULKARNI, J.S. PETHKAR AND R. VIJAYAKUMAR
Indian Institute of Tropical Meteorology, Pune 411 008, India

Introduction

The interannual variability of atmospheric flows as recorded in temperature, wind speed etc., in the total atmospheric column has been investigated extensively and major dominant periodicities such as the quasi-biennial oscillation (QBO) and the El Nino-Southern Oscillation (ENSO) cycle have been identified (Lamb, 1972, Philander 1991). However, such dominant periodicities are superimposed on 'background continuum noise' which also contribute appreciably to the total variance (Lorenz 1990). Predictability of atmospheric interannual variability therefore requires identification and quantification of the physics of multiple scale interactions (Barnett 1991). Multiple scale interactions in atmospheric flows result in long-range spatiotemporal correlations seen as the fractal geometry to the global cloud cover pattern and the inverse power law form for the atmospheric eddy energy spectrum documented by Lovejoy and Schertzer (1986). Such non-local connections are ubiquitous to real world dynamical systems and are recently identified as signatures of self-organized criticality (Bak, Tang and Wiesenfeld 1988). The physics of self-organized criticality in atmospheric flow is not identified. A recently developed cell dynamical system model for atmospheric flows (Mary Selvam 1990, Mary Selvam et al 1992) predicts the observed long-range spatio-temporal correlations as intrinsic to quantum-like mechanics governing atmospheric flows. In this paper it is shown that the model predictions are in agreement with continuous periodogram analysis of seasonal (SON : September to November) mean sea surface pressure (COADS 1985) for the 28-year period 1961-1988 for 225 grid points in the Indian ocean and adjacent regions.

Model predictions

The cell dynamical system model (Mary Selvam 1990, Mary Selvam et al 1992) predictions are as follows. (1) The atmospheric eddy energy spectrum follows the universal inverse power law form of the statistical normal distribution, i.e., the square of the eddy amplitude represents the eddy probability corresponding to the normalised standard deviation t given as

$$t = (\log L / \log T_{50}) - 1 \quad (1)$$

where L is the period length in years and T_{50} the period up to which the cumulative percentage contribution to total variance is equal to 50. (2) The atmospheric flow structure consists of an overall logarithmic spiral trajectory with the quasiperiodic Penrose tiling pattern for the internal structure. The dominant periodicities are therefore expressed in terms of the golden mean [$\Gamma = (1+\sqrt{5})/2$] equal to 1.618 as $tX/\Gamma = 2.2t$, $tX = 3.6t$, $t\,\Gamma X = 5.8t$, $t\,\Gamma^2 X = 9.5t$, $t\,\Gamma^3 X = 15.3t$, $t\,\Gamma^4 X = 24.8$ where $X = 2+\Gamma$ and t, the primary perturbation time period is the annual cycle of solar heating. The above model predicted interannual cycles may be related to the reported (Ghil and Vautard 1991) periodicities as follows. The 2.2 year cycle may correspond

to the high frequency quasi-biennial component of ENSO, while the 3.6 and 5.8 year cycles may refer to the low frequency component of ENSO. 9.5, 15.3 and 24.8 year cycles correspond to interdecadal oscillations.

Data and analysis

The seasonal (SON) mean sea surface pressure for the 28-year period 1961-1988 for 225 grid points in the Indian Ocean and adjacent regions within 50°E to 138°E longitudes and 30°N to 40°S latitudes were used for the study. The broadband power spectra of the sea surface pressure time series were computed by an elementary but very powerful continuous periodgram analysis technique of Jenkinson (1977). The method provides a quasi-continuous form of the classical periodogram, allowing systematic allocation of the total variance and degrees of freedom of the time series to logarithmically spaced elements of the frequency range (0.5, 0). The power spectrum was represented as the percentage contribution to total variance corresponding to the respective t values for the periodicities (Eq. 1). The statistical normal distribution represented as the percentage probability corresponding to the normalised standard deviation t values were then compared with the power spectrum and the "goodness of fit" was tested using the statistical chi-square at 95% confidence level.

The power spectrum for a representative grid point is plotted in Figure 1 (left) as the cumulative percentage contribution to total variance versus the corresponding t values as continuous line. The statistical cumulative normal percentage probability distribution is plotted as crosses in Figure 1 (left). The short horizontal line in the lower part of the spectrum represents the level above which the spectrum is the same as the normal distribution as determined by the chi-square test.

The diagram on the right side of Figure 1 gives the plot of normalised variance (continuous line) and phase (crosses) for periodicities 2 to 25 years.

The results of periodogram analyses of sea surface pressure time series for 225 grid points is summarised in Table 1 to give the total number of spectra which exhibit the following model predicted universal characteristics: (1) normal distribution, (2) $T_{50} < 5$ years where T_{50} is the period up to which the cumulative percentage contribution to total variance is equal to 50, (3) dominant periodicities (years) in the period ranges, 2-3, 3-4, 4-8, 8-12, 12-20 and 20-28 years. The above period ranges were chosen with reference to model predicted intrinsic periodicities 2.2, 3.6, 5.8, 9.5, 15.3 and 24.8 years.

Discussion and conclusion

The power spectra of interannual variability of sea surface pressure are found to follow the universal inverse power law form of the statistical normal distribution as seen from Table 1 and the representative spectrum at Fig. 1 (left). The above results is consistent with model prediction of quantum-like mechanics governing atmospheric flows since the square of the eddy amplitude, namely, the variance represents the probability density. Inverse power law form for power spectra imply long-range temporal correlations (or persistence) and are signatures of self-organized criticality. Universal quantification for self-organized criticality in terms of the statistical normal distribution implies predictability

of the total pattern of fluctuations namely, the interannual variability of sea surface pressure in this study. It is seen that the broad-band eddy continuum at Fig. 1 (right) consists of several dominant wavebands, the bandwidth increasing with period length. Further, there is a progressive rotation of the phase angle with increasing period in each waveband indicating a continuous spiral-like structure for the eddy continuum. The above result is in agreement with model prediction of logarithmic spiral structure for atmospheric flow pattern.

Periodicities up to 5 years alone contribute as much as 50 percent of the total variance (see Table 1, $T_{50} < 5$ years). Short period (up to 5 years) dominant periodicities such as the ENSO therefore determine to a large extent the near future sea surface pressure pattern. Also, relatively recent and short period time series such as the 28-year period (1961-1988) used in this study are sufficient to identify the universal structure of atmospheric interannual variability.

The model predicted "magic numbers" for intrinsic robust circulation time periods are found to be present in almost all the cases.

The model predicted high and low frequency components of ENSO centred respectively around 2.2 years and 3-8 years are present in almost all the time series (Table 1). The model predicted interdecadal oscillations 8-28 years are also present in about one third of the cases.

The important conclusions of the present study are as follows. The ENSO interdecadal oscillations are intrinsic to atmospheric flows and are generated by the annual cycle of solar heating. Universal spectrum for interannual variability of atmospheric flows rules out linear trends i.e. prolonged destructive weather patterns such as droughts or floods. Persistent energy input to atmospheric circulation from greenhouse gas warming or solar variability may result in intensification of internal circulation structure, namely, the high frequency components of ENSO.

Acknowledgement

The authors are grateful to Dr.A.S.R. Murty for his keen interest and encouragement during the course of this study.

Figure 1

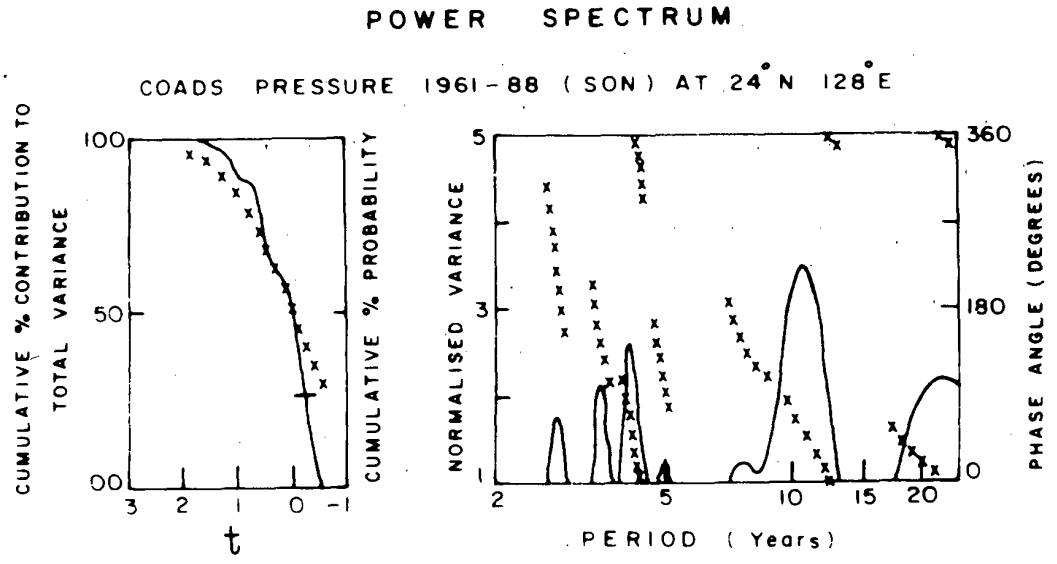


Table 1 : Results of periodogram estimates for 225 grid points

spectra same as normal distribution	spectra with T_{50} <5 years	spectra with dominant peak periodicities (years) in the range 2-3	3-4	4-8	8-12	12-20	20-28
220	190	223	187	219	72	93	55

References

Barnett, T.P., 1991 : The interaction of multiple time scales in the tropical climate system. *J. Climate* **4**, 269-285.

Bak, P., C. Tang and K. Wiesenfeld, 1988 : Self-organized criticality. *Phys. Rev. A*, **38**, 364-374.

COADS, 1985 : Comprehensive Ocean Atmosphere Data Set Release 1, NOAA, Boulder, Co., USA.

Ghil, M., and R. Vautard, 1991: Interdecadal oscillations and the warming trend in global temperature time series, *Nature* **350**, 324-327.

Jenkinson, A.F. 1977 : *A powerful elementary method of spectral analysis for use with monthly, seasonal or annual meteorological time series.* (U.K. Meteorol. Office, London) Met. O 13 Branch Memorandum No. 57, 1-23.

Lamb, H.H., 1972 : Climate : present, past and future, vol. I. Fundamentals and climate now. (Methuen and Co., Ltd., London) pp. 613.

Lorenz, E.N., 1990 : Can chaos and intransitivity lead to interannual variability? *Tellus* **42A** 378-389.

Lovejoy, S., and D. Schertzer, 1986 : Scale invariance, symmetries, fractals and stochastic simulations of atmospheric phenomena. *Bull. Am. Meteorol. Soc.* **67**, 21-32.

Maddox, J., 1988 : Licence to slang Copenhagen? *Nature* **332**, 581.

Mary Selvam, A., 1990 : Deterministic chaos, fractals and quantum-like mechanics in atmospheric flows. *Can. J. Phys.* **68**, 831-841.

Mary Selvam, A., J.S. Pethkar and M.K. Kulkarni, 1992 : Signatures of a universal spectrum for atmospheric interannual variability in rainfall time series over the Indian region. *Int'l. J. Climatol.* **12**, 137-152.

Philander, S.G., 1990 : *El Nino, La Nina and the Southern Oscillation.* (Academic Press, New York) International Geophysical Series 46, pp. 291.

AGRICULTURAL DROUGHTS IN INDIA

By

A.R.SUBRAMANIAM
Department of Meteorology and Oceanography, Andhra University
Visakhapatnam, India.

ABSTRACT

The paper deals with many approaches available in literature for the delineation of drought, and their limitations. A rational approach to Agriculture drought studies based on Statistical technique is suggested. Further Agricultural droughts have been classified based on departure of index of moisture adequacy from the minimal value corresponding to average agricultural yield crop value. The approach consider different limits for different crops for determination of agricultural droughts. This technique is applied to study of agricultural droughts of the principal crops in different states of India (Kerala, Andhra, Karnataka, Maharashtra, Orissa, M.P. and Rajasthan). The paper deals with the application of the method for determining Agricultural drought intensity, frequency and proneness of different crops in different areas in various states. It is felt this approach will be of great assistance in delineation of Agricultural droughts particularly in dry climates where marginal water supplies are available.

THE EVALUATION OF ROMANIA'S REGIONAL CLIMATIC ANOMALIES BY MEANS OF LINEAR PREDICTION MODELS

Aristiţa Busuioc
National Institute of Meteorology and Hydrology, Bucharest-Romania
Sos.Bucureşti - Ploieşti 97

ABSTRACT

The paper presents a linear prediction model by a multivariate regression method. The analysis of the principal components and the factor analysis are applied to the initial predictor set in view of obtaining a new set of independent variables. The skill of the model and statistical significance are defined, showing the procedure of choosing the optimum model.
Finally, the paper presents the results of the application of a autoregressive model for the prediction of the annual temperature and precipitation anomalies corresponding to various regions in Romania.

1. INTRODUCTION

The prediction of short-time climate variations in the range from a month up to a few years has a very great practical importance. The results of the global climatic models show considerable incertitudes mostly related to the regional predictions (Barnett, T.P., 1981).

The problem of choosing the most adequate predictors is a difficult one. The studies carried out emphasized the fact that the predictive performance of the models increases and the statistical significance generally decreases with the increase of the number of predictors. Therefore, the main problem when developing such models is that of acieving a proper balance between the performance and the statistical significance.

This paper presents a linear prediction model by the multivariate regression method. Paragraph 2 describes the algorythm of the model and the procedure for choosing the optimum model and paragraph 3 presents the results obtained for the prediction seasonal and annual temperature and precipitation anomalies corresponding to different regions in Romania.

2. MODEL DESCRIPTION

If X of an n x p -dimension represents the predictor matrix of n-observations on p-variables, while Y being the predictant variable, then a multiple regression model between Y and X writes as follows:

$$Y = X\beta \quad , \text{ while } \hat{\beta} = (X'X)^{-1} \qquad (1)$$

where $\hat{\beta}$ = the β - unknown coefficient estimation.

In case a set of Y dependent m-variables is being estimated according to the same independent X-variables, then the respective case is reduced to the treatment of the m-multiple regression of the form /1/.(Haan,T.C., 1979).

As in the practical approaches the predictor variables are not always independent, original correlated variables are being therefore transformed into uncorrelated ones or orthogonal principal components as expressed by the relation below:

$$Z = XA \qquad /2/$$

where Z represents the p-orthogonal component matrix, while A contains the eigenvectors of the correlation matrix $R=XX'/(n-1)$ (in this relationship, the elements of matrix X are standardized by using the rela-

tionship $X_{ij}=(X_{ij}-\overline{X}_j)/S_j$, $\overline{X}_j, S_j$-represents the mean and the mean square deviation of variable j).

The fraction of the total variance of X, explained by the j^{th} component is given $\lambda_j / \sum_{i=1}^{p} \lambda_i$ (λ_j - the eigenvalue corresponding to the j^{th} eigenvector). In many situations the first q components justify a great part of the X variance (over 90%) indicating the fact that the last p-q components are not necessary for the explanation of the variance (so they can be eliminated).

The correlation between the i^{th} variable and the j^{th} component is given by the relationship,

$$Cor\ (X_i, Z_j) = \lambda_j^{1/2} \cdot a_{ij} \qquad (3)$$

where a_{ij} is the i^{th} element of the eigenvector a_j. By examining the correlations given by the relationship (3) those variables which are not correlated to any of the components can be eliminated. If a component is highly correlated with several variables, then this component is a reflection of these variables. Therefore, these correlations can be used to associate a physical significance to the components.

The principal components once obtained are used as independent variables in the multiple regression model presented by the relationship (1).

In order to simplify the interpretation of the principal components they are rotated by the "varimax" method (Haan, 1979). The rotated principal components are given by the relationship:

$$Z^* = XR^{-1} L^* D_\lambda^{1/2} \qquad (4)$$

where $D_\lambda^{1/2}$ - diagonal matrix containing the square root of the eigenvalues λ_i, L^* - the rotated matrix of the factor loading matrix $L = AD_\lambda^{1/2}$.

The multiple regression between Y and Z^*, expressed in terms of the original variables X, has the form:

$$Y_i = \overline{Y} + \sum_{j=1}^{p} \beta_j \left\{ \sum_{k=1}^{p} E_{kj} [(X_{ik} - \overline{X}_k)/S_k] \right\} \qquad (5)$$

were E_{kj} are the elements of matrix $E = R^{-1} L\ D_\lambda^{1/2}$.

The skill of the model is defined by the relationship (Barnett, 1981)

$$S_H = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} y_i^2} \qquad (6)$$

and its statistical significance is given by:

$$F = \frac{(\beta' X' Y' - n\,\overline{Y}^2)/(p-1)}{(Y'Y - \hat{\beta}' X' Y')/(n-p)} \qquad (7)$$

which has an F distribution with p-1 and n-p degrees of freedom.

Generally S_H increases with the increase of p order of the model, while its statistical significance decreases. Therefore, the main problem when constructing a model of optimum order, is to achieve a balance between the skill and the statistical significance. This is achieved by constructing a hierarchy of models generated by the successive addition of new predictors.

A set of p predictant variables initially settled, is considered at q time-lags back and by an adequate rewriting of the variables, a new set of predictors is obtained of a number equal to r=pxq (the new order of the model). Varying q, a model class is obtained, the optimum order r of the model being settled for that number of q time - lags back for which the model has a satisfactory significance. In case the same variables X, predicted in K time-lags ahead, are considered are predictants the case is that of an autoregressive model as in this paper.

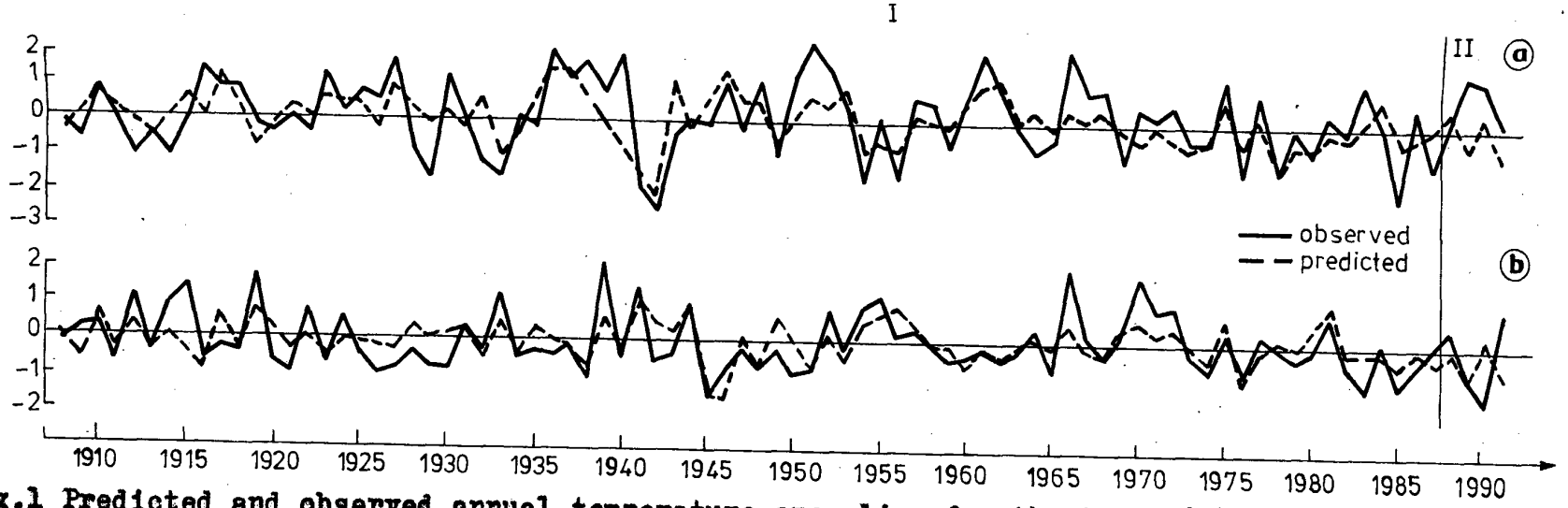


Fig.1 Predicted and observed annual temperature anomalies for the South (a) and annual precipitation anomalies for the South-east (b)

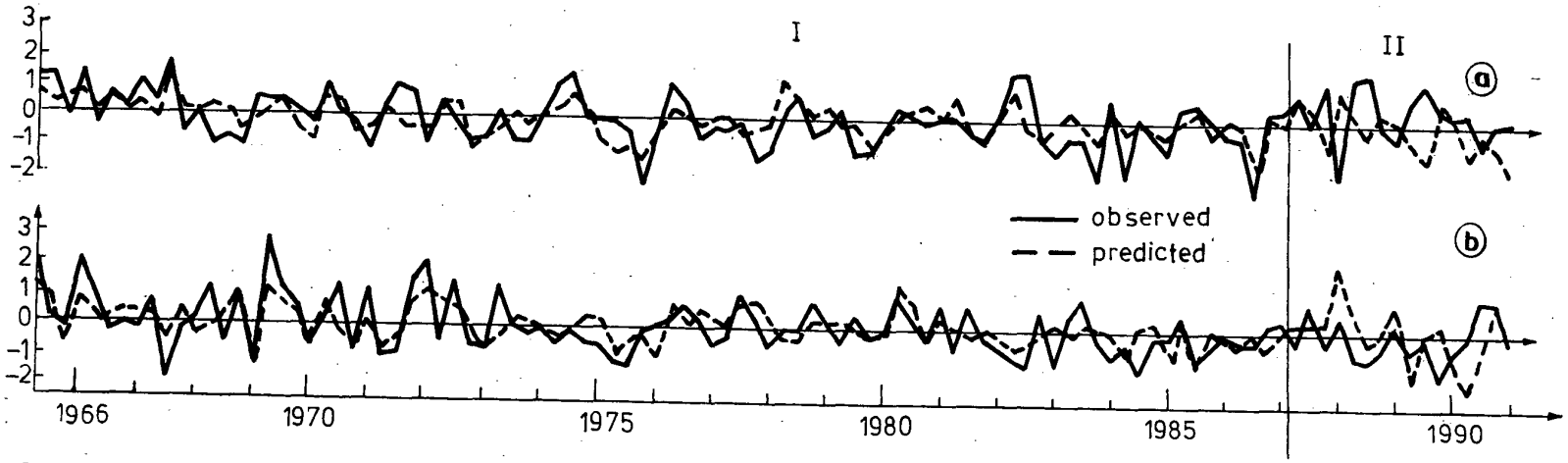


Fig.2 Predicted and observed seasonal temperature anomalies for the South (a) and seasonal precipitation anomalies for the South-east (b)

3. RESULTS OBTAINED

The model presented at paragraph 2 was applied for the prediction of the annual (with one time lag ahead) and seasonal (with 1-4 time - lags ahead) temperature and precipitation anomalies corresponding to various regions of Romania. These anomalies have been calculated by spatial averaging of the values $(X_i-\bar{X})/\sigma$ ($\bar{X}$ - the multiannual mean, σ = the mean square deviation) for various regions over the 1899-1987 interval.

As initial predictors, the temperature and precipitation anomalies were taken for combinations of two of such regions, considered at various time lags back, the optimum model being calculated for each case.

For the prediction of the annual anomalies the optimum model was of order 32 (corresponding to the 4 variables considered at 8 time-lags back) and for the seasonal anomalies, the model of order 36 (corresponding to the same variables considered at 9 time-lags back).

The skill of the model was between 0.4 and 0.6 (the better skill being for the seasonal predictions).

The level of the statistical significance was ranging between 10% and 5%, the best significance level (5%) being for the seasonal predictions.

Fig.1 (a,b) presents as an example the real and the predicted values of the annual temperature(south) and precipitation (south-east) anomalies. Fig.2 (a,b) presents the real and predicted values of the seasonal temperature and precipitation anomalies predicted with 4 time-lags ahead (for the same regions).

In the second part of the graphs the predictions for 1988-1991 are presented by using independent data set.

Although there are certain discordances between the observed and the predicted values, the inter-seasonal (interannual) variation of the predicted values generally show the same evolution as the real ones but with a different amplitude.

Better results for the prediction of the seasonal temperature anomalies (as regarding the statistical significance as well as the skill) have been obtained by means of a regressive model of order 16 using as predictors the SST anomalies from 4 stations in the North Atlantic, considered at 4 time-lags back) (Busuioc, 1992).

4. CONCLUSIONS

The statistical modelling technique presented allows to obtain regional forecasting information useful for the seasonal predictions with various anticipations as well as for annual predictions.

From the analysis of the results obtained differences occur as regarding the optimum order of the model for the prediction for various time scales (year, season). The way of choosing the pairs of initial predictor variables is also important. The more homogeneous from the climatic variability standpoint are the regions considered for the calculation of the anomalies, the better results are.

As it was mentioned, a very important problem is the choice of the proper predictors, among which the ocean variables playing a major part. The results can be improved by considering other sets of predictor variables, physically justified as well as possible.

REFERENCES

1. BARNETT,T.P., (1981) Statistical Prediction of North American Air Temperature from Pacific Predictors.Monthly Weather Review,5,1o21-1o41
2. BUSUIOC,A., (1992) Prediction of Regional Temperature Anomalies in Romania by Using SST Anomalies from the Northern Atlantic.European Geophysical Society XVII General Assembly,Edinburgh(6-1o April), Book of Abstracts, Part II.
3. HAAN,T.C., (1979) Statistical Methods in Hydrology. The IOWA State University Press/Ames.

Modelling the dependence of whitecap on windspeed: hierarchical models, and shrunken parameter estimation

Iognáid Ó Muircheartaigh,
University College Galway, Ireland
and
Edward C. Monahan,
Marine Sciences Institute,
University of Connecticut, U.S.A.

1 Introduction

Gaver and Lehoczky (1987) describe the concept of *hierarchical* stochastic models. A wide variety of stochastic modelling situations arising in pracical applications can be conveniently approached by using such models. These models incorporate two types of variability: variability *within* a unit or population member and variability *between* units. This paper applies the concept to the problem of modelling the dependence on windspeed of oceanic whitecap coverage, and oceanic whitecrest coverage; in the former case eleven, and in the latter case ten, distinct data sets are available. Hierarchical models are utilized to determine the degree of compatibility between the data sets, and to enable estimation of model parameters in a manner which incorporates both types of variability described above.

2 Hierarchical Models

For a full description of hierarchical models see Gaver and Lehoczky (1987). In the present application, we are given k datasets (units), and our observations (within units) consist of measurements of windspeed (u) and whitecap coverage (w). On the basis of previous studies, we transform the data as follows:

$$W = w^{1/3} \quad (1)$$

$$U = u(g\nu)^{-1/3} \quad (2)$$

(where g is the accelaration due to gravity, and ν is the kinematic viscosity of sea water), and apply the following *hierarchical* model to the transformed data:
within dataset variation:

$$W_{ij} = \beta_i U_{ij} + \epsilon_i \quad (3)$$

$$\epsilon_i \sim N(0, \sigma_i^2) \quad (4)$$

for $i = 1, \ldots, k$ and $j = 1, \ldots, n_i$,

and between dataset variation:

$$\beta_i \sim N(\beta, \sigma_\beta^2) \quad (5)$$

for $i = 1, \ldots, k$. This is often referred to as the *superpopulation* from which the individual β_i can be considered drawn at random.
Following standard methodology, the likelihood may be written

$$\mathcal{L} = \prod_{i=1}^{k} \prod_{j=1}^{n_i} \mathcal{L}_{ij} \quad (6)$$

where

$$\mathcal{L}_{ij}(\beta,\sigma_\beta^2) \propto$$

$$\int_{-\infty}^{\infty} \frac{1}{\sigma_i} e^{-\frac{1}{2}\left(\frac{W_{ij}-\beta_i U_{ij}}{\sigma_\beta}\right)^2} \frac{1}{\sigma_\beta} e^{-\frac{1}{2}\left(\frac{\beta_i-\beta}{\sigma_\beta}\right)^2} d\beta_i \quad (7)$$

Letting $l_{ij} = ln(\mathcal{L}_{ij})$, and omitting details, we obtain

$$l_{ij} \propto \frac{\left(U_{ij}W_{ij}\sigma_\beta^2 + \beta\sigma_i^2\right)^2}{\sigma_i^2\sigma_\beta^2\left(U_{ij}^2\sigma_\beta^2 + \sigma_i^2\right)}$$

$$-\frac{\left(W_{ij}\sigma_\beta^2 + \beta^2\right)}{\sigma_i^2\sigma_\beta^2} - ln(U_{ij}^2\sigma_\beta^2 + \sigma_i^2) \quad (8)$$

and hence the loglikelihood may be written

$$l(\beta,\sigma_\beta^2,\sigma_i^2) = \sum_{i=1}^{k}\sum_{j=1}^{n_i} l_{ij} \quad (9)$$

3 Parameter estimation

Parameter estimation for this problem involves three separate but related stages, viz,.

3.1 Empirical Bayes estimation of σ_i^2

Each σ_i, $(i = 1,\ldots,k)$, is estimated by the residual mean square error, $\hat{\sigma}_i^2$, for the regression of W on U with no intercept term included in the model. Following conventional empirical Bayes methodology, these parameters are assumed known for the further stages of the estimation procedure, as described below.

3.2 Maximum likelihood estimation of superpopulation parameters

The likelihood, as described in equation (9), and assuming σ_i^2 known, is a function of two variables β and σ_β^2. This function was numerically maximized using the IMSL statistical package, and confidence regions for the ML estimates $\hat{\beta}$ and $\hat{\sigma}_\beta^2$ were obtained by using the likelihood ratio test procedure which specifies that all (β,σ_β^2) values such that

$$-2\left[ln\left(\frac{\mathcal{L}(\beta,\sigma_\beta^2)}{\mathcal{L}(\hat{\beta},\hat{\sigma}_\beta^2)}\right) \leq \chi^2_{(1-\alpha)}(2 \text{ d.f.})\right] \quad (10)$$

constitute an approximate $100(1-\alpha)\%$ confidence region for (β,σ_β^2). The issue of whether or not all the data sets may have come from populations with identical β's is then addressed by determining whether or not the confidence region derived includes $\sigma_\beta^2 = 0$.

3.3 Empirical Bayes estimation of the individual β_i

From equations (3) and (4), it is easy to show that the posterior probability density function of β_i is

$$f(\beta_i|U_{ij},W_{ij},\sigma_i^2) \sim N(\mu_{ij},\sigma_{ij}^2)$$

for $i = 1,\ldots,k$ and $j = 1,\ldots n_i$, where

$$\mu_{ij} = \frac{\frac{U_{ij}W_{ij}}{\sigma_i^2} + \frac{\beta}{\sigma_\beta^2}}{\frac{U_{ij}^2}{\sigma_i^2} + \frac{1}{\sigma_\beta^2}} \quad (11)$$

$$\sigma_{ij}^2 = \frac{1}{\frac{U_{ij}^2}{\sigma_i^2} + \frac{1}{\sigma_\beta^2}} \quad (12)$$

Accordingly, combining the information in a given data set with the (empirical Bayes) estimates of the superpopulation parameters, it

follows that, if $\mathbf{U_i}$ represents the set of U observations for data set i, and $\mathbf{W_i}$ the corresponding set of W observations, then

$$f(\beta_i|\mathbf{U_i},\mathbf{W_i},\sigma_i^2) \sim N(\hat{\beta},\hat{S}_i^2)$$

where

$$\hat{\beta}_i = \frac{\sum_{j=1}^{n_i} \mu_{ij}/\sigma_{ij}^2}{\sum_{j=1}^{n_i} 1/\sigma_{ij}^2} \quad (13)$$

$$\hat{S}_i^2 = \frac{1}{\sum_{j=1}^{n_i}(1/\sigma_{ij}^2)} \quad (14)$$

$\hat{\beta}_i$ is an individualized (pooled, shrunken) estimate of the parameter β_i for the population from which data set i was selected. It has the very reasonable property that when σ_β^2 is large (i.e. large variations between the data sets) then the contribution of the superpopulation factor (i.e. the data sets other than data set i) is small, and when σ_β^2 is small (i.e. the data sets relatively uniform), the contribution of the superpopulation factor is substantial.

4 Data and analysis

4.1 Data

The models described in section 2 were used to describe the dependence of oceanic whitecap and oceanic whitecrest coverage on windspeed. In the case of whitecaps, eleven distinct datasets (average number of obvervations: 40) and in the case of whitecrests, ten distinct datasets (average number of observations: 44) were available. Due to space limitations, it is not possible in this paper to give a complete description of these datasets, but this can and will be provided on request.

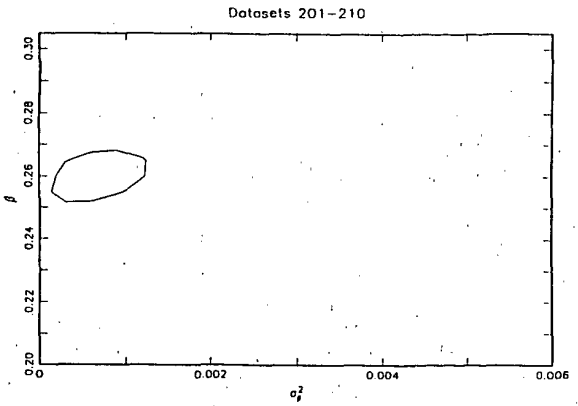


Figure 1: 95% confidence contour

4.2 Analysis

4.2.1 Superpopulation parameter estimation

The results of maximizing the likelihood described in equation (9) are presented in Table 1.

DATA SETS	$\hat{\beta}$	$\hat{\sigma}_\beta^2$
101 - 112	0.470	0.0014
201 - 210	0.260	0.0007

Table 1: Superpopulation parameter estimates.

In Figure 1 we present the 95% confidence contour for (β,σ_β^2) described in equation (10) for datasets 201-210. This region does not include $\sigma_\beta^2 = 0$, and accordingly the hypothesis that $\sigma_\beta^2 = 0$ would be rejected at the 5% significance level. Two points are worth noting in relation to this: firstly, a 99% confidence region would include $\sigma_\beta^2 = 0$; secondly, by taking account of the size of $\hat{\sigma}_\beta$, the formula given in equation (13) gives appropriate weighting to

the diversity or otherwise of the datasets in producing shrunken estimates of β_i. Identical conclusions apply to datasets 101-112, but are omitted for reasons of space limitations.

4.3 Shrunken estimates of regression parameters

In Tables 2 (for whitecaps) and 3 (for whitecrests) we present the simple least squares estimates $\tilde{\beta}_i$ and the shrunken or pooled estimates $\hat{\beta}_i$ [as described in equation (13)] of $\beta_i, i = 1, \ldots, k$. In parentheses, after each such estimate, we give the standard error of the estimate.

DATA SET	$\tilde{\beta}_i$ (ls)	$\hat{\beta}_i$ (shrunken)
101	0.442(0.016)	0.462(0.008)
102	0.446(0.023)	0.464(0.012)
103	0.534(0.023)	0.480(0.009)
104	0.508(0.013)	0.481(0.007)
105	0.277(0.021)	0.400(0.013)
106	0.407(0.025)	0.458(0.011)
107	0.577(0.017)	0.500(0.009)
109	0.468(0.026)	0.469(0.015)
110	0.552(0.026)	0.490(0.013)
111	0.430(0.010)	0.452(0.007)
112	0.480(0.013)	0.475(0.009)

Table 2: Estimates of β_i, whitecaps.

In examining the results in Tables 2 and 3, it will be noted that the shrinkage of the estimates of β_i is much more pronounced for the whitecrest than for the whitecap data. This is due to the fact that the superpopulation variance (σ_β^2) is substantially smaller (by a factor of about 9) for the whitecrest data. This in turn leads to the superpopulation mean (β) being given much larger weighting [via the formula in equation (13)] in the case of the whitecrest data.

DATA SET	$\tilde{\beta}_i$ (ls)	$\hat{\beta}_i$ (shrunken)
201	0.216(0.016)	0.256(0.008)
202	0.282(0.014)	0.266(0.005)
203	0.232(0.012)	0.257(0.006)
204	0.443(0.031)	0.267(0.003)
205	0.344(0.010)	0.283(0.005)
206	0.260(0.012)	0.264(0.006)
207	0.300(0.015)	0.275(0.008)
208	0.262(0.014)	0.264(0.007)
209	0.250(0.004)	0.255(0.003)
210	0.224(0.006)	0.243(0.005)

Table 3: Estimates of β_i, whitecrests

4.4 Conclusions

This methodology provides a simple means of achieving a balance between the two strategies of combining a number of disparate data sets into a single data set, on the one hand, and treating each set separately, on the other. It is intuitively appealing in that the amount of pooling (shrinkage) is proportional to the extent of homogeneity among the datasets.

Acknowledgement

This work has been supported in part by ONR Grants N00014-90-J-1538 and N00014-91-J-1444.

References

[1] Gaver, D.P., Lehoczky, J.P. (1987) Statistical Analysis of Hierarchical Stochastic Models : Examples and Approaches. *Annals of Operations Research* **8** 217–227.

ON REGULARIZATION OF BEST LINEAR ESTIMATOR IN LINEAR INVERSION WITH NONNEGATIVE COVARIANCE STRUCTURE

Hoang H.S.,Talagrand O.[1]

Laboratoire de Meteorologie Dynamique, Ecole Normale Superieure

24, rue Lhomond, 75231 PARIS Cedex 05 (FRANCE)

[1] On leave from Institute of Theoretical Physics, P.O. Box 429 Bo Ho, Hanoi 10000 Vietnam

Abstract. *The statistical regularizators for best linear (BL) estimators for unknown vector of parameters in general linear inversion model with arbitrary nonnegative covariance structure are obtained on the basis of theory of regularizing operators.*

1° Intruduction. Problem Statement. Inverse and ill-posed problems are very important in several domains of climate studies such as estimation of atmospheric parameters using passive measurements of thermal radiation through linear inversion process, assimilation of satellite radiances in numerical weather prediction ....The need to investigate ill-posed problems originated two branches of mathematical analysis : theory of generalized inverses and theory of regularization /1,2/ . Most investigated regularization methods are developed for regularizing the **C'**-generalized solutions /2,3/. In this paper we concentrate attention to the case when the *a priori* knowledge of statistical properties of signal and noise is available /4,9/ . Concretely, consider the observation model $(\mathbf{z},\mathbf{H},\mathbf{V}^2)$

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{v} \qquad (1)$$

where $\mathbf{z} \in \mathbf{R}^N$ is vector of observations, $\mathbf{H}$ is observation matrix , $\mathbf{v}$ is experimental error, $\mathbf{E}(\mathbf{v}) = 0$, $\mathbf{E}(\mathbf{v}\mathbf{v}') = \mathbf{V}^2$, $\mathbf{V}^2$ may be singular. It is known that the best linear estimator (BL-estimator) for unknown vector $\mathbf{x} \in \mathbf{R}^n$ can be represented in the following form

Lemma 1.(*General Gauss-Markov Theorem /1/*) . For the observation model (1), the BL-estimator for $\mathbf{x}$ is given by

$$\begin{aligned} \overline{\mathbf{V}} &= \mathbf{V}(\mathbf{I} - \mathbf{H}\mathbf{H}^+) \\ \mathbf{G} &= \mathbf{H}^+[\mathbf{I} - (\overline{\mathbf{V}}^+\mathbf{V})'] \\ \hat{\mathbf{x}} &= \mathbf{G}\mathbf{z} \end{aligned} \qquad (2)$$

In (2) $\mathbf{H}'$ is the transpose of $\mathbf{H}$.
From (2) it follows the generalized least-squares (GLS) estimator if $\mathbf{V}^2$ is nonsigular

$$\hat{\mathbf{x}} = (\mathbf{H}'\mathbf{V}^{-2}\mathbf{H})^+\mathbf{H}'\mathbf{V}^{-2}\mathbf{z} \qquad (3a)$$

and least squares (LS) estimator if $\mathbf{V}^2 = \mathbf{I}$

$$\hat{\mathbf{x}} = (\mathbf{H}'\mathbf{H})^+\mathbf{H}'\mathbf{z} = \mathbf{H}^+\mathbf{z} \qquad (3b)$$

In (2)(3) $\mathbf{H}^+$ is the pseudoinverse of $\mathbf{H}$; $\mathbf{V}$ is symmetric square root of $\mathbf{V}^2$.

Let only the δ -approximation of value φ be given where φ stands for any of $\mathbf{H}$, $\mathbf{V}^2$, $\mathbf{z}$, i.e. , instead of knowing the true value φ we have $\tilde{\varphi}$,

$$|\varphi - \tilde{\varphi}| \leq \delta \qquad (4)$$

where $|\cdot|$ denotes the euclidian norm in linear space. Since pseudoinverse operator, in general, is sensitive to initial given data the direct calculation of BL-estimator based on (2),(3) is ill-posed problem /1-4/ (when $\mathbf{H}$ and $\tilde{\mathbf{H}}$ have different ranks). The object of study in present paper is to obtain the

regularizators for estimators (1). Mention that the regularizators for estimators (3a,b) are investigated in /4/ for the case of filtering the state of dynamic system affected by the correlated random noises.

2°Approximator for BL-estimator (2) The main idea underlying in design of the stable procedures in this paper consists of introducing two families of estimators depending on two small positive parameters α and β. The first one playing the role of approximate estimator for (2) is given by

$$x_\alpha = H'(HH' + \alpha^2 V^2)^+ z \tag{5}$$

Lemma 2. The limit $x_l = \lim_{\alpha \to 0} x_\alpha$ always exists and is equal to estimator (2).

Comment. Computation of $\hat{x}$ based on (5) is considerablly more simple than (2).

Proof of Lemma 2. Consider matrix

$$S_\alpha = \alpha^{-2}.H'(HH'.\alpha^{-2} + V^2)^+ \tag{6}$$

Applying the Theorem 4.9 /1/ we have

$$(HH'\alpha^{-2} + V^2)^+ = (\bar{V}'\bar{V})^+ + \alpha^2 S_1 + O(\alpha^4)$$
$$S_1 = (I - \bar{V}^+V)(HH')^+(I - \bar{V}^+V)'$$

where α is small enough , $\bar{V}$ is defined as in Lemma 1. However according to definition of pseudoinverse matrix there exists some matris W such

$$(\bar{V}'\bar{V})^+ = (\bar{V}'\bar{V})W = \bar{V}'W_1, W_1 = \bar{V}W$$

Taking limit of S_α as $\alpha \to 0$ we see

$$H'(\bar{V}'\bar{V})^+ = H'\bar{V}'W_1 = 0$$
$$\lim S_\alpha = H'(I - \bar{V}^+V)(HH')^+(I - \bar{V}^+V)'$$

However since $H'\bar{V}^+ = 0$ then $\lim S_\alpha = H^+(I - (\bar{V}^+V)')$ that proves the Lemma $\otimes$

3°Regularizator for BL-estimator (2) Lemma 2 plays the important role in design of stable estimation procedures. To use it, let us define the second family of estimators depending on paramater β

$$x_\beta = \alpha^{-2}H'(HH'\alpha^{-2} + V^2 + \beta.I)^{-1}z \tag{7}$$

In what follows for random vectors x , y

$$\bar{x} \hat{=} E(x), K_{xy} \hat{=} E(xy'), K_x \hat{=} E(xx')$$

$E(.)$ denotes a mathematical expectation.

Lemma 3. For the estimator (7)

$$\lim_{\beta \to 0} x_\beta = x_\alpha \tag{8}$$

Proof. Consider the observation model

$$\zeta = Hy + v, \quad E(y) = 0,$$
$$E(yy') = I.\alpha^{-2}, \quad E(yv') = 0 \tag{9}$$

Then we have

$$K_{y\zeta} = E(y\zeta) = H'.\alpha^{-2},$$
$$K_\zeta = E(\zeta\zeta') = HH'\alpha^{-2} + V^2$$

Since (Theorem 9.2.2 , /1/)

$$K_{\zeta y} \in R(K_\zeta),$$

$R(K)$ denotes linear space generalized by columns of K , we see that there exists some matrix Y that $K_{\zeta y} = K_\zeta Y$ .Therefore

$$x_\beta = K_{y\zeta}(K_\zeta + \beta I)^{-1}z = Y'K_\zeta(K_\zeta + \beta I)^{-1}z$$

and from the Lemma 3.3 /1/

$$\lim_{\beta \to 0} x_\beta =$$
$$= Y'K_\zeta K_\zeta^+ z = K_{y\zeta}K_\zeta^+ z = x_\alpha \otimes$$

Now suppose instead of true φ only $\tilde{\varphi}$ and information (4) are given,where without loss of generality φ stands for any of K_ζ and $K_{\zeta y}$ too. Denote by $\tilde{x}$ the estimator defined on the basis of $\tilde{\varphi}$ .It means that $\hat{x}$ is defined by (2) with true value φ , while $\tilde{\hat{x}}$ is of the form (2) too but with $\varphi \hat{=} \tilde{\varphi}$.

Theorem 1. Let $\beta_0(\delta)$ and $\beta_1(\delta)$ be some nondecreasing continuous and nonnegative functions such that

$$\frac{\delta}{\beta_1(\delta)} \le \beta_0(\delta)\ ,\ \beta_0(0)=0\ .$$

Then $\forall\,\varepsilon>0$ there exists $\delta_0>0$ such that for any $\delta \le \delta_0$,

$$\left|\tilde{x}_\beta - x_\alpha\right| \le \varepsilon \qquad (10)$$

if the parameter β satisfies the condition

$$\frac{\delta}{\beta_1(\delta)} \le \beta \le \beta_0(\delta) \qquad (11)$$

Theorem 1 may be proven by the way similar to that of proving Theorem 4 in /4/.

From (10) it follows that there are many ways to choose the suitable *regularizing* parameter β .A problem of optimal choice of β is of interest for many authors and is described in /2,3/.We mention only that the simple choice of β can reduce considereblly the computational cost of regularizator. One simple way is to choose β satisfying

$$\delta / \beta(\delta) \to 0 \ \text{as}\ \delta \to 0 \qquad (12)$$

The combination of Lemmas 2-3 ,Theorem 1 now gives the following main result

Theorem 2.(*On regularizator for (2)*). Let $\hat{x}$ be estimator defined by (2). Let $\tilde{x}_\beta$ be determined by (7) where $\tilde{\varphi}$ stands for φ, and $\tilde{\varphi}$ satisfy (4). Then there is the possibility to choose the parameter β as a function of δ in such a way that $\forall\,\varepsilon>0$ there exists $\delta_0>0$ such that for any $\delta \le \delta_0$

$$\left|\tilde{x}_\beta - \hat{x}\right| \le \varepsilon \qquad (13)$$

4° Regularizator for unknown vector with a priori information. Let us be given in addition to (2)

$$\mathbf{E}(x)=\bar{x},\ \mathbf{E}[(x-\bar{x})(x-\bar{x})']=M \qquad (14)$$

$$\mathbf{E}[(x-\bar{x})v'] = -N \qquad (15)$$

The BL-estimator for the model (1)(14)(15) is of the following form /5/

$$\hat{x} = \bar{x} + K.(z - H.\bar{x})$$

$$K = K_{y\zeta}K_\zeta^+,\ y \hat{=} x - \bar{x},\ \zeta = z - H\bar{x} \qquad (16)$$

The solutions of many problems in the fields such as assimilation of observations, linear inversion from satellite measurements... may be expressed in form (16). For initialization problem in dynamic system /10/,to estimate $\{x_0, x_1, \ldots, x_N\}$ based on observations $\{z_0, z_1, \ldots, z_N\}$ one minimizes LS criterion J which is a "distance" function between observations and system states. Taking a derivative of J by x_i gives the system of normal equations for determining $\{x_0, x_1, \ldots, x_N\}$. A simplification occurs assuming that a solution $\{x_0, x_1, \ldots, x_N\}$ is uniquely defined by initial state x_0. Expressing all states δx_i of linearized pertubations equation through δx_0 leads to considerable reduction of number of control variables . Mention that if the amount of a priori information related to covariance structure of noises (and of initial state x_0) is available we can accordingly improve the exactness of estimator for δx_0 ,considering the Markov estimator (2) (and BL-estimator (16)) . It would be to emphasize that in practice any a priori information we have must be given only approximately.Therefore a great attention has to be paid to construction of estimators robust to model errors especially when in the domain of δ -approximation of φ the dynamic system may be unstable and/ or observation matrix H may have the different ranks .

The similar situation takes the place in the general linear inversion scheme . Namely , according to scheme of /11,12,13/

$$\hat{x} = x_0 + W.(y_m - y_c\langle x_0\rangle) \qquad (17a)$$

where $\hat{x}$ is vector of retrieved parameters, x_0 - first- guess value of vector, y_m-vector of multi-channel radiance, $y_c\langle x_0\rangle$-vector appropriate to first guess, W is 'inverse matrix'. In minimum variance formulation

$$\mathbf{W} = (\mathbf{K}.\mathbf{C})'.(\mathbf{K}.\mathbf{C}.\mathbf{K}' + \mathbf{E})^{-1} \qquad (17b)$$

where one assumes, in scheme of (15), the uncorrelation between retrieved parameters and measurement noise, and nonsingularity of $\mathbf{K}_\zeta = \mathbf{KCK}' + \mathbf{E}$. The similar formulas may be obtained for direct use of satellite sounding radiances in NWP (see /12/ for detail). Thus (16) is given in general form for which any assumption on nonsingularity of covariance matrices as well as on uncorrelation between vectors of parameters, of noise and residual vector is not required. One can see that in (16) there is only one source, namely $\mathbf{K}_\zeta^+$ which could make the process of computation of $\hat{\mathbf{x}}$ to be unstable. By the similar way as outlined in proving Lemma 3, Theorems 1-2, it can be shown that by appropriate choosing the parameter β in the estimator

$$\tilde{\hat{\mathbf{x}}}_\beta = \tilde{\bar{\mathbf{x}}} + \tilde{\mathbf{K}}_{y\zeta}.(\tilde{\mathbf{K}}_\zeta + \beta.\mathbf{I})^+(\tilde{\mathbf{z}} - \tilde{\mathbf{H}}\tilde{\bar{\mathbf{x}}}) \qquad (18)$$

we can approach as closely as required to the BL-estimator (16), dependingly on the degree of uncertainty δ given a priori. To regularize (18), the theory presented above is equally applied letting in addition φ stand for any of $\bar{\mathbf{x}}$, $\mathbf{M}$, $\mathbf{N}$ too.

Mention, to advoid singularity of $\mathbf{K}_\zeta$ in /5,6/ the formula like (18) for $\hat{\mathbf{x}}_\beta$ is proposed, therefore $\hat{\mathbf{x}}$ in (16) is obtained by process $\lim_{\beta\to 0}\hat{\mathbf{x}}_\beta$ (see Lemma 2).

5° Concluding remarks. The regularizators presented in this paper are constructed in framework of stochastic version of regularization method ,when a priori knowledge of statistics of signal and noise is available. That is why the form of proposed regularizators is different from the well-known regularizators for $\mathbf{C}'$-generalized solution /3,7,8/ which is given by $\mathbf{x}_{\mathbf{C},\alpha} = (\mathbf{H}'\mathbf{H} + \alpha^2\mathbf{C}'\mathbf{C})^{-1}\mathbf{H}'\mathbf{z}$.

Up to now, for the nonlinear estimation problems the linearization technique is commonly used. Since after linearization one comes to model of the form (1) with complicated covariance structrure, one possible way is to use the information given in the form of $|\varphi - \tilde{\varphi}| \le \delta$ to find the unknown parameters. Thus, the results derived in this paper are expected to be useful for solving some nonlinear estimation problems where some amount of a priori information is available in order to get the robustness of estimation procedure.

6°Acknowledgement. The authors are mostly grateful to Ministere de la Recherche et Technologie for its financial support during visit of first author in Labo de Meteorologie Dynamique (LMD, ENS, Paris, France) and to Director of LMD, Prof. R. Sadourny for his valuable encouragement and suggestions in this work.

REFERENCES

/1/Albert A.(1972) Academic Press, NY & London, /2/ Tikhonov A.N. (1963) Soviet Math. Dokl. Vol. 4, pp. 1035-1038, /3/ Morozov V.A.(1984) Berlin,West Germany, Springer-Verlage, /4/ Nguyen T.L., Hoang H.S.(1983) Autom. Remote Contr., N-4, pp.58-73 , /5/ Lipsher R.S., Shiryayev A.N. (1974) Nauka, Moskow, /6/ Ljung L., Soderstrom T.(1982) Cambridge -Mass., MIT, Press, /7/ Bertero M.,et al.(1988) Proc. of IEEE, Vol.76, N°8, August, pp. 869-889, /8/ Nashed M.Z. (1971) L.B.Rall,NY: AP, /9/ Nguyen T.L., Hoang H.S.,Pham T.L. (1990), Proc. of Int. Symp. Assim. of Observ. in Meteorol. & Oceanog.(ISAOMO), WMO, pp. 437-442, /10/ Talagrand O.,Courtier P.(1986) Tellus, 38A, 97-110, /11/ Eyre J.R. (1987) Q.J.R. Meteorol.Soc. 113, 279-292, /12/ Eyre J.R. (1990) Proc. ISAOMO, WMO, pp. 117-121, /13/ Rodgers C.D.(1976) Rew. Geophys. & Space Phys.,V-14, N-4, 609-624.

STEPWISE LOGISTIC DISCRIMINATION ANALYSIS APPLIED TO THE EXAMINATION OF METEOROLOGICAL INFLUENCES ON VEGETABLE HARVEST YIELDS [1]

LUE Chun Lian, MATHES Harald *, CHEN Shun Hua
Nanjing Institute for Meteorology, 210044 PR China
* University of Munich, Akademiestr. 1, 8000 Munich 40, Germany

The paper presents a statistical method to investigate the causal relationship of meteorological conditions on harvest yields. If we suppose it is only important to consider whether the yields are of good, medium or bad size, then the harvest variable can be categorical. In this case the *multivariate logistic regression model* can be applied see DAY & KERRIDGE(1967). Owing to the great number of exogenous variables, a selection method is used to reduce the meteorological variables to a small number of high–significant causes. The procedure of *crossvalidation* is applied to check the validity of the modeled relationship with the selected variables. This means that for the estimation of the model a data point is omitted and the corresponding endogenous variable is forecasted by the exogenous data. This procedure is repeated until all data points have been omitted once. The prognosis form of logistic regression is identical to the *logistic discrimination analysis.*

1. The Model

Suppose that $X = (X_0,X_1,...,X_p)'$ is an exogenous variable vector with X_0 as dummy variable for a constant term ($X_0 \equiv 1$) and the assignment to one of m possible categories is described by the endogenous variable Y with the realizations $Y = 1,...,m$. Further we assume that a sample x_i and y_i with $i = 1,...,n$ of n independent observations of $(X',Y)'$ is available, then the logistic discrimination analysis (LDA) can be formulated by the following conditional probability of Y_i with given x_i:

$$(1) \qquad \pi_{is} = P(Y_i = s|x_i) = \frac{\exp(u_s(x_i,\beta))}{\sum_{r=1}^{m} \exp(u_r(x_i,\beta))} \quad \text{for } s = 1,...,m \text{ and } i = 1,...,n;$$

[1] The research was supported by Stiftung VW and guided by Prof. Dr. Hans Schneeweiß.

$$\text{with } u_s(x_i,\beta) = \sum_{j=0}^{p} \beta_{sj}\, x_{ij} \text{ for } s = 1,\ldots,m-1 \text{ and } u_m(x_i,\beta) \equiv 0.$$

The linear functions u_s, $s = 1,\ldots,m$ are called logistic discrimination functions and β_{is} are the logistic discrimination coefficients. An exogenous data point is assigned to the category with the highest conditional probability in (1).

Maximizing the conditional likelihood with respect to β, where $\underline{X}$ contains the n observed exogenous data points,

$$(2) \qquad L(\beta|\underline{X}) = \prod_{i=1}^{n} \prod_{s=1}^{m} \pi_{is}^{y_{is}} \quad \text{with} \quad y_{is} = \begin{cases} 1 & \text{if } y_i = s \\ 0 & \text{otherwise} \end{cases},$$

$$\beta = (\beta_0',\ldots,\beta_p')', \ \beta_j = (\beta_{1j},\ldots,\beta_{qj})' \text{ and } q = m-1$$

leads to the ML–estimate of β, which is computed by an iterative procedure, see for example MATHES & LUE(1991).

2. Variable Selection Method

The following stepwise selection procedure is based on a significance criterion of appropriate likelihood ratios. [2]

Step 0: The estimate $\hat{\beta}_0$ is obtained by maximizing (2) when the discrimination functions contain only the variable X_0.

Step 1: At step 1 the variable X_{j_1} joins X_0 in the set of significant variables, if it holds:

$$(3) \qquad \lambda_{j_1} = \max_{j=1,\ldots,p} \lambda_j \quad \text{with} \quad \lambda_j = \left\{-2 \ln \frac{L(\hat{\beta}_0)}{L(\hat{\beta}_0,\hat{\beta}_j)}\right\}, \ j = 1,\ldots,p$$

and if λ_{j_1} exceeds some pre–chosen constant K. If the last condition fails the procedure stops and the discrimination functions contain only X_0. Under appropriate regularity conditions it is well–known that λ_{j_1} is the test statistic for the likelihood ratio test of H_0: $\beta_{j_1} = 0$ versus H_1: $\beta_{j_1} \neq 0$. Under H_0 and as $n \to \infty$, λ_{j_1} will have a χ^2 – distribution with q degrees of freedom.

---

[2] The describtion contains only the forward selection. For backward selection and further details see HOMSER, WANG, LIN & LEMESHOW(1978).

Step 2: The Variable X_{j_2} with $j_2 \neq j_1$ enters the set if it fullfils the following conditions ($\hat{\beta}_{j_2}$ maximizes the following expression brackets):

$$(4) \qquad \lambda_{j_1,j_2} = \max_{j=1,\ldots p} \left\{ -2 \ln \frac{L(\hat{\beta}_0, \hat{\beta}_{j_1})}{L(\hat{\beta}_0, \hat{\beta}_{j_1}, \hat{\beta}_j)} \right\} \quad \lambda_{j_1,j_2} > K.$$

The procedure continues until all of the exogenous variables have entered the discrimination functions, or at a step, say s, $\lambda_{j_1,\ldots j_s} < K$ or the exogenous observation points can be completly assigned correctly to their observed categories by the discrimination functions (so called 100% matching ratio).

3. Analysis of meteorological influences on low supply of vegetables in spring and autumn in Guangzhou

The endogenous variable is constructed by the vegetable amount supplied in march and september in Guangzhou from 1954 to 1983, excluding the years 1960 to 1962 (outliers), owing to Chinese cultural revolution. The harvest yields are adjusted by means of a linear trend model, because the technological developments have to be taken into account.

$$(5) \qquad z_t = \hat{z}_t + \hat{e}_t = \hat{a} + \hat{b}\,\hat{z}_t + \hat{e}_t \quad \text{with} \quad t = 1954,\ldots,1959,1963,\ldots,1983$$

We got a R^2 of 0.21 for the march sample and 0.51 for the september data. The residuals $\hat{e}_t$ are standardized by their corresponding estimated value of crop yield $r_t = \hat{e}_t/\hat{z}_t$. To obtain the categorical endogenous variable we divided the r_t values into three categories. The following boundaries were chosen for the march data [3]:

(6) $r_t \in]-\infty,-0.1] \Rightarrow y_t = 1$ (low amount); $r_t \in]0.2,\infty[\Rightarrow y_t = 3$ (high amount) and the rest belongs to category 2 (middle amount).

The exogenous variables are taken from the 5 months prior to the harvest month and the harvest month itself. [4] We use the average decade values for the variable groups precipitation (P), temperature (T) and sunshine (S), whereby the group

---

[3] The boundaries of the september sample are similar.

[4] For our purposes the three 10–day periods in a month shall be called decades.

letters of the march harvest is preceded by M (e.g. MP or MT) and those of the september harvest by S. Each variable group has 18 variables (3 decades x 6 months) which are enumerated currently (e.g. MP1 to MP18). Additionally we have 12 specially constructed T variables due to their important effects during crop growing.

Applying the stepwise selection method of chapter 2 leads to the following significant variables until a 100% matching ratio is reached: [5]

March: MT30 – MIT in March of th CY;
MT25 – AVT in January of the CY;
MP2 – P in the 2nd decade of October of PY
MS18 – S in the 3rd decade of March of CY;
MT6 – T in the 3rd decade of November of the PY;

Sept.: SS6 – S in the 3rd decade of May of the CY;
ST18 – T in the 3rd decade of September of the CY;
SP4 – P in the 1st decade of May of the CY;
SS1 – S in the 1st decade of April of CY;
ST24 – number of days in Sep. with MATD $\geq$ 30°C of the CY.

The results of crossvalidation are for the march sample 18/27 and for september 17/27. The selected variables show that especially the temperature variables of the harvest months are responsible for the group separations. For purpose of comparison further methods of discrimination analysis and variable selection procedures were applied which have confirmed these results.

REFERENCES

ANDERSON J.A.(1972) Separate sample logistic discrimination, Biometrika 59, 19–35.

DAY N.E. & KERRIDGE D.F.(1967) A general maximum likelihood discrimination, Biometrics 23, 313–323.

HOMSER D.W. & WANG C.Y. & LIN I.C. & LEMESHOW S.(1978) A coputer program for stepwise logistic regression using maximum likelihood estimation, Coputer Programs in Biomedicine 8, 121–134.

MATHES H. & LUE C.L.(1991) A quadratic logistic discrimination algorithm in matrix form, Computational Statistics Quarterly 4, 269–280.

---

[5] Abbreviations: current year (CY); previous year (PY); average temperature (AVT); minimum temperature (MIT); maximum temp. of a day (MATD)

A Spatio - Temporal Model for Rainfall

by

SUNDAR RAJAN, R., SHUKLA, G.K., KUNDU, D.
DEPARTMENT OF MATHEMATICS, IIT KANPUR, INDIA

ABSTRACT

In the present approach we have considered a simple approach to model rainfall spatio-temporally at a point and time of interest. Let us say we have a linear regression model with only one predictor variable which is rainfall at some location and known fixed time lag h. We propose to extend this model by also estimating the location of the predictor variable. Note that this extension will not necessarily result in a linear model.

Weather phenomena are well known for their dependence on the notion of "local direction of weather movement", and isotropic models may prove inadequate. We propose an alternative approach which permits directions with refernce to this "local dirction of weather movement", while ensuring that the model does not depend on adhoc global positioning conventions. For this purpose we introduce a function u(x,y), which returns a direction which we denote by $\overrightarrow{u(P)}$, a unitvector, where P is the point (x,y). Rotating u through 90 degrees anticlockwise we get a vector which we denote by v(x,y). The unitvector u(x,y) denotes this "local direction of weather movement", u(x,y) can point in different directions for different points P=(x,y) and may be assumed to have nice properties like continuity etc. Let C be the covariance

function between rainfall at two locations $P_1=(x,y)$ and $P_2=(x',y')$ at time t and $t'(\leq t)$ respectively, where P_2 is close to P_1. $u(x,y)$ has also the interpretation that C is a function of coordinates of P_2 in the local $(u(P_1),v(P_1))$ co-ordinate frame with origin at P_1 and $t-t'$ alone. Note that "P_2 *close* to P_1" is an essential requirement as the unit vector u(P) can point in different directions for different Ps. In actual model building and for purposes of developing estimation procedures for model parameters here u(x,y) is assumed to be known a priori.

Eventhough with reference to the local u-vco-ordinate frame with origin at P the <u>direction</u> does matter the process is still assumed to be homogeneous (i.e., stationary in time and under translation of origin). It is infact on the basis of this assumed hormogeneity that we propose to estimate the <u>location</u> of the predictor variable.

The model for rainfall at $P_1=(x,y)$ on day t denoted by $R_{(x,y,t)}$, is assumed to be given by

$$R_{(x,y,t)} = a\, R^{*}_{(c,d,t-h)} + b + \epsilon_{(x,y,t)} \qquad (1)$$

where

$$R^{*}_{(c,d,t-h)} = R_{(x',y',t-h)},$$

(x',y') represents the point $x\vec{i} + y\vec{j} + c.u(P_1) + d.v(P_1)$,

$\vec{i}$ and $\vec{j}$ are the latitude-longitude axis unit vector and,

(c,d) determine the position of (x',y') in the $(u(P_1), v(P_1))$ co-ordinate frame with origin at P_1.

$\epsilon_{(x,y,t)}$ is assumed to satisfy

$$E(\epsilon_{(x,y,t)}) = 0,\ V(\epsilon_{(x,y,t)}) = \sigma^2,$$

$$E(\epsilon_{(x,y,t)}\ \epsilon_{(x',y',t')}) = 0 \text{ if } x \neq x', y \neq y'.$$

While we have assumed an extremely simple model equation, it should be noted that this is solely to develop the methodology by first working out the estimation procedure in simple cases before going on to add complicating features. In view of the unconventionality of the approach this seems advisable.

Covariance between rainfall at two spatial location P_1 and P_2, P_1 is our point of interest is assumed to decrease rapidly to zero as P_2 approaches the boundary of the circle of radius r centered at P_1 and becomes negleigible when P_2 is outside the circle centered at P_1 of radius r, r assumed to be a known fixed constant.

By this formulation we have introduced the notion of weather circle (sphere in 3-dimensions) to formalise the hypothesis that weather in a location has to necessarily arrive through its neighbours. Therefore for purposes of modelling we can consider only the neighbourhood of the locality of interest.

The model parameters to be estimated are a,b,c,d. c,d to be estimated subject to the condition that $c^2 + d^2 \leq r^2$.

There are no conventional methods in Statistics to our knowledge to tackle this problem. Since, normally, rainfall at

each spatial location is considered a distinct variable, estimating the location appears to amount to selecting the variable to be included into the model. However we choose to think of the spatial location (like time) as an aspect of the same physical variable (rainfall) so that, in a sense the model being proposed may be called autoregressive.

The point of this approach is to take advantage of the assumed homogeneity of the process to develop a single model which by a mere shift of origin to the locality of interest can supply a fit of the process for that locality. It is this requirement of Portability that makes the estimation of predictor variable's location with respect to the locality of interest necessary. Process here means the rainfall occurrence process.

The methods for estimating the parameters a,b,c,d are proposed.

IDENTIFICATION OF NON-LINEAR STOCHASTIC MODELS FROM SEISMIC RECORDS

ALVARO M.D. NUNES
Universidade de Macau
P.O. Box 3001
MACAU (via Hong Kong)

SYNOPTIC ABSTRACT

An autoregressive stochastic model (rth order) is proposed, as mathematical model, for seismic records of explosions and earthquakes. These models are based on deterministic formulations from explosions and earthquakes theories. Statistical properties of the estimation procedure are derived. Some examples as applications of the model to seismic records are considered.

1. INTRODUCTION

We give a brief summary of the relevant parts of the deterministic theory concerning the far-field displacement behaviour of underground explosions and earthquakes. The functional form for the reduced displacement potential was suggeted by Haskell (1967) and revised by Von Seggern and Blandfor (1972), giving

$$\Psi(\tau) = \Psi(\infty)\left\{1 - e^{-k\tau}\left[1 + k\tau - h(k\tau)^2\right]\right\} \tag{1}$$

where k is a time scale constant, h is a dimensionless constant, $\tau = t - r/a$ is the "retarded time" at distance r from the source. As far as the theory for earthquakes is concerned most results in this field are given in the frequency domain (ω) field and solely in terms of the amplitude spectrum $|U(\omega, r)|$, without phase information. The "ω- square" model due to AKi (1967) and revised by the suggestions of Von Seggern and Lambert (1970) that terms involving $\cos\theta$ can be neglected, then gives

$$|U(\omega,r)| = \frac{WD_0L}{r}\left\{1 + \left(\frac{\omega}{k}\right)^2\right\}^{-1} \tag{2}$$

A time - domain version for (2) is given by Harkrider (1976). He finds

$$U(t) = D_0K^2te^{-kt} \tag{3}$$

2. NONLINEAR STOCHASTIC MODEL

A general stochastic difference equation of the form

$$A(B)X(t) = M(B)\varepsilon(t) = 2(T), \qquad t = 0, 1, 2, \ldots \tag{4}$$

where $\varepsilon(t)$ is a zero mean Gaussian white noise process and A(B), M(B) are respectively polynomials of degree p and q in the backward shift operator B has a solution of form

$$X(t) = C_s(t-s) + I_s(t-s) \tag{5}$$

where s is some arbitrary time origin, such as the time at which the process was first observed. The complementary function, C_s, is the general solution of the difference equation

$$A(B)C_s(t-s) = 0 \tag{6}$$

Note that, like the complementary function, the autocorrelation function of X(t) also satisfies (5) (Priestley, M.B., 1980). It follows that the (statistical sense) spectrum of X(t) is algebraically equivalent to the (deterministic sense) spectrum of C_s, that is the Fourier transform of C_s.

3. EXPLOSIONS AND EARTHQUAKES

From (1) We deduce that

$$U_r(\tau) \approx (\alpha r)^{-1} \psi(\infty) \ell^{-hk} \left\{ k^2(1+2h)\tau - hk^3\tau^2 \right\} \tag{7}$$

Hence, from the above discussion, we see that with suitable boundary conditions the radial displacement function can be regarded as the complementary solution for a stochastic difference equation having general form

$$(1-\ell^{-k}B)^p X(t) = M(B)\varepsilon(t) \tag{8}$$

The model for the observed seismic record can then be written

$$(1-\ell^{-k}B)^p (1+B)^L Y(t) = \varepsilon(t) \tag{9}$$

Concerning the earthquakes we first examine AKi's ω-square model

with $VK_L = K_T$ as described by (3) in the time domain. Following the above procedure for explosions we are led to the single parameter second-order autoregressive model.

$$X(t) - 2\ell^{-k}X(t-1) + \ell^{-2k}X(t-2) = \varepsilon(t) \tag{10}$$

$$(1 - \ell^{-k}B)^2 X(t) = \varepsilon(t)$$

4. ESTIMATION OF THE PARAMETERS

We consider the approach for estimation the parameters by (Man and Wald, 1943) with amendments due the fact that the coefficients in our model are non linear functions of a single parameter. Given the data and the filter parameter L, (9) is a special case of the general model

$$X(t) + \alpha_1(k)X(t-1) + \alpha_2(k)X(t-2) + \cdots + \alpha_p(k)X(t-p) = \varepsilon(t) \qquad t = 1, 2, \ldots, n \tag{11}$$

where $\{\varepsilon(t)\}$ is a sequence of mutually independent N(O, σ^2) variates and $\alpha_1(k)$, $\alpha_2(k)$, ..., $\alpha_m(k)$ are known functions of the unknown parameter K. The estimate for σ^2 is simply

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{m}\sum_{j=0}^{n}\alpha_j(k)x(t-j)^2 \tag{12}$$

The asymptotic variance of the maximum likelihood estimate for k is then

$$Var(\hat{k}) = \left\{E_k\left[-\frac{d^2 \log_e L}{dk^2}\right]\right\}^{-1} \tag{13}$$

It follows that the sample etimate for this variance is

$$s^2(k) = \hat{\sigma}^2\left[\sum_{i=1}^{m}\left\{\sum_{j=1}^{n}\left(\frac{d\alpha_j}{dk}\right)X(t-j)^2\right\} + \sum_{i=1}^{m}\left\{\sum_{j=0}^{n}\alpha_j X(t-j)\right\}\left\{\sum_{j=1}^{n}\left(\frac{d^2\alpha_j}{dk^2}\right)X(t-j)\right\}\right]^{-1}_{k=\hat{k}} \tag{14}$$

with s(k) being the (estimated) standard error for k.

5. CONCLUSIONS

The asymptotic consistency of all the above estimators, without the normality assumption, can be deduced from the results of Hannan (1971) for least-squares estimation in time series models. The theoretical developments considered in this paper will be applied later on to smoothed real seismic records.

6. REFERENCES

Box, G.E.P. & Jenkins G.M. (1970). Time series analysis forecasting and control. Holden- Day, San Francisco.

Harkrider, D.G. (1976). Potentials and displacements for two theoretical seismic sources. Geophys, J.R. Astr. Soc. 47, 97-133.

Hannan, E.J. (1971). Nonlinear time series regression J. Appl. Prob. 8 767-780.

Mann, H.B. & Wadd, A. (1943). On the statistical treatment of linear stochastic difference equations. Econometrica, 11, 173-220.

Nunes, A.M.D., and Subba Rao, T. (1985). Identification of non linear (quadratic) systems using higher order spectra. Presented at the 7th IFAC Symposium and Identification and System Parameter Estimation, University of York, England.

Priestly, M.B. (1982). Spectral Analysis and Time Series. Academic Press, London.

APPLICATION OF THE COPLOT DISPLAY TECHNIQUE IN CLIMATIC REGIONALIZATION

Y. GOLDERICH AND A. RAVEH, BAR-ILAN UNIVERSITY, RAMAT GAN, ISRAEL

ABSTRACT

A new approach for dividing a geographical region into climatic sub-regions is presented by applying a new graphic display method for multivariate data analysis. This method, COPLOT, which is a variant of the BIPLOT technique, enables us to study the similarity of climatic stations, the correlation structure among climatic variables and the mutual relationships between the stations' observations and the variables.

COPLOT maps the rows of a matrix in such a way that similar rows (observations) are closely located on the map. Each variable is represented individually by an arrow. A measure of goodness-of-fit is computed and associated for each variable separately. COPLOT enables the simultaneous study of observations and variables for a set of data.

The input matrix consists of 47 weather stations in Israel with 14 climatic parameters. On this matrix a cluster analysis (with and without the preliminary factor analysis), and the COPLOT technique is applied. The decision on grouping areas is based on visual inspection of the COPLOT output display, with the aid of the cluster analysis results.

The next stage involves the reducing of the number of variables taking part in the final COPLOT display into six climatic variables, according to their goodness-of-fit, (as measured by the maximal correlations between the variables and their associated arrows) and reapplying both the and cluster COPLOT techniques. The final quasi objective new climatic regions suggested and their boundaries do not always match previous climatic or physiographic regions.

To: Yair Goldreich <f49036%barilvm.bitnet@utcs.utoronto.ca>
Subject: Re: Fifth Int. Meeting Stat. Clim.

Non-Parametric Trend Statistics and Rank Correlations of Long European Sea Level Pressure Time Series.

Michael Denhard and Christian-Dietrich Schönwiese

J.W. Goethe University, Institute for Meteorology and Geophysics,
Praunheimer Landstr. 70, D-6000 Frankfurt a.M. 90

1 Introduction

Usual linear trend analysis often is a misleading technique. A more appropriate non-parametric trend test was introduced by H.B. Mann (1945). This technique is modified and applied to annual and seasonal averages of European sea level pressure time series covering approximately the last century. The results are spatial and temporal distributions of significant trends and 'trend-regions', including similar or different years of initiation as well as termination of the trends detected. Furthermore, the time series correlations between the stations are calculated using linear as well as rank correlation methods and coherence analysis. There arise not only outstanding differences in the spatial trend behavior but also in the correlation patterns.

2 The Data

The monthly mean values of sea level pressure data (SLP) are based on a number of sources (Spangler and Jenne (1985)), including European SLP maps. The data were recently updated and corrected by Fuchs (1991). Only those time series are accepted having gaps of less than four years and being located below 600 m above sea level. Gaps are closed using 30 year mean values. Homogeneity errors in the SLP time series were eliminated as far as information on these errors was available. But nevertheless tests of homogeneity show that there still remain discontinuities in the time series. This should be kept in mind when interpreting the results of statistical tests. A regional analysis of the SLP data however, helps to avoid misinterpretation. The present study is based on 61 European SLP time series covering the period from 1889 to 1988.

3 Methods

To detect a trend in a time series the non-parametric Mann-Kendall statistic is used. It compares the null hypothesis of randomness and the alternative hypothesis of trend for a given sample of measurements. The trend-coefficient Q is defind as

$$Q = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} sgn\,(X_i - X_j) \qquad (1)$$

The distribution of Q tends to normality (Gaussian distribution) if the length N of the time series is at least ten. Its variance is then defined as

$$\sigma^2 = \frac{1}{18}[N(N-1)(2N+5) - \sum_{i=1}^{k} b_i(b_i-1)(2b_i+5)] \qquad (2)$$

where k is the number of different ties and b_i is the number of ties i. A trend may be called significant if the Q value exeeds three times its standard deviation σ. In order to find a trend in a time series every sub-series has to be tested against trend. In this study only those sub-series are considered, containing at least ten elements. This kind of variable trend analysis enables the detection of every trend no matter where it starts or ends.

The time series correlations between the stations are analysed using the linear correlation of Pearson (1966) as well as the rank correlations of Kendall and Spearman (see Kendall (1970)). Additionally, the Fisher (1970) transformation is applied to the linear correlation coefficient. Thus the SLP data do not have to be normally distributed even if the linear correlation coefficient of Pearson is calculated. The coherence is defined as

$$ch(f) = \frac{co^2(f) + qu^2(f)}{S_{xx}(f)S_{yy}(f)} \qquad (3)$$

where co is the cospectral density, qu is the quadrature spectral density and S_{xx} (S_{yy}) is the spectral density of the time series x (y) at the frequency f. To obtain the spectral estimates the method developed by Blackman and Tukey (1958) is used.

4 Correlation and Coherence

The highest significance of correlation between SLP time series is usually obtained by the linear correlation of Pearson. It is a matter of fact that the correlation of the sea level pressure is in first order linear. Although there are differences between linear and rank correlation methods, all correlation coefficients show nearly the same spatial distribution. Fig. 1 shows contour lines of Kendall's rank correlation for Europe related to the station of Akureyri (Iceland). It is obvious that the correlations depend on the geographical situation and are not simply a function of distance. Positive as well as negative correlations are found, having different spatial distributions in January and July. In general the correlations are much stronger in winter than in summer. Consequently the atmosphere shows different kinds of order.

Fig. 2 shows some examples of the interannual behavior of the correlations with Akureyri. A significant negative correlation, being present all over the year, can be found only between Akureyri and Ponta Delgada (North Atlantic Oscillation). The situation is different if stations on the continent are considered. Some stations do not have any significant correlations with Akureyri at all. Others show an interannual variability with a strong positive (northern Europe) or negative (southern Europe) maximum of correlation in Winter.

The coherence analysis reveals similar results as they are found for correlation. Again there is a much stronger coherence in winter than in summer. But the spatial distribution of the coherence depends not only on the geographical situation, it is also a function of frequency.

5 Variable Trend Analysis

If Kendall's method (1) is used as a variable trend statistic, a trend is characterized by its beginning, its end and its Q value. Fig. 3 shows the SLP time

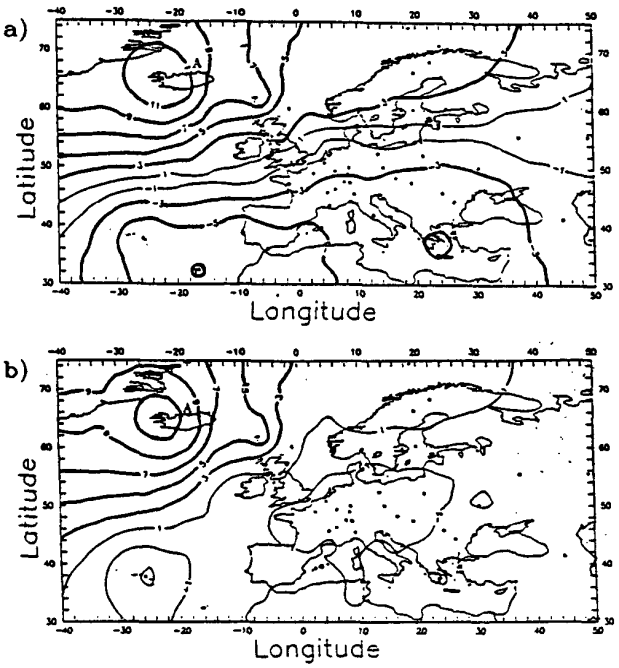


Figure 1: Contour lines of standardized correlation in Europe related to the station Akureyri for January (a) and July (b). The correlations depend on the geographical situation and show that the atmosphere is much more organised in January than in July.

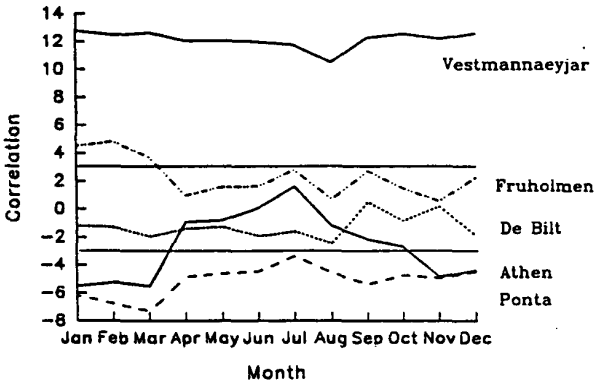


Figure 2: Interannual dependencies of standardized correlations with Akureyri. The correlations show an interannual variability which can have different shapes but which mostly has a stronger correlation in winter.

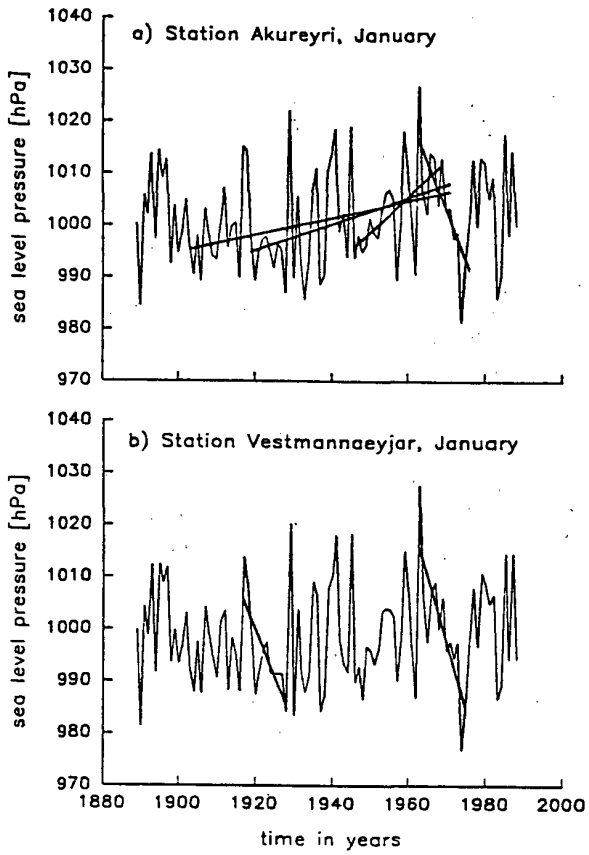


Figure 3: SLP time series of Akureyri and Vestmannaeyjar (Iceland) in January. Significant positive and negative trends are indicated (3σ threshold). These trends differ considerably although both stations are located very close to one another ($\sim$ 270 km).

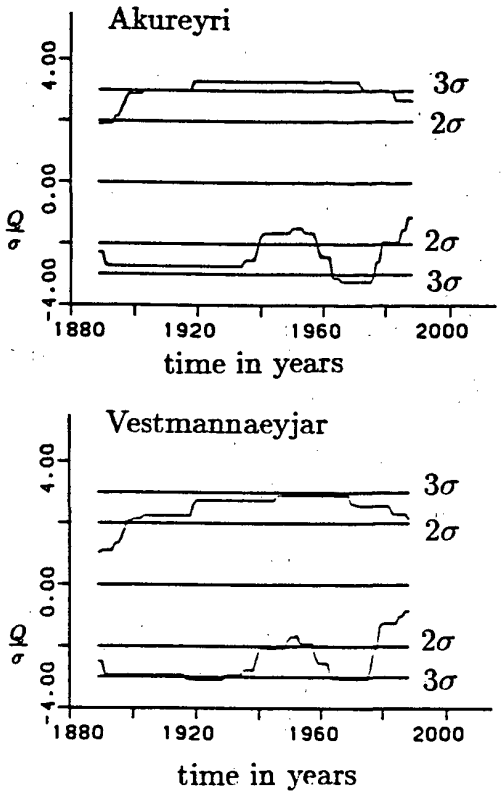


Figure 4: 'Trend-profiles' of the SLP time series of Akureyri and Vestmannaeyjar in January. The 'trend-profiles' of Akureyri and Vestmannaejar are much more similar than the trend pattern.

series of Akureyri and Vestmannaeyjar in January, both located in Iceland. Significant positive as well as significant negative trends are found and indicated (3σ threshold). Although the annual variability of both time series is very similar, these trend patterns are different.

In order to identify and to classify the various trends, a so-called 'trend-profile' is introduced. Every sub-series j of the original time series has its own trend coefficient Q_j. For the overlapping time period of different sub-series only the maximum (minimum) value of the trend coefficient is chosen, if Q_j is positive (negative). A positive or a negative 'trend period' is defined as the period of time in which the trend-profile exeeds a certain level of significance.

Fig. 4 shows the 'trend-profiles' of the SLP time series already presented in Fig.1. The positive trends of Akureyri are merged to one significant trend period from 1903 to 1971. In contrast to the significant trends, the trend-profiles of both stations are very similar. Consequently there is a good correlation between both stations, which is not caused by long term significant trends. The threshold for significance is the most important parameter in trend analysis. A change of this threshold causes totally different significant trends. If the threshold is lowered many more single trends are found. Their number can be enormous, because every subseries of the examined time series may be a significant trend. Therefore, it is quite useful to combine all trends in a trend-profile, but it is also true that information is lost while merging the single trends.

Fig. 5 shows trends in Europe for January and July. The beginning and the end of the trend periods are not indicated, because it would not be possible to recognize anything at all. A station with a

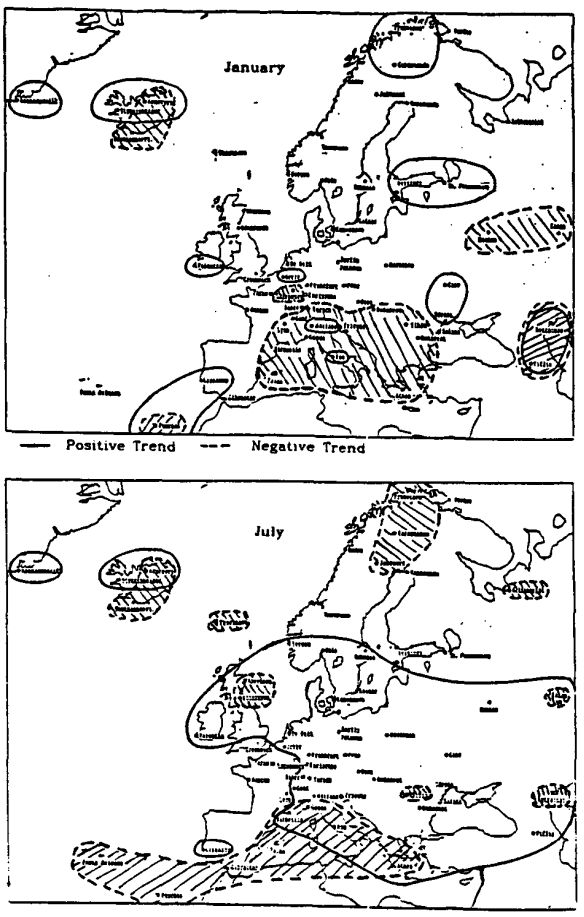


Figure 5: Trends in Europe during 1889 to 1988. A station has a significant trend, if at least one significant trend period is found (the 3σ threshold is used). The trends seems to be local phenomena, having a random nature.

positive trend symbolize that at least one significant trend period is found. There are a number of different trends even at stations located very close to each other. The trends seems to be local phenomena, having a random nature. No systematic change or even an annual cycle can be found. There is no systematic long term change as it would be expected from greenhouse forcing.

In comparison with correlation and coherence, the spatial distribution of trends is totally different. A significant correlation does not cause the same significant trends. This means that the correlation is caused by the year to year variability and has to be separated from long-term trends.

References

Blackman, R.B., Tukey, J.W. (1958): The measurement of power spectra, Dover, New York.

Fisher, R.A. (1970): Statistical methods for Research Workers, Oliver & Boyd, Edinburgh.

Fuchs, T. (1991): 'Globale Statistik langer Luftdruckreihen', Instiut for Meteorology and Geophysics, J.W. Goethe University, Frankfurt,Germany.

Kendall, M.G. (1970): Rank Correlation Methods, London: Griffin, Fourth Edition.

Mann, H.B. (1945): Nonparametric Test Against Trend, Econometrica, 13, 245-259.

Pearson, E.S., Hartley, H.O., 1966: Biometrica Tables for Statisticians, Vol. 1 and 2, Cambridge University Press, 1:1969, 2:1972.

Spangler, W.M.L., Jenne, R.L. (1985): World Monthly Surface Station Climatologie, NCAR, Boulder, USA.

Climatological aspects of the road maintenance in the Finnish Lapland during winter

by
Jaakko Helminen
Finnish Meteorological Institute
P.O.Box 503
SF-00101 Helsinki
Finland

1. Introduction

The wintertime road maintenance problem covers the following three main categories: the snow ploughing, the prevention of slipperiness by salting, and the prevention of slipperiness by sanding. From the administrative point of view it would be tempting to introduce a meaningful "winter index", which would enable one to allocate the wintertime road maintenance allowances optimally between the different road districts in the country. Because the maintenance activities are strongly related to certain weather events, one natural starting point for the considerations would be to try to develop a wintertime road maintenance cost model based on variables measuring the effects of these weather events.

From the climatological point of view the wintertime road maintenance needs vary regionally considerably in Finland. Roughly speaking the coastal areas close to the Gulf of Bothnia and the Gulf of Finland differ much from the inland areas. More to the north and east the snow ploughing represents the most important part of the needs. Especially in Lapland, the northernmost part of Finland, the snow ploughing costs dominate quite markedly the overall wintertime road maintenance costs.

In the following we consider some preliminary results concerning the wintertime road maintenance cost model development for Lapland. To simplify the problem only climatological aspects are included, although also matters, like the traffic density and the road classification, pose important conditions.

2. Data

To assess the needs for snow ploughing we consider the following variables: amount of snowfall (water equivalent in mm) and the relative frequency of at least moderate snowfall events, driving snowstorm incidents, and saltation events. For this purpose we use the synoptic weather station network of the Finnish Lapland as given in fig. 1.

The quite limited amount of the wintertime road maintenance cost data restricts the model development severely. At the moment monthly maintenance costs only for years 1981-83 and 1986 of January, February, March, November, and December have been available. Fortunately from the point of view of wintertime road maintenance conditions all these months can be considered to be winter months in Lapland. Therefore a composited, winter period model was warranted in this case.

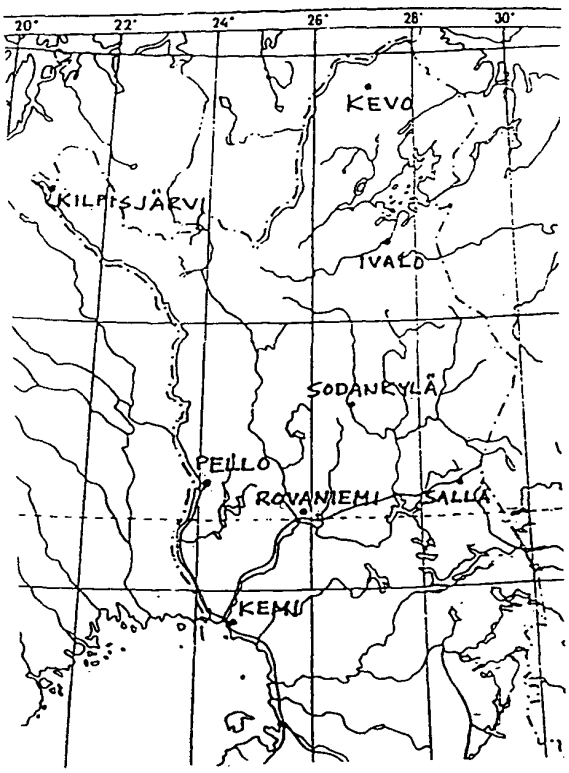


Fig. 1: The seven synoptic weather stations of the Finnish Lapland used in this study.

3. Methods

Taking into account the short wintertime road maintenance cost data record the choice of an appropriate method becomes very important. Even then one can hope nothing more than to cover the most relevant features of the snow ploughing cost problem.

From the statistical point of view there is an obvious need for a robust approach. From the point of view of a multiple linear regression model this means to prefer e.g. the least absolute values or the minimax criterion approach to the classical least squares one (Barrodale and Roberts (1973), Barrodale and Roberts (1974) and Barrodale and Phillips (1975)).

From the climatological point of view there are well-known problems in the quality of visual observations. Concerning the weather events causing needs for snow ploughing the observations of driving snowstorm incidents and saltation become very important. Tests based on physically favourable conditions for these events usually match poorly with the visual observations and moreover, the visually observed frequencies of the events tend to underestimate the true frequencies quite considerably.

In Finnish Lapland the visually observed occurence of driving snowstorm events seems not to be a severe problem in this context, because the precipitation, temperature, and wind measurements reveal these events fairly adequately. On the other hand the saltation, i.e. snow flowing with the wind in a thin layer above the snow surface, seems to be reported quite poorly if at all. From the point of view of the snow ploughing saltation seems to be a major problem on the roads of Lapland, because the roads are usually not elevated with respect to their surroundings and the use of snow fences is quite limited.

The only way to assess the role of the saltation was to parameterize its frequency with respect to temperature and wind speed observations. Based on Gray and Male (1981) the saltation was considered to be strong enough, when

$$T \leq -1.0\,^{\circ}C \qquad (1)$$

throughout the past 48 hours with respect to the observation time and

$$v \geq 5 \text{ m/s}. \qquad (2)$$

The temperature condition (1) prevents to consider cases, where crust causes the saltation to cease and the wind speed condition (2) ensures the saltation to be able to cause traffic problems.

4. Results and discussion

In the calculations the multiple linear regression model based on the least absolute values criterion was used. The dependent variable was the monthly road snow ploughing cost of the Lapland road district. As independent variables the precipitation (snowfall amount; mm) and the relative monthly number of occurences of at least moderate snowfall, driving snowstorm, and saltation were considered.

The stepwise calculations revealed quite clearly that the effect of saltation cannot be ignored. The results were fairly sensitive to the variation of the limit values in conditions (1) and (2). However, the values chosen were based on the best estimates available.

The results of the calculations are given in fig.2.

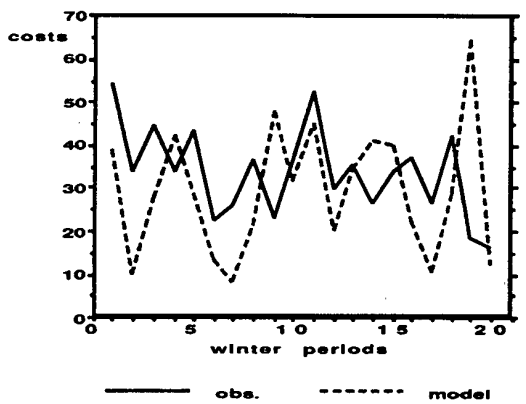


Fig. 2: The observed and model calculated snow ploughing costs for the Lapland road district. The winter periods 1-20 refer to January, February, March, November, and December of 1981-83 and 1986. The costs are given in 10^5 Fmk.

The results showed by no means a perfect agreement, but according to the preliminary results due consideration for preventive actions against the effects of the saltation seemed to be warranted. To obtain a better agreement with the observed snow ploughing costs also the wintertime road maintenance policies should be considered. These could have some reflectations to the representativeness areas of the different weather stations considered. In fact this representativeness problem should take into account the road network, the traffic density, the classification of various roads etc.. These aspects stand for the conditions the user puts on the problem. So far they have been taken into account only quite crudely.

However, the main problem is the restricted availability of the wintertime road maintenance cost data and also the fact that only monthly data could be used. Therefore the costwise effects of the different weather phenomena under consideration could not be discussed separately. Only some suggestions could be given.

5. Summary

Based on a robust multiple linear regression approach (the least absolute values criterion) the dependence of the snow ploughing costs on various weather phenomena was considered. Taking into account the short cost data records the main conclusion was to emphasize the importance of the saltation in this context.

Further considerations would need both a broader view of the various aspects affecting the cost formation and a more detailed , e.g. daily, knowledge of the costs.

Acknowledgements: During this investigation the author has had fruitful discussions with Mr. Erkki Nysten (The Finnish Meteorological Institute) and Dr. Esko Kuusisto (The Hydrological Office). Also the help of Mr. Ari Venäläinen (The Finnish Meteorological Institute) and Mr. Olli Penttinen (Finnish

National Road Administration) is gratefully acknowledged.

References:

Barrodale, I., and F. D. K. Roberts (1973), An improved algorithm for discrete L_1 approximation, SIAM Journal on Numerical Analysis, 10, 839-848.

Barrodale, I. and F. D. K. Roberts (1974), Solution of an overdetermined system of equations in the l_1 norm, Communications of the ACM, 17, 319-320.

Barrodale, I., and C, Phillips (1975), Algorithm 495: Solution of an overdetermined system of linear equations in the Chebyshev norm, ACM Transactions of Mathematical Software, 1, 264-270.

Gray, D. M., and D. H. Male (eds.) (1981), Handbook of Snow, Pergamon Press, Toronto, Canada.

Robust Procedures for Multivariate Data Analysis

Todorov, V.K., Neykov, N.M. and Neytchev, Pl.N.
National Institute of Meteorology and Hydrology
Bulgarian Academy of Sciences,
Trakia blvd, Sofia 1184, Bulgaria

Some robust multivariate statistical techniques are considered. Special attention is devoted to principal components, cluster and discriminant analysis. The behavior of the robustified procedures is studded via artificial data sets and Monte Carlo simulation. It is concluded that some of them perform well.

1. Introduction

The classical mean **m** and covariance matrix **S** usually constitute the basis for many classical multivariate procedures, in particular discriminant and cluster analysis, principal component analysis, factor analysis, canonical correlation and correspondence analysis, outlier detection, hypothesis testing etc. If outliers (atypical values, anomalous observation, gross errors) are present in the data they will influence the estimates **m** and **S** and further will worsen the performance of the classical multivariate procedures based on these estimates. Therefore the problem of robust estimation of **m** and **S** has been given extensive treatment in the recent years (see Huber (1981), Hampel et al (1984), Rousseeuw and Leroy (1987)).

Although there exists a lot of theoretical results the robust covariance matrices and the robustified by them multivariate procedures are not widely used. Our efforts were directed mainly towards robustifying several well known classical techniques of multivariate statistical analysis, studying their behavior on simulated and real data and developing relevant, practice-oriented software. Some of the obtained results are presented in this paper.

2. Robust Means and Covariance Matrices

Let consider a point cloud of n data points in the p-dimensional space, represented by the vectors $x_1, x_2, \ldots, x_n \in R^p$. The classical measures of the center of the point cloud and the scatter of the points about this center are the usual sample mean vector **m** and the sample covariance matrix **S**.

A class of robust alternatives to the sample mean and covariance matrix are the M-estimates, defined by Maronna (1986). From practical point of view the M-estimates of multivariate location and scatter can be considered as weighted version of the sample mean covariance matrix, the weights being determined by the data. The robust M-estimates of location M and scatter C are computed iteratively by:

$$M = \sum_{i=1}^{n} w_1(d_i)\, x_i \Big/ \sum_{i=1}^{n} w_1(d_i)$$
$$C = \sum_{i=1}^{n} w_2^2(d_i)(x_i-M)(x_i-M)^T \Big/ \left(\sum_{i=1}^{n} w_2^2(d_i)-1\right) \quad (1)$$

where the weights, w_1 and w_2, depends on Mahalanobis distance:

$$d_i^2 = (x_i - M)^T C^{-1} (x_i - M), \quad (2)$$

One choice for w_1 and w_2 is discussed by Campbell (1980):

$$w_1(d_i) = \omega(d_i)/d_i$$
$$w_2(d_i^2) = [w_1(d_i)]^2$$
$$\omega(d) = \begin{cases} d & \text{if } d \le d_o \\ d_o \exp\{-\frac{1}{2}(d-d_o)^2/b^2\} & \text{if } d > d_o \end{cases} \quad (3)$$

with $d_o = \{\sqrt{(2p-1)} + c)\}/\sqrt{2}$ and $c = 2.25$.

If $b=\infty$ Huber's weights (HUB) are obtained. If $b=1.25$ Hampel's type weights (HAM) are obtained.

Maronna (1976) defined the weights (MAR):

$$\omega(d_i) = (p+1) / (1 + d_i^2)$$
$$w_1(d_i) = \omega(d_i) = w_2(d_i^2) \quad (4)$$

The sample mean **m** and covariance matrix **S** can be taken as initial values for the iterative process and it terminates when a given convergence criterion is satisfied (see Devlin et al. 1981).

A disturbing aspect of the M-estimation approach is that the breakdown point is no more

than 1/(p+1) (p is the number of variables) as pointed out Hampel et al (1986). When a high level of contamination is expected it is appropriate to use High Breakdown Point (HBP) estimators. Such estimators are the Minimum Volume Ellipsoid (MVE) and the Minimum Covariance Determinant (MCD), described by Rousseeuw and Leroy (1987), which can cope with up to 50% of outliers.

The MVE and MCD estimators for a data set of n points in p dimensions are defined as follows:

$M_{MVE}(X)$ = center of the minimum volume ellipsoid covering (at least) $h=[n/2]+1$ points of X,

$M_{MCD}(X)$ = mean of the $h=[n/2]+1$ points of X for which the determinant of the covariance matrix is minimal.

Both MVE and MCD yield a robust covariance estimate at the same time:

$C_{MVE}(X)$ = the minimum volume ellipsoid itself, multiplied by a suitable factor,

$C_{MCD}(X)$ = classical covariance matrix of the selected h points multiplied by a suitable factor.

Rousseeuw and van Zommeren (1990) recommend these estimates to be used as initial values for a reweighted estimator or the M-estimates. They described several algorithms based on random search for computing an approximate MVE and MCD estimates. In order to unmask multivariate outliers and leverage points they propose the classical mean and covariance in the expression of the Mahalanobis distance to be replaced by their HBP analogs. Neykov and Neytchev (1991) implement a version of their algorithm using the widely distributed BMDP package. An alternative approach to computing the MCD estimates, using a general purpose optimization algorithm known as "simulated annealing", is proposed by Todorov (1991). A variety of classical, robust M, HBP and their one-step improvements may be obtained interactively by the computer program ROCOV written by Todorov et al. (1990b).

If a grouping structure present in the data set, robust group means and covariance matrices can be computed for each of the groups and then the obtained with them weights can be used for computing robust within-group, between-group and pooled covariance matrices. Details can be found in Campbell (1982) or Todorov et al (1990a):

3. Robust Principal Component Analysis

Principal component analysis is widely used techniques for reducing of dimensionality of multivariate data. It is essentially based on computation of eigenvalues and eigenvectors of the sample covariance or correlation matrix and therefore its results can be extremely sensitive to the presence of even a few atypical observations in the data. Details about principal component analysis subjected to robustification may be found in Devlin et al (1981), Campbell (1982), Zhou (1989), Branko (1990), Naga (1990).

Zhou (1989) has published a computer program ROPCA computing robust estimates of principal components based on the covariance matrix given by one of the three robust methods: multivariate trimming (MVT), M-estimates using Huber weights or T-weights introduced by Devlin et al, (1981). With minor changes to the original program, we enhanced it by adding options for computing several HBP estimators (and called it HROPCA). To illustrate the results produced by this program we consider the following example. A data set of moderate dimension (p = 6) is generated from symmetric contaminated normal distribution SCN(**0**, **P**) with a given covariance matrix **P** and zero mean. The eigenvalues of **P** are shown in the first row of Table 1 and SCN is obtained as a mixture:

$$SCN_p(0,P) = (1-\alpha)NOR_p(0,P) + \alpha NOR_p(0,K*P)$$

where NOR(**0**, **P**) is multivariate normal distribution with zero mean and covariance P.

Table 1. Eigenvalues for the generated data. OLS* stays for OLS estimation over clear data.

Method	V1	V2	V3	V4	V5	V6
target	2.029	1.499	1.4999	.943	.028	.0002
			α = 0.10, K = 9			
OLS*	2.3980	1.5418	1.2302	.8069	.0212	.0002
OLS	3.8987	2.7428	1.6425	1.0269	.0270	.0002
MVT	2.6133	1.5176	1.2159	.6383	.0224	.0002
MAR	4.2816	3.3789	1.7885	1.5546	.0515	.0004
HUB	3.1649	2.6447	1.3956	1.0245	.0268	.0002
MVE	2.6952	2.2546	1.2329	.6194	.0184	.0002
			α = 0.25, K = 25			
OLS*	2.3501	1.4956	1.0207	.6737	.0208	.0001
OLS	18.3964	11.4121	8.8378	4.0910	.1067	.0007
MVT	10.4347	5.0113	2.6802	1.1678	.0345	.0005
MAR	12.3433	10.2774	7.3129	4.4876	.1312	.0013
HUB	11.1754	8.8458	6.8622	3.1474	.0848	.0006
MVE	2.6889	1.5905	1.0338	.7211	.0242	.0003

The fraction of contamination α was chosen as 10% or 25% and the variance inflation factor K is 9 or 25 (the second choice results in 38 "good" and 12 "bad" observations).

In Table 1 are presented the results obtained with HROPCA, using classical and some robust methods of estimation over the generated data sets. By HUB and MAR are denoted M-estimators using weighting functions (3), resp. (4). While the M-estimates (MAR, HUB) and MVT perform well in cases of moderate contamination, in cases of higher contamination the results obtained with MVE technique are more adequate.

4. Robust Discriminant Analysis

The classical Linear Discriminant Analysis assign an object to one of g homogeneous populations on the basis of p-dimensional feature vector x by the Linear Discriminant Functions:

$$d_i(x) = -(x-\mu_i)^T \Sigma^{-1}(x-\mu_i) + \ln(\alpha_i), \qquad (5)$$

where $i=1,\ldots,g$; α_i is the prior probability that an object comes from population P_i; μ_i and Σ are the sample means and pooled covariance matrix. If there are outliers in the training data set they will influence the estimates, thus decreasing the quality of the classification. A number of robust versions of LDF exist, reducing the influence of outliers on the quality of discrimination, but none of the widely available computer packages possesses options for robust discrimination.

If only two groups are considered one may take advantage of the fact that linear discriminant analysis and multiple linear regression are numerically equivalent in this case. Robust regression to device a robust discriminant analysis can be used. This approach is ready-to-use in practice, since several computer programs performing bounded influence (GM) linear regression are available (see Dutter (1988), Marazzi (1987), Samarov and Welsch (1982), Neykov and Neytchev (1989), Neytchev et al (1990)). Robust selection of the most discriminative variables can be easily achieved as point out Krusinska and Liebhart (1988). In Todorov et al (1991) the performance of various GM-estimators, as applied to the problem of discrimination, in small sized training samples is investigated through the use of a Monte Carlo simulation. It is concluded that the GM-classifiers perform well. The main disadvantage of GM-estimators is their low breakdown point. So LDF based on HBP estimators of location and scatter is proposed by Todorov et al (1991).

Usually LDA is performed in a stepwise manner in order to select the subset of the most discriminating variables. A common measure used for this purpose is the Wilks' lambda statistic:

$$\text{Wilks' lambda} = \frac{\det W}{\det T}, \qquad (6)$$

where W is the within sums-of-crossproducts matrix and T is the total sums-of-crossproducts matrix. Because of the low breakdown point of M-estimators, when a large fraction of outliers is expected it is more appropriate to compute the lambda statistic relative to MVE or MCD as was proposed in Todorov et al (1990). The determinants of W and T = W + B in (6) are replaced by their robust analogues. Todorov et al. (1990) showed that the HBP analog of Wilks' lambda, used over contaminated data yielded the same order of entering variables as the classical one used over the original data.

5. Cluster Analysis

One of the best known techniques for partitioning a data set into k groups is the minimization of the within cluster sum of squares. If the cluster center is represented by its mean, the presence of outliers in the data may be very harmful to the partitioning. Whatever the reasons were for the presence of atypical observations in the data, these observations will distort the formation of clusters, since they will be also assigned to the nearest cluster. Furthermore, they will influence the estimation of the centers of the clusters as well as the estimate of the weighting matrix W if it is not taken to be the identity. In practice the outliers usually show up as clusters consisting of only one or two data units. Such atypical features of the data set tend to worsen the performance of any clustering algorithm and especially those based on the variance criterion Kaufman and Rousseeuw (1990).

In Todorov et al (1991) a partitioning procedure is proposed, which is similar to the well-known *K-means* algorithm, but instead of the ordinary estimates of the mean vectors their M-estimates are used. Therefore it is named K-M-means procedure. One way to illustrate the usefulness of a robust approach is to introduce artificial outliers into an existing data set and study their effect. Our example is based on the well-known Fisher's Iris data which has been widely used as a test data in cluster analysis. (3 groups - *Iris Setosa*, *Iris Versicolor* and *Iris Virginica* with 50 observa-

tions each, represented by 4 measurements - *sepal length, sepal width, petal length* and *petal width*).

In Table 2 the configuration of outliers is presented (the original values and the ones they were substituted with). The ninth of the table shows the value of the weight given to the corresponding outlier by HUB (HAM assigns all of them zero weights). All the rest observations get unit weight.

Table 2. Original data, outliers and weights.

Group	Original				Substituted				Weights
	X1	X2	X3	X4	X1	X2	X3	X4	(HUB)
1	4.9	3.1	1.5	0.2	4.9	3.1	16.5	0.2	.153
	5.0	3.3	1.4	0.2	5.0	12.3	1.4	0.2	.116
	5.1	3.8	1.6	0.2	5.1	3.8	1.6	10.2	.493
	5.4	3.4	1.7	0.2	5.4	3.4	1.7	29.2	.190
3	6.5	3.0	5.2	2.0	16.5	3.0	5.2	2.0	.172
	7.4	2.8	6.1	1.9	7.4	2.8	6.1	21.9	.233

In Table 3 are presented the reclassification tables for CKM, HUB and HAM. The association of the obtained partition with the true one is assessed by a measure presented by Rand (1971) (noted by *).

Table 3. Reclassification tables obtained with the classical KM algorithm and with M-K-algorithm with Huber (HUB) and Hampel (HAM) weights.

CKM (0.734*)			HUB (0.864*)			HAM (0.855*)		
47	2	1	47	0	3	47	1	2
3	47	0	0	50	0	0	48	2
0	40	10	0	15	35	0	14	36

REFERENCES:

Branco, J.A., and A.M.Pires, 1990: Robust estimation of principal components, *COMPSTAT'90 Short Communications*,95-96.

Campbell, N.A., 1980: Robust procedures in multivariate analysis I: Robust covariance estimation, *Appl. Statist.* **29**, 231-237.

Campbell, N.A., 1982: Robust procedures in multivariate analysis II: Robust canonical variate analysis, *Appl. Statist.* **31**, 1-8.

Devlin, S.J., R. Gnanadesikan, and J.R. Kettenring, 1981: Robust estimation of dispersion matrices and principal components, *J.Amer.Statist.Assoc.* **76**, 354-362.

Dutter, R., 1987: BLINWDR: A computer program for robust and bounded influence regression, *Statistical Data Analysis based on the L_1-norm and related methods*, Y.Dodge, Ed., North Holland,

Hampel, F.R., E.M. Ronchetti, P.J. Rousseeuw and W.A. Stahel, 1986: *Robust Statistics. The Approach Based on Influence Functions*, Wiley, New York.

Huber, P.J., 1981: *Robust Statistics*. Wiley, New York.

Kaufman, L. and P. Rousseeuw, 1990: *Finding Groups in Data*, Wiley, New York.

Krusinska E., J. Liebhart, 1988: Robust selection of the most discriminative variables in the dichotomous problem with application to some respiratory disease data, *Biometrical Journal* **30**, 295-304.

Marazzi, A., 1987: Solving bounded influence regression problems with ROBSYS, *Statistical Data Analysis based on the L_1-norm and related methods*, Y.Dodge, Ed., North Holland,

Maronna, R.A., 1976: Robust M-estimators of multivariate location and scatter, *Annals of Statistics* 4, 51-67.

Naga, R. A., G. Antile, 1990: Stability of robust and non-robust principal components analysis, *Computational Statistis and Data Analysis*, **10**, 169-174.

Neykov N. M. and Pl. N. Neytchev, 1988: Robust and bounded-influence regression using BMDP3R and BMDPAR, *COMPSTAT'88 Short Communications and Posters*. Copenhagen, Denmark.

Neykov N. M. and Pl. N. Neytchev, 1991: Unmasking multivariate outliers and leverage points by means of BMDP3R, *Directions in Robust Statistics and Diagnostics, Part II*, W. Stahel and S. Weisberg, Eds, Springer-Verlag.

Neytchev, P. N., N. M. Neykov and V. K. Todorov, 1990: Fitting models to data by REGRESS-PC program system, *Stochastic Methods in Experimental Sciences*, W. Kasprzak, and A. Weron, Eds, World Scientific Publishing Co, Singapore.

Rand, W. M., 1971: Objective criteria for the evaluation of clustering methods, *J.Amer.Statist.Assoc.* **66**, 846-850.

Rousseeuw,P.J. and A. M. Leroy, 1987: *Robust Regression and Outlier Detection*, Wiley, New York.

Rousseeuw, P.J., and B. C. van Zommeren, 1990: Unmasking multivariate outliers and leverage points (with discussion), *J.Amer.Statist.Assoc.* **85**, 633-643.

Samarov, A., and R. E. Welsch, 1982: Computational procedures for bounded influence regression, in H. Caussinus, P.Etinger, and R. Tomassone Eds., *Proceedings in Computational statistics, COMPSTAT'82*, Physica Verlag, Heidelberg.

Todorov, V. K., 1991: Computing the minimum covariance determinant estimator (MCD) by simulated annealing, *Computational Statistics and Data Analysis* (in press).

Todorov, V. K., N. M. Neykov and P. N. Neytchev, 1990a: Robust variable selection in the discriminant analysis based on MVE and MCD estimators, in Momirovic, K. and V. Mildner, Eds., *Proceedings in Computational statistics, COMPSTAT'90*, Physica Verlag, Heidelberg.

Todorov, V. K., N.M. Neykov, and P. N. Neytchev, 1990b: A program for computation of robust covariance matrices, *Proceedings of DIANA III*, 4-8 June 1990, Becine, Czechoslovakia.

Todorov, V. K., N. M. Neykov and P. N. Neytchev, 1991a: Robust two-group discrimination by bounded influence regression, *Computational Statistics and Data Analysis* (submitted).

Todorov, V. K., N. M. Neykov, and P. N. Neytchev, 1991b: K-M-means: A robust procedure for clustering by means of robust M- estimators, a manuscript.

Zhou, Di., 1989: ROPCA: A FORTRAN Program for robust principal components analysis, *Computers & Geosciences* **15**, 59-78.

A NEW TOOL TO DISCRIBE THE GLOBAL ATMOSPHERE —RELATIVE DISTRIBUTION FUNCTION

Zhang Xuewen Ma li

Meteorological Institute of Xinjiang ,Urumqi, China, 830002

1.The distribution of molecular speed

Suppose there are a great deal of molecules of certain gas in a bottle. According to the kinetic theory of gas, All the gas molecules not only move very quickly but also with various speed.

In 1860, J.C.Maxwell worked out a important relation between the molecular speeds v and the concerned molecular numbers N (Fig.1).

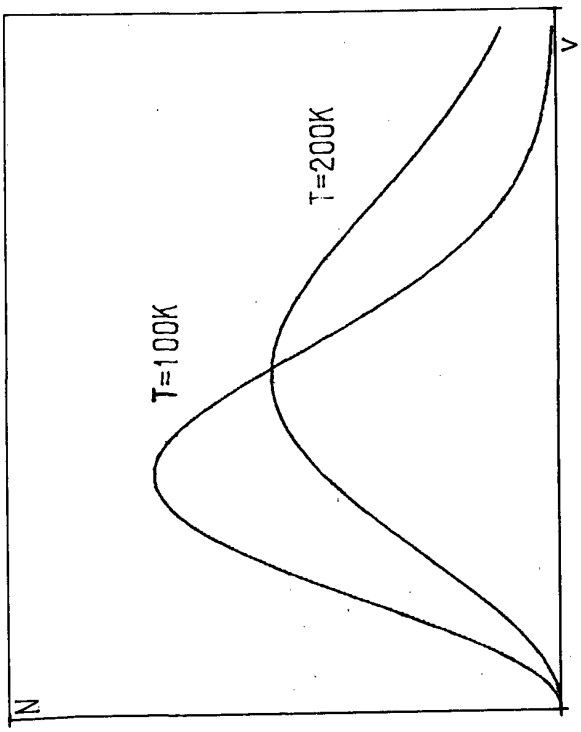


Fig.1 the distribution of molecular speed

We have to point out that the relationship between molecular speeds and molecular numbers is exactly matained at any time, although the speed of each molecule changes very quickly.

2. The distribution of atmospheric speed

We have treated gas in a bottle as a isolated system and get the distribution of molecular speed as Fig.1. Now let us research a similar problem in the atmosphere. From certain point of view, we could treat the global atmosphere as a isolated system also. Now the problem is: could we work out a corresponding curve in the global atmosphere which has a similar meaning with gas?

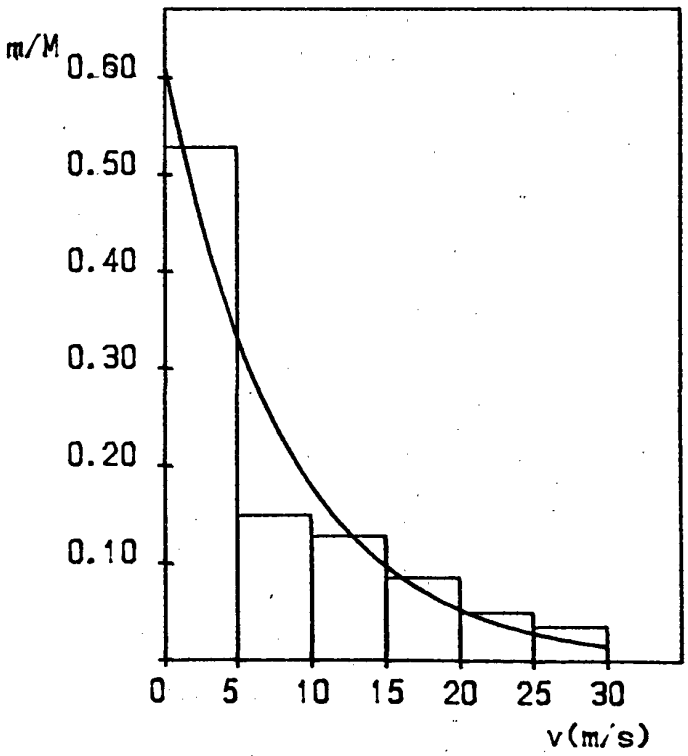


Fig.2 The distribution of wind speed of global atmosphere

After some statistical analyses from the data of global atmospheric wind speed, we get Fig.2. Comparing with Fig.1, horizontal axis of Fig.2 stands for wind speed instead of

molecular speed, and the meaning of veritical axis is the same as Fig.1 which expresses a ratio value of mass which has a given speed v to the mass of total system.

Fig.2 tells us

* There is a very simple curve to express the relationship between the wind speed and the air mass in global atmosphere. We had investigated those curves day by day, even in the cold current period, and discovered this curve is very steady. Obviously this is an important feature of global atmosphere which we have never known.

* We could use a negative exponent equation to express this relationship between the air mass to its wind speed (Zhang, 1990)

$$f(v)=\frac{1}{\bar{v}}\exp(-v/\bar{v}) \text{-------(1)}$$

Here the f(v) expresses the ratio of air mass which speed is v to the total mass of atmosphere. $\bar{v}$ is the mean speed of global air mass. It is a statistical parameter of atmosphere.

Maxwell derived a theoretical equation

$$f(v)=\sqrt{2}\pi N\left(\frac{m}{\pi kT}\right)^{\frac{3}{2}} v^2\exp\left(-mv^2/2kT\right)$$
$$\text{----------(2)}$$

to describe the distribution of molecular velocity. So Eq.1 is the Maxwell equation in atmosphere.

3. Relative distribution function

If we analyse the global atmosphere as above but change the variable wind speed to other meteorological elements such as air temperature T...., then we can get many other curves. Fig.3 and Fig.4 are two examples.

In Fig.3,4 the meaning of vertical axis is the same as Fig.1, but one figure stand for the distribution of air temperature, the other one stand for the total energy.

Fig.2-4 are three examples of the distribution of certain meteorological element. Generally to say, we could research the distribution of any element if we like. Base on this idea, we introduce a general concept to describe the distribution feature of any material system and we call it(Zhang, 1986)

RELATLVE DISTRIBUTION FUNCYION--R.D.F.

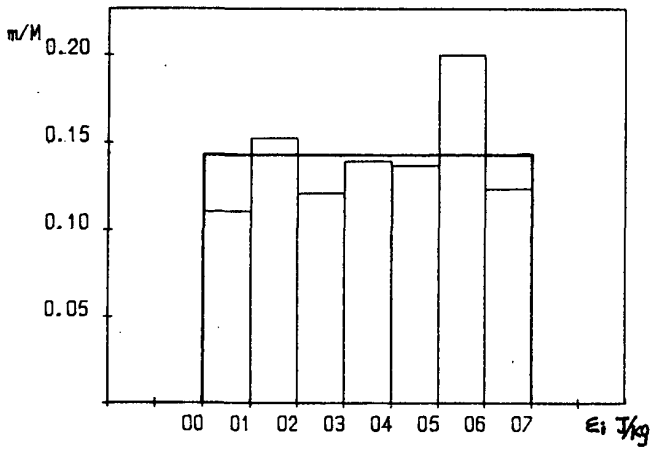


Fig.3 The distribution of temperature of global atmosphere

If there is a material system (relative isolate), whose total mass is **M**. Then we could partition it into many parts (such as partitioning gas into a great deal of molecules). For each part(each partical), it has a value for variable x(such as wind speed v or temperature T ...). Then the relative distribution function--R.D.F is such a function $f(x)$ which describes the relationship between the value of variable x and its corresponding relative mass $dm/(Mdx)$. So we could write

$$f(x)=dm/(Mdx) \text{----------(3)}$$

In statistical physics, there are many relative distributions researched (as Eq.2), but in meteorology they have been ignored for a long time, (but cloud specturn is an exception).

We think the study of R.D.F's of various meteorological element is an important task for us.

4. Some results

We had worked out many R.D.F's of meteorological element, some of them were gathered in Tab.1.

Tab.1 The distribution types of R.D.F of various meteorological element

No.	meteorological element	distribution type
1	speed distribution	negative exponent
2	pressure distribution	uniform
3	temperature distribution	uniform
4	specific humidity	negative exponent
5	potential temperature	Gamma (n=2)
6	interal energy	uniform
7	potential energy	negative exponent
8	total energy	Gamma (n=3)

To work out the R.D.F.'s as Tab.1, we utilize global data from surface to upper air, from the polar to the equator(Zhang,1992). Here we have to point out that for geting a R.D.F., only the given time data needed.

Tab.1 display that all R.D.F.'s are continuous function. we think this is a good character. We proved R.D.F. is on an equality with probability density distribution function. So, we named many R.D.F.'s with probability language such as exponent,uniform, Gamma,... distribution.

Because we can get a R.D.F. for any meteorological element at any given time, so it excite a problem: Does the R.D.F.'s of different time has the same feature ? After the analyse, we discover the feature of R.D.F. is very steady. that is to say

* The form (type) of analytic function of R.D.F. does not change day by day even season by season.

* In any R.D.F., there are several statistical parameters. These parameters change a very small percentage day by day. We could say that the weather disturbance can't affect the figure of R.D.F. seriously.

5. Maximun entropy principle

R.D.F. is not only a good tool to analyse the real data, but also a bridge to guide us to a theory field--Maximun Entropy Principle.

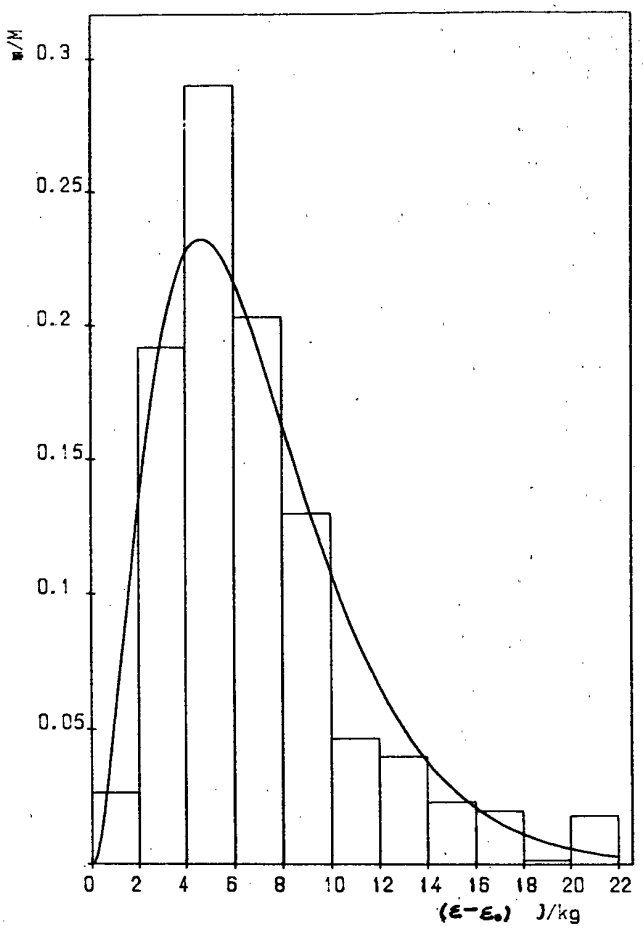


Fig.4 The distribution of total energy of global atmosphere

In information theory, we have known that the entropy value H(x) may be defined

$$H(x)=-\int_{-\infty}^{\infty} f(x)\ \ln f(x)dx \qquad \text{----------(4)}$$

Here x is a random variable, $f(x)$ is its probability density distribution function. Because this function is on equality with R.D.F., so we could let $f(x)$ stand for a R.D.F. Thus for every R.D.F., we could worked out an entropy value by Eq.(4). Base on this idea, we could connect R.D.F. with a concered entropy value.

At the same time we have known that there is an important law about entropy. In physics, this is the second law of thermodynamics,in the information theory,this is the maximun entropy principle. Because we have connected R.D.F. with entropy concept,so we could connect it with entropy principle further.

Since 80's we had utilized this principle to illstrate seveal meteoroloical problems (Zhang, 1987, 1991a, 1991b). For save words we can't introduce them in detial.

We have pointed out the steadiness of R.D.F. has an obvious feature and known in physics. If a material system achieve its equilibrium status, its entropy achieve its maximun value. Then, this system will stay at this status for ever.

So we belive the appearance of the steadiness of R.D.F. in a material system is controled by the entropy principle, Whether it is a physical system as molecular speed, or a meteorological system as wind speed of global atmosphere etc..

6. Weather forecast and climatic change

In atmosphere, the main priciples are the law of Newton, the conversation of mass, the first law of thermodynamic. Does the second law of thermodynamic--the law about entropy, controls the atmosphere also ?

We think this is an important problem which we haven't researched in detial. Using R.D.F. we could step in this important field.

If we believe the steadiness of R.D.F. is an apperance of the entropy law, we think that we could conbined these equations of R.D.F. with the dynamic equations in the numerical weather forecast. There are a lot of numbercial experments waiting us to test.

In the atmosphere, each R.D.F. is a analystic equation for any given time. We had pointed out there are several marameters in each R.D.F.. If we analyse the R.D.F. season by season we could discover these parameters changed slowly. We could treat this changing as climatic change of interannual. Then, if we study the change of the parameters of R.D.F. year by year or century by century, then it becomes a subject of climatic change. So we have seen the R.D.F. is also a tool to study the climatic change.

7. Conclusion

7.1 RALATIVE DISTRIBUTION FUNCTION--R.D.F is a important concept which have used in statistics physics for a long time, but until now it have not used in the study of atmosphere.

7.2 After some reseach, we discover some important features of R.D.F. of atmospher. It can be expressed by a analytic equation, has seveal parameters, and possess the steadiness.

Tab.1 are some R.D.F's which we had discovered.

7.3 How to introduce the entropy law into atmosphere science is a important problem. The R.D.F. is a effective tool to do this.

7.4 The R.D.F. is a tool for numberical weather forecasting and for the study of climatic change.

Reference

Zhang Xuewen, 1988, Relative distribution function and meteorological entropy, ***Acta Meteorologica Sinica,*** **44** (2), 214-219

Zhang Xuewen, 1987, The distribution rule of the sepcific humity in global atmosphere, ***Acta Meteorologica Sinica,*** **45** (2), 251-253

Zhang Xuewen and Zheng Guoguang, 1989, Steady stratus cloud's droplet spectrun equation derived from entropy principle, ***Plateau Meteorology,*** **8,** (3), 273-278

Zhang Xuewen and Ma Li, 1990, Energy distribution and wind velocity distribution in atmosphere, ***Xinjiang Meteorologicval Monthly,*** **13,** (1), 4-10

Zhang Xuewen and Yang Xiusong, 1991, The relationship between the area and the depth of the storms derived from entropy principle, ***Plateau Meteorology,*** **10,** (3), 225-232

Zhang Xuewen, el., 1991, Rain rate intensity-duration formula derived from entropy priciple, ***Scientia Atmospherica Sinica,*** **15,** (6), 17-25

Zhang Xuewen and Ma Li, 1992, Entropy Meteorology, Meteorological Press, Beijing, (in press)

Application of Statistical Techniques in Hydrology Illustrated with Case Studies from Malabar Coast, India

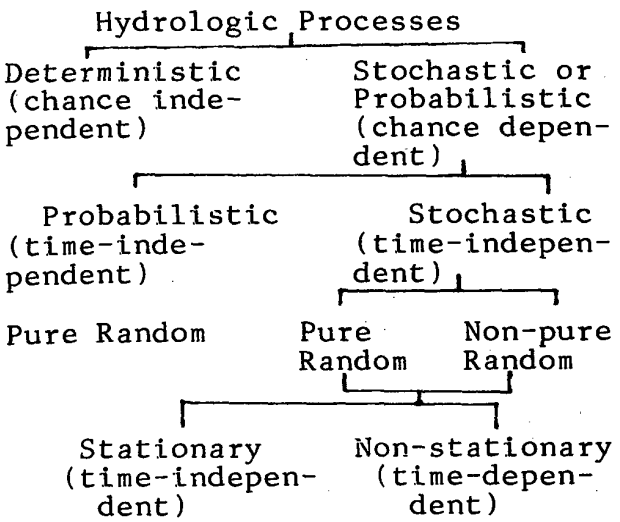
E.J. JAMES
Centre for Water Resources Development and Management, Kozhikode, India

ABSTRACT

The application of statistical techniques in hydrology is highlighted with case studies from Malabar Coast in India. The frequency analysis of rainfall, flood and lowflow and the regionalisation techniques have been illustrated. The application of regression and correlation in hydro-geomorphologic, lowflow, sedimentation and estuarine studies on the Malabar Coast is furnished. The role of statistics in rainfall-runoff modelling has also been briefly highlighted with examples.

1. Introduction

The hydrologic processes in a particular environment are either stochastic or combined deterministic-stochastic processes; the demand for water is also a random variable (Yevjevich 1972). The hydrologic processes may be classified as given below:



The hydrologic data series may be broadly classified into complete duration, partial duration and extreme value. Certain case studies are presented to illustrate the application of statistics in hydrology. The area of the Malabar Coast considered is about 40000 sq.km. with 41 west-flowing and 3 east-flowing rivers which are short, monsoon-fed and flowing through a steep topography.

2. Frequency Analysis and Regionalisation

One of the important problems in hydrology deals with interpreting past record of hydrologic events in future probabilities of occurence (Chow 1964). The frequency analysis is generally used in estimating frequencies of rainfall, floods, droughts, water quality and waves. Regionalisation of such results are also possible.

A generalised map of 1-day rainfall for 2 year return period for the Malabar Coast is given in Fig.1; for all durations and return periods studied, higher rainfall is observed in the

highland zone of the eastern belt. The rainfall depth-duration-frequency relationship for the southern part of the coast is given in Fig.2. A typical flood frequency curve using Gumbel's distribution for the northern part of the coast is given in Fig.3. The flow-duration curve for a typical sub basin (40 sq.km.) of the region is given in Fig.4.

3. Regression and Correlation

The relationships of variables in hydrology may either show the cause and effect at one hydrologic site or correlation of the effects at the neighbouring sites. Simple linear and curvilinear regression and correlation are the tools generally used by hydrologists.

A simple relationship between elongation ratio and monthly (summer) mean/peak flow is given in Fig.5. In a lowflow study for the

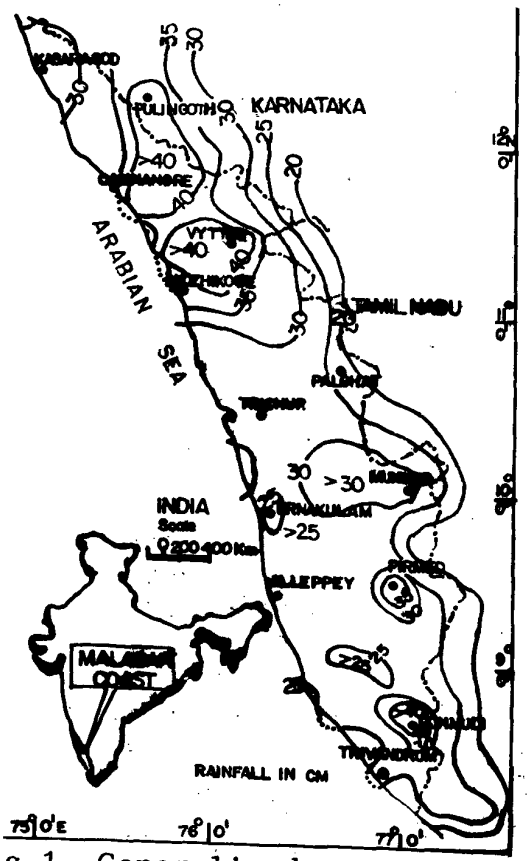


Fig.1. Generalised map of 1-day rainfall for 2-year return period

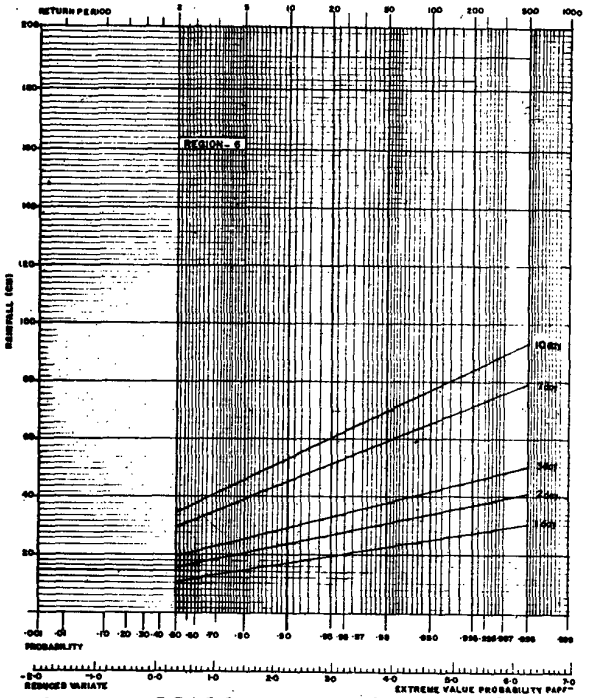


Fig.2. Rainfall-DDF relationship for southern Malabar region

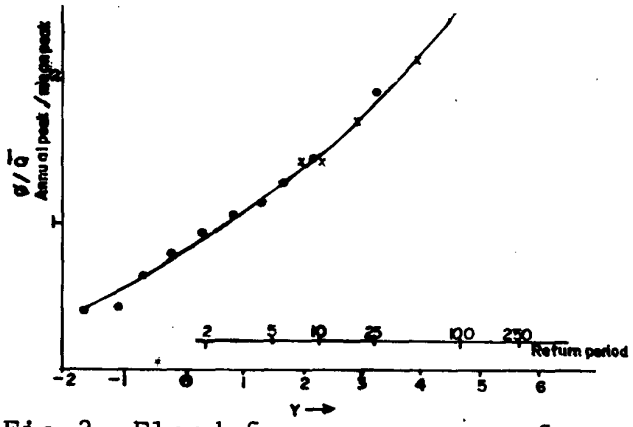


Fig.3. Flood frequency curve for northern Malabar region

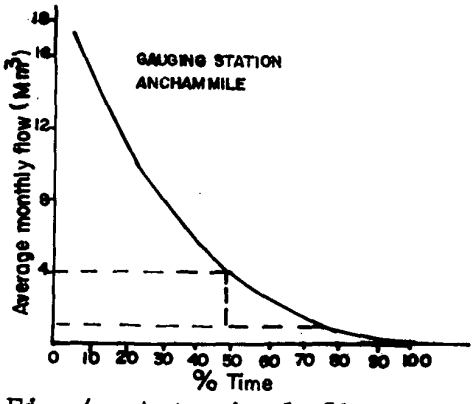


Fig.4. A typical flow-duration curve

region, a regression equation in the following form was obtained for Q 95(1), the 95 percentile 1-day flow:

$$Q\ 95\ (1) = -\ 30.7+21.3\ BFI\ +0.0023\ A\ +\ 0.043\ S1085\ +\ 0.0046\ SAAR \quad (R=0.85)$$

in the above equation, BFI is the baseflow index, A the area of the basin, S 1085 the stream slope measured between 10% and 85% upstream and SAAR the average annual rainfall.

The sediment rating curve obtained for one of the tributaries of the Periyar (40 sq.km.) is given in Fig 6.

The following relationship was established between surface salinity at a cross-section of the Beypore estuary and the river discharge:

$$S_{sf} = 3.9649-0.5314\ R_6$$

in which R_6 is the average discharge of six days ending on the day of sampling.

4. Rainfall-runoff Models

The statistical techniques play a major role in conceptual rainfall-runoff models; stochastic models are also extensively used in hydrology.

In a simple regional monthly model attempted for the river basins of Malabar Coast, rainfall-runoff relationships are obtained in a quadriatic form. The equation is treated as a multiple linear regression with three variables, having a measure of association, namely the multiple correlation coefficient which is the ratio of standard deviation of the estimated runoff to that of the observed (Minikov and Rao 1983). The coefficients are estimated by the method of least squares. The significance of statistical dependence between the rainfall and runoff is checked by the significance of correlation. For regionalisation, the coefficients are plotted against monthly runoff coefficients divided by the unit area of the basin and are plotted against corresponding monthly rainfall values which are also divided by the unit area of the basin. The best fit is then arrived at; Fig.7 shows the application of such a model in an adjacent sub basin. The derivation of unit hydrograph following the approach of Nash (1959) employs the method of moments to determine the two parameters, namely the number of equal linear reservoirs and the storage coefficient for each linear reservoir, which parameters determine the shape of unit hydrograph (Fig.8). The simulation of daily runoff was done using the TANK model (Sugavara 1967). The result for a sub basin of Chaliyar is shown in Fig.9.

The use of stochastic model, such as that of Thomas-Fiering (1962) for the Chaliyar basin is illustrated in Fig.10.

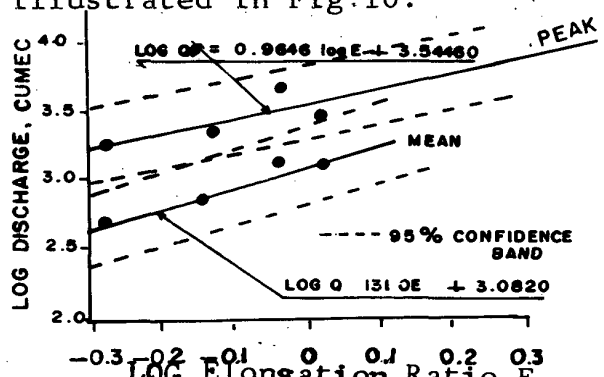


Fig.5. Relationship between stream-flow and elongation ratio

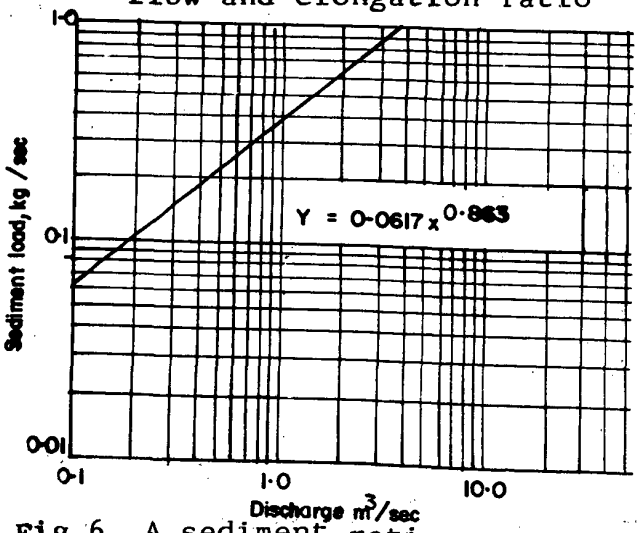


Fig.6. A sediment rating curve

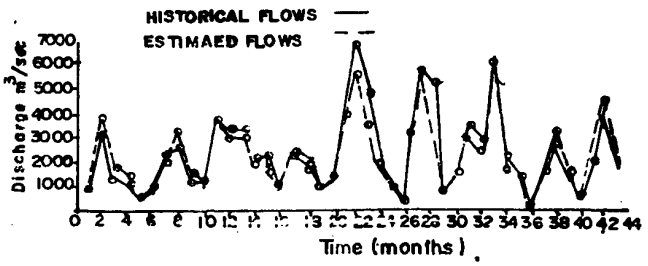


Fig.7. Application of regional monthly rainfall-runoff model

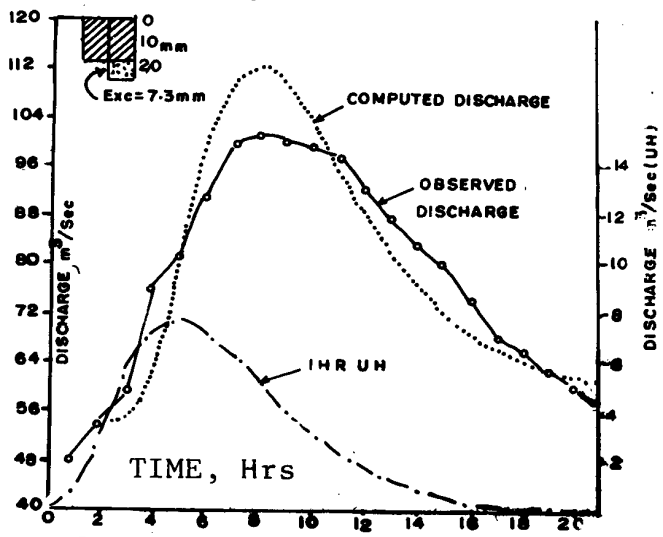


Fig.8. Unit hydrograph-Nash's approach

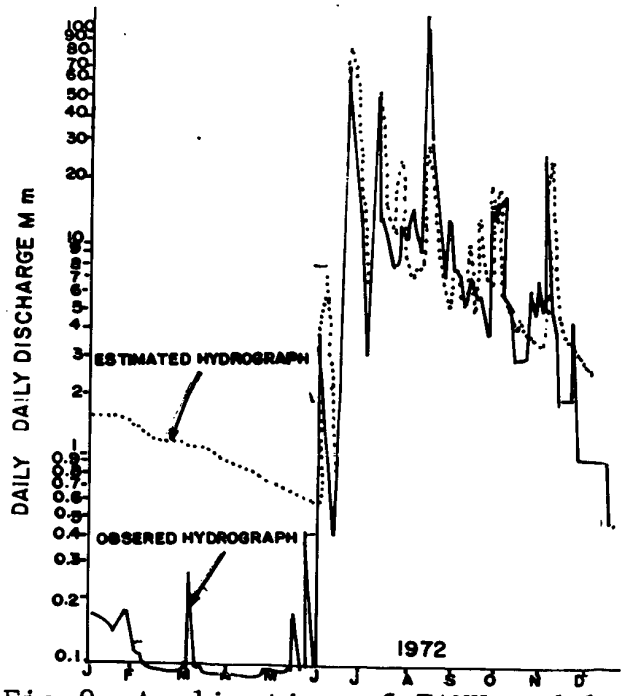


Fig.9. Application of TANK model

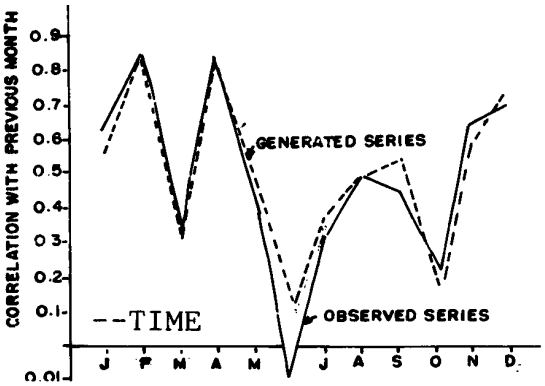


Fig.10. Thomas-Fiering model

5. Conclusions

In order to highlight the role of statistics in hydrologic studies, research carried out on the Malabar Coast with special reference to frequency analysis of rainfall, floods, lowflows, and application of regression and correlation in hydro-geomorphologic lowflow, sedimentation and estuarine studies, has been presented. Certain conceptual and stochastic models tested for the region also have been presented.

REFERENCES

Chow V.T., 1964: Handbook of Applied Hydrology, McGraw Hill Book Company, New York, section 8.

Nash, J.E., 1959: Systematic determination of unit hydrograph parameters, J.Geophys Res., vol.64, 111-115.

Minikov, M. and A Ramachandra Rao, 1983: Regional monthly rainfall-runoff model, J Water Resour.Plg. Mangt.,ASCE vol.109,no.1 ,75-83.

Sugavara,M.,1967: On the analysis of runoff structure about several Japanese rivers, J.Geophys., Japan, vol.33, no.4.

Thomas, H.A. and M.B. Fiering, 1962: Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation, In Design of Water Resources Systems, Harward Univ. Press, Cambridge, 459-493.

Yevjevich, V., 1972: Prabability and statistics in hydrology, Water Resources Publication, Fort Collins, Colorado, 11-13.

CUB: A General Purpose Coverage Algorithm

Albert R. Boehm

Hughes STX Corporation
109 Masschusettes Ave.
Lexington, MA 02173

1. BACKGROUND

Coverage is defined here as the fraction of a domain that contains some specified event. The domain may be a line, an area, a volume, or a hyper volume. Events in general are correlated within the domain.

For meteorologists, cloud cover is the most familiar coverage distribution. However, coverage can apply to the frequency that a fraction of an area will be below freezing or the fraction of a radar scope that shows echoes. It can be used in time to give the fraction of time an airport is below one mile visibility. Numerous other applications are possible.

Coverage can also be applied to other statistical problems. The distribution of the extreme is the same as the distribution of zero coverage. Similarly, the error in an empirical cumulative distribution (the credibility density function for a given percentile) can be given by a coverage distribution (Boehm, 1988).

Hall (1988) presents a variety of coverage processes and associated formulas. However, the work presented in this paper is influenced more by rank order statistics as detailed by David (1981). The actual path of this present work is the direct result of the prior achivements of Irving Gringorten.

Gringorten (1979) developed graphs for coverage along a line or within an area based on Monte Carlo simulations. Later Burger hand fitted algorithms to a new set of simulations based on a sawtooth wave model (Burger and Gringorten, 1984), Gringorten (1984) and Gringorten and Boehm (1987). Burger (1985) used his fitted algorithms to make a cloud atlas. This cloud cover methodology was favorably reviewed by Henderson-Sellers (1991).

The Burger algorithms work well with sky cover distributions as seen by a surface observer, however, they have some important limitations. They are based on a single correlation function resulting from sawtooth waves of constant wavelength. Moreover, the single wavelength sawtooth correlation does not match too closely empirical cloud-free line-of-sight correlations as observed by satellite nor does it match observed cloud correlations in time.

The goal then was to develop a more general coverage distribution, one that could be used with alternate correlation structures, in time as well as space, and have the capability of generalization by being more closely related to statistical theory.

2. COVERAGE DISTRIBUTION ATTRIBUTES

Coverage is a mixed distribution, that is, it has a continuous density between zero and one but has true point probabilities at zero and at one.

The mean of the coverage distribution, μ_c, is equal to the mean (over the domain) expected probability of the event:

$$\mu_c = \frac{1}{m} \sum_{j=1}^{m} c_j = \frac{1}{m} \sum_{j=1}^{m} \left(\frac{1}{n} \sum_{i=1}^{n} b_{ij} \right) \qquad (1)$$

where m is the number of realizations, and n is the number of points in the domain. b_{ij} is a binary value - one if the event occured at point i during realization j, zero otherwise. In theory, both m and n can be infinite in which case the sumations become integrals.

If the expected probability is the same everywhere in the domain, then the mean probability of a point is equal to the mean of the coverage distribution. The symbol P_0 is often used in this case standing for the probabiltiy of a single point (0 area) being covered.

The variance, σ^2, of the coverage distribution is equal to the mean covariance (over all pairs of points in the domain):

$$\sigma^2 = \frac{1}{m} \sum_{j=1}^{m} \left\{ \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} b_{kj} \, b_{lj} \right\} \qquad (2)$$

For example, if at one time the sky were divided into 100 points and each point was labled cloudy or not (b = 1 or 0) then the terms within the braces {} are the mean covariance between all pairs of points at that time. Then the mean over many times of this mean covariance will equal the variance of the coverage distribution.

Similar, results hold for higher moments. That is, the pth raw moment of the coverage distribution is equal to the mean of the sum of p binary products for all possible sets of p points.

Thus, if the point probabilities and

correlation structure is known, all the moments of the coverage distribution can be calculated. Conversely, if the coverage distribution is known as a function of domain size, a homogenious correlation structure can be completly specified.

3. THE EQUICORRELATED MODEL

While all moments of the coverage distributions can be calculated using the above approach, summing p products as p get larger becomes tedious and time consumming. An approximation was desired that is both reasonably accurate and tractable. The equicorrelated model has these properties.

Consider a set of points in which every pair of points has the same correlation, r_e. For isotropic correlation, this can only occur if all points are the same distance from each other - in two dimensions an equilateral triangle. This geometry is not typical of climatological problems. However, the mean correlation, $\bar{r}_e$, is readily calculated:

NOTE: For clarity in the following formulas r is used for r_e.

$$\bar{r}_e = n + n(n-1)r \qquad (3)$$

where the n term is due to the ones on the diagonal of the correlation matrix. The inverse, $\underline{R}_e^{-1}$, the determinant, $|\underline{R}_e|$ and the eigenvalues, $\grave{e}$, are all known and simple in form (Westlake, 1968, p142):

$$\underline{R}_e^{-1} = \frac{1 + r(n-2)}{1 + r(n-2) - r^2(n-1)} \quad \text{for diagonal elements and}$$

$$= \frac{-r}{1 + r(n-2) - r^2(n-1)} \quad \text{for off diagonal elements} \qquad (4)$$

$$|\underline{R}_e| = r^n\left(\frac{1}{r} - 1\right)^{n-1}\left(\frac{1}{r} + n - 1\right) \qquad (5)$$

$$\grave{e}_i = 1 - r \quad \text{for } i = 1, 2, \ldots, n-1$$
$$= 1 + nr - r \quad \text{for } i = n \qquad (6)$$

also the sum of the squares of the diagonal elements of the equivalent triangular matrix (designated T2 for trace elements squared) is simple:

$$T2 = n(1 - r^2)^2 \qquad (7)$$

The equicorrelated matrix is completely specified by its single correlation value and its size n. To approximate a more general correlation matrix, the mean correlations are of the two matrices set equal, and also one of the above parameters are set equal. Which one is used depends on the ease of calculating it for the general matrix. The T2 parameter is generally easy enough to calculate.

4. COVERAGE FOR THE EQUICORRELATED MODEL

The equicorrelated model can be simulated by the set of equations (David, 1981, p108):

$$Y_i = r^{1/2}X_0 + (1 - r)^{1/2}X_i \quad \text{for } i = 1, 2, \ldots, n \qquad (8)$$

where the X's are independent random standard normal variables and X_0 is common to all i values of the equicorrelated Y's. Without loss of generality, we can let $X_1 \le X_2 \le \ldots \le X_n$ so that the Y's are also ranked. Thus, the ith largest Y is given by the weighted sum of a standard normal variable and the ith largest variable from a set of independent normal variables.

The distribution of the ith largest independent normal variable can found through use of the inverse of the cumuative normal and the beta distribution (David, 1981, p8):

$$X_i = \Phi^{-1}[I_p(i, n-i+1)] \qquad (9)$$

where Φ is the cumulative normal and $I_p(a,b)$ is the cumulative (Incomplete) Beta distribution. While (9) is not to hard to evaluate with approximations readily available on most computer systems, the problem lies in adding the two variables in (8).

5. THE BOEHM DISTRIBUTION

There is a good approximation for the beta distribution that is easier to use in (8). This distribution was derived for use with probability forecasting. If variables are distributed as multivariate normal, than the conditional probability that a threshold t will not be exceeded is given by:

$$Pr(x<y) = \Phi\left[\frac{\ddot{y} - R\ddot{X}}{(1 - R^2)^{1/2}}\right] \qquad (10)$$

where R is the multiple correlation coefficient (not to be confused with matrix $\underline{R}$ as used above) and X is the predictor (or linear sum of predictors weighted by regression coefficients). See Boehm (1976). In either case X is normally distributed and (10) can be inverted, and the frequency of a given probability or less being forecast is:

$$Fr(\ Pr<p) = \Phi[(\ddot{p}S - \ddot{y})/R] \qquad (11)$$

where $S = (1 - R^2)$ and the quotes (") over a variable indicate it is the equivalent normal deviate of the variable, $p = \Phi^{-1}(p)$.
If (11) is differentiated (using the differential of an integral), the result is the probability density, f(p), of a given probability forecast:

$$f(p) = \frac{S}{R}EXP\left\{\frac{\ddot{p}^2}{2} - \frac{1}{2}\left[\frac{\ddot{y} - \ddot{p}S}{R}\right]^2\right\} \qquad (12)$$

This density has a mean, μ_f,

$$\mu_f = y, \tag{13}$$

a median, $\tilde{\mu}$,

$$\tilde{\mu}_f = \tilde{y}/S, \tag{14}$$

and a variance, σ^2, in probability space,

$$\sigma^2 = \Phi_2(\tilde{y}, \tilde{y}, R^2) - y^2 \tag{15}$$

where $\Phi_2(x,y,r)$ is the bivariate normal with correlation r integrated from minus infinity to limits x and y. The variance in equivalent normal deviate space is in equivalent normal space:

$$\sigma^2 = R^2/(1 - R^2) \tag{16}$$

The relavent point in this context is that (12) is a good aproximation for the beta distribution.

6. THE CUB ALGORITHM

In (8) the first variable is normal with a mean of zero and variance of one. The second variable is skewed when coverages near zero or one are required, and it is skewed more as n goes higher. However, David (1981, p272) shows that in the limit as n goes to infinity, that the result of adding the two variables is a normal variable. Thus, we are encouraged to use the form:

$$Pr(C<c) = \Phi[\ (P_0 - \mu_r)/\sigma_r], \tag{17}$$

where C is coverage and c is a coverage threshold, P_0 is the probability of a point in the area being covered. μ_r and σ_r are the mean and standard deviation of the correlated variable Y_i in (8). μ_r and σ_r vary with c. They are calculated by adding the weighted means of the two variables in (8) and also getting their resulting variance:

$$\mu_r = 0 \ + \ (1 - r)\mu_i \tag{18}$$

Recall the first variable has zero mean. μ_i is the mean for the ith variable in the equicorrelated model. For the median i, the mean is 1/2. For the extreme i, the mean equivalent normal deviant is the median of the beta distribution which for the extreme is simply the exponential distribution. The median in this case is found by solving,

$$1/2 = p^n \quad \text{to get} \quad p=(1/2)^{(1/n)} = 2^{(-1/n)} \tag{19}$$

and μ_e for the extreme is the equivalent normal deviant of p. Linear interpolation is satisfactory between a coverage of 1/2 and the extreme so that:

$$\mu_i(c) = \Phi^{-1}\{1/2 + (c - .5)[2^{(1-1/n)} - 1]\} \tag{20}$$

The one in the exponent comes from multiplying by 2.

The standard deviation is not so simple. While the result (8) approches normal for large n, the exponential distribution at extreme coverage, even when the equivalent normal deviate transformation is used, is far from normal. In this case, a two point method is used to calculate the standard deviation. One point is the extreme mean as found by (19). The other point that is used is the probability value, P_e, that would occur at the extreme if there was no correlation, that is,

$$P_e = (P_0)^n \tag{21}$$

Since two points define a line and the slope of a line in equivalent normal space is the standard deviation, the standard deviation at the extreme, σ_e, is calculated:

$$\sigma_e = [\Phi^{-1}(P_0) - \mu_e]/\Phi^{-1}(P_e) \tag{22}$$

At the median Y_i is a very close approximation to the Boehm distribution and (15) and (16) can be used to calculate the median variance, σ_m. The cumulative bivariate normal has a closed solution when both x and y in (15) equal zero:

$$\Phi_2(0,0,\sigma_m^2) = 1/4 + \sin^{-1}(\sigma_m^2)/(2\pi) \tag{23}$$

Substituding the variance of the median, 1/(n+2), of n independent samples for Φ_2 and using (16) to solve for σ_m^2 in equivalent normal space:

$$\sigma_m^2 = \frac{\sin[\pi/(2n+4)]}{1 - \sin[\pi/(2n+4)]} \tag{24}$$

An empirical fit based on many simulations and a fair amount of trial and error is used to interpolate the variance between the median and the extreme:

$$\sigma_i^2(c) = [(2c - 1)^2]^{\ln(n)}[\sigma_e^2 - \sigma_m^2] + \sigma_m^2 \tag{25}$$

Eq (25) gives the variance of independent samples at coverage c. To calculate variance of correlated samples, the variances are added per weights defined by (8):

$$\sigma_r^2 = 1 r \ + (1 - r)\sigma_i^2 \tag{26}$$

The results of (26) and (20) are substituted into (17) to provide the probability of coverage.

7. COMMENTS

Eq (21) worked well on 48 and 64 bit computer floating point words, but it became numerically unstable on 32 bit machines. A patch was develped that allows use of CUB on 32 bit computers but it is somewhat messy. A more elegant fix is being developed. In the meantime, the patch can be bypassed by using double precission with the above formulas.

The resulting algorithm is only about 15 lines of FORTRAN code and provides good approximations to a large variety of coverage problems. It has been used extensivly to fit ground and satellite cloud coverage. For sky

cover, n set at a constant 80 has produced very good statistical fits with mean absolute deviations in the density measuring only a few percent.

The acronym CUB stands for Coverage Using the Boehm distribution.

This research was sponsored by the Philips Laboratory, Geophysics Directorate under contract F196228-88C-0089.

8. REFERENCES

Boehm, A., 1976: Transnormalized Regression Probability, AWS-TR-75-259, Air Weather Service, 52 pp.

Boehm, A., 1988: Bayesian Error Update of Distributions, Preprints Fourth International Meeting on Statistical Climatolgy, New Zealand.

Burger, C., and Gringorten, I., 1984: Two Dimensional Modeling for Lineal and Areal Probabilities of Weather Conditions, AFGL-TR-84-0126, ADA 147970, 58 pp.

Burger, C., 1985: World Atlas of Total Sky Cover, AFGL-TR-85-0198, ADA 170474, 111 pp.

David, H., 1981: Order Statistics, John Wiley & Sons, New York, 360 pp.

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

Gringorten, I., and Boehm, A., 1987: The 3D-BSW Model Applied To Climatology of Small Areas and Lines, AFGL-TR-87-0251, ADA 199114, 91 pp.

Hall, P., 1988: Introduction to the Theory of Coverage Processes, John Wiley & Sons, New York, 408 pp.

Henderson-Sellers, A., 1991: An Investigation of the Burger Distribution to Characterize Cloudiness, J. of Climate, 4, No 12,

Fifth International Meeting on Statistical Climatology
June 22-26, 1992, Toronto, Canada

A STOCHASTIC APPROACH TO SIMULATION OF MULTIVARIATE RAINFALL PROCESSES

Van-Thanh-Van Nguyen
Chavalit Chaleeraktrakoon
Department of Civil Engineering and Applied Mechanics
McGill University
Montreal, Quebec, Canada.

ABSTRACT

Information on physical and statistical characteristics of rainfall processes has proven to be essential for various types of hydrologic studies concerning the planning, design and operation of many water resources systems. In particular, rainfall data are usually required as input to watershed simulation models in order to simulate the physical processes of runoff. Hence, knowledge about the space–time variability of observed rainfall time series is necessary to study the behaviour of the runoff process or to assess the reliability of water resources projects. However, the observed rainfall records, even if they are available, contain only a finite amount of information about their historical patterns. A stochastic simulation of rainfall processes is thus needed to generate many sequences of synthetic rainfall time series which could have similar statistical properties to the observed series. A large number of generated input rainfall sequences would provide more adequate information to assess the response and reliability of a water resources system.

Most previous investigations, however, have considered rainfall simulation at a single site, but few studies have dealt with multisite simulation. The past efforts in the area of multisite rainfall modelling have relied heavily on the second–order tools (e.g., autoregressive moving average models). These second– order models might not be appropriate for cases in which one has to consider a large number of sites, because it is necessary to estimate many parameters inherent in the models. The objective of this study, therefore, is to propose a stochastic method for simulating rainfall sequences at various locations simultaneously based on the theorem of sigular value decomposition (SVD) of matrices. The SVD theorem has been found to be very efficient in dealing with large matrix computations. The proposed method is shown to be able to preserve various desirable statistical properties of the multivariate rainfall process (e.g., temporal and spatial correlation structures,...)

An illustrative application of the proposed simulation method will be presented using daily rainfall data from a network of five recording raingages in the Montreal region (Quebec, Canada). The observations during summer seasons (June–September) over a concurrent five–year period (1969–1974) were used. The suggested simulation method is relied on the multivariate normal distribution. A fractional power transformation is used to transform the skewed observed distribution to normal at each station. A generation of 100 synthetic rainfall samples was performed. It was found that the proposed simulation method could reproduce well the observed means, standard deviations and spatial correlations of the historical rainfall series.

The Structure and Behaviour of Precipitation Fields During Short-Term Intense Rain Events: An exploratory investigation of non-linear techniques.

Claire M. Cosgrove and Michael Garstang
Department of Environmental Sciences
University of Virginia
Charlottesville, Virginia 22903

Abstract

A case study approach is used to characterise the behaviour of short-duration, intense rain systems. The data analysed was collected by a network of tipping bucket rain gauges with a temporal resolution of 1-minute and a volume resolution of 0.25mm. This network is situated over the Kennedy Space Flight Centre in central Florida. Based on predetermined criteria for delineating rain events, nearly 200 such rain events were identified as having occurred during a 12 month period. The climatological statistics revealed that 5% (11 events) of the rain events produced 50% of the total rainfall recorded over the network while 19% (37 events) produced 88% of the total rainfall. On average, the duration of the extreme events was approximately two hours with nearly 40% of the rainfall being measured in the first quartile, i.e. 30 minutes. This translates into nearly 90% of the annual rainfall being measured in 74 hours of which, 40% of the total amount is collected in 18.5 hours, or in 0.2% of the year.

In attempting to analyse precipitation fields on monthly and seasonal basis, we are dealing with an extremely variable parameter in both temporal and spatial scales. It is the intent of this paper to present an alternative approach where the rain field as measured by a relatively dense network, is treated as a single entity made up of the collective rain gauges. Instead of treating the rainfall measurements as a fixed geographic entity, the measurements are treated in a lagrangian framework moving with the rain system. By merging the rain gauge measurements into a continuous time series with the measurements in chronological order, the spatial component of the rain system moving over the network is also taken into account. Non-linear dynamics theory is applied to the analysis of the short time series of precipitation observations.

An objective of this paper is to explore the use non-linear systems techniques for characterising the precipitation fields at a small-scale. These techniques are applied to assess the highly variable nature of rain and to determine from point datasets whether rainfall is purely stochastic or whether there is a deterministic chaos component. A final question to be addressed is whether the structural and behavioural characteristics of a single rain event are common to other rain systems. If this concept of universality is present within different systems, then a potentially valuable insight into climate forecasting and modelling could be gained with use of these non-traditional methods.

THE BI-VARIATE FREQUENCY DISTRIBUTION OF TEMPERATURE AND DEW POINT IN TEXAS AND THE SOUTHEAST

Jon W. Zeitler

Southeast Regional Climate Center
Columbia, South Carolina
(803) 737-0800

ABSTRACT

Human and livestock heat stress, crop condition, energy systems design, and material degradation are all affected by the combination or concurrence of heat and atmospheric water vapor. Heat is typically measured in terms of dry-bulb temperature, while atmospheric water vapor may be measured in a variety of ways, one of which is dew point temperature (dew point). Historically, few studies of dew point have been performed. Additionally, very few investigations of concurrent climate variables have been undertaken. New observing networks will provide additional sites for measurement of concurrent climate observations on at least an hourly basis. However, there will remain a need to estimate probabilities of occurrence for concurrent climate variables at locations for which data is not available.

The objective of this study is to examine a method for fitting daily mean temperature and dew point to a bi-variate normal distribution, such that joint-event probabilities of temperature and dew point may be calculated. Six stations in Texas were selected to represent the various climate conditions found in the state. Twenty years of observations (1961-1980) at three-hour intervals were used to calculate daily mean temperature and dew point values. January, April, July, and October were selected to represent seasonal patterns and reduce autocorrelation between successive periods. Autocorrelation between successive daily observations was examined for significance at lags one through six. Lag four (observations at four-day intervals) provided the optimum reduction in autocorrelation while maintaining an adequate sample size for fitting the bi-variate normal distribution. Using criteria limits of skewness and kurtosis from each variable's distribution, 18 of the original 48 variables (24 station-months x two variables) were normally distributed. Power transformations were applied to non-normal variables to derive new normally distributed variables. Many of the original variables were negatively skewed, indicating use of a power transform greater than unity.

The normally distributed original and transformed variables were fit to a bi-variate normal distribution. Joint-events were defined for each station-month, and expected values were calculated for each joint-event from the theoretical distribution. Observed values were calculated for each joint-event from the original data. A chi-square goodness-of-fit test was applied to each station-month to determine the degree of fit at the 0.05 significance level. Results indicated that daily mean temperature and dew point values can be successfully fit for stations distant from the Gulf of Mexico in January and April. Little success in fit was found for stations near the coast in nearly every month, and for all stations in July.

The overall conclusions of the study are that daily mean dew point and to a lesser extent daily mean temperature values are not normally distributed at coastal locations in most months and at all locations during July for the six locations in Texas. Similarly, attempts to fit the original or transformed normally distributed variables to a bi-variate normal distribution were unsuccessful in July at all stations and at coastal locations in most months. Based on these conclusions, a statistical examination of temperature and dew point values for some select stations in the Southeast (Alabama, Florida, Georgia, North Carolina, South Carolina, Virginia) will be performed such that a different approach might be developed for calculating joint-event probabilities. Results for the Southeast can be compared to those found in Texas.

Entropy Value Is Also A Statistical Parameter To Express Climate Status

Ma Shuhong
(Chinese Academy of Meteorological Sciences)

Znang Xuewen
(Meteorological Institue of Xinjiang, China)

1. Introduction

For a long time, we had threated the mean value of meteorological ele-ment as a climate parameter. thus to analyse the geographic distribution of such mean value is a very important work. But it is also clear that if we only research the mean value in climatology, it is not sufficient.

Here we will point out that the entropy value is also a important climate parameter. We may analyse the distribution of entropy value of certain meteorological element on the map as we had done it for the mean value.

In statistics, we try to regard any meteorological element value at given point as a random variable. Then for a random variable x we can get its probability distribution p(x) for discrete variable or its probability density distribution function f(x) for continuous variable.

Therefore, the mean value of random variable is defined as

$$\bar{x}=\begin{cases}\sum x_i\, p(x_i) & \text{discrete}\\ \int_{-\infty}^{\infty} x\, f(x)dx & \text{continuous}\end{cases}$$

or

$$\bar{x}=E(x)$$

In information theory, we study the entropy value H(x) of random variable x, here (Reza, 1962)

$$H(x)=\begin{cases}\sum -p(x_i)\ln p(x_i)\\ \int_{-\infty}^{\infty} -f(x)\ln f(x)dx\end{cases}$$

Since probability value p(x) is the function of random variable x,according to the former equation, we can take entropy value H(x) as a special mean --the mean of -lnp(x), or -lnf(x), so we get

$$H(x)=\begin{cases}E(-\ln p(x))\\ E(-\ln f(x))\end{cases}$$

So, we have known, like mean value x, the entropy value H(x) is also a statistical parameter of meteorological element x.

Furthermore, if we can work out many entropy value of every station for meteorological element in certain area, we can also plot them on map and analyse its isoline of entropy value for that meteorological element. So, like the mean value figure, we can get a figure about the distribution of entropy value. Based on this figure, we could research the geographical distribution of entropy value.

We think this is a new discipline in climate research.

2. Example

Using above-mentioned formula and associated climate data, we can easi-ly plot the entropy value distributions for various climate parameters. We just give out several cases here.

2.1 Distribution of entropy value for daily temperature mean

The daily mean of temperature is a random variable, which has certainly

corresponding entropy. We have calculated the entropy values of temperature means for each division and 4 seasons in Xinjiang region. Figure 1 is one case among them.

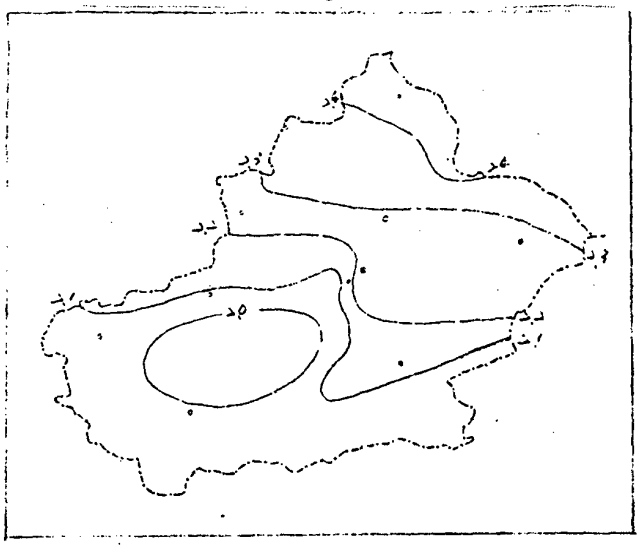


FIG. 1 Distribution of entropy value for daily temperature mean of spring in Xinjiang region (nat)

2.2 Distribution of entropy value for daily precipitation

If there occurs rainfall in a certain day, the value of rainfall is a random variable, which should have corresponding entropy value. We calcu-lated the entropy value of daily pre-cipitation for China area. Eigure 2 is the entropy value distribution for one season.

FIG.2a Distribution of entropy for daily precipitation of continental spring in china(nat)

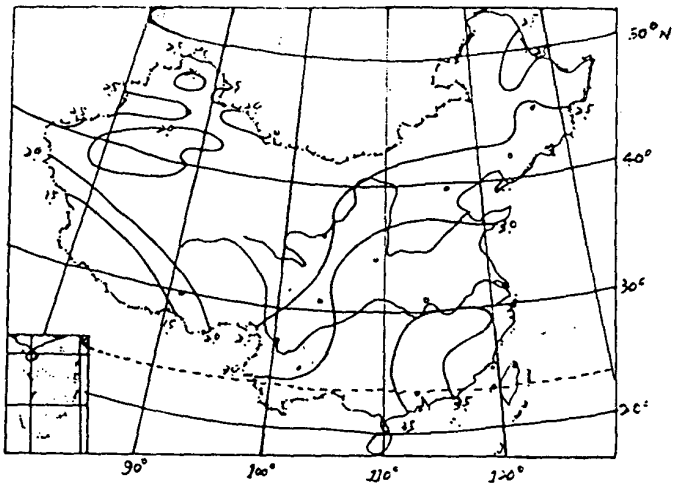
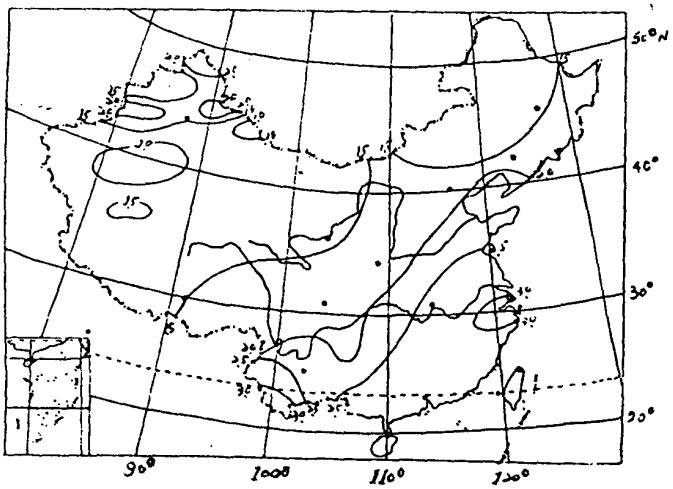


FIG.2b Distribution of entropy for daily precipitation of continental summer in China(nat)



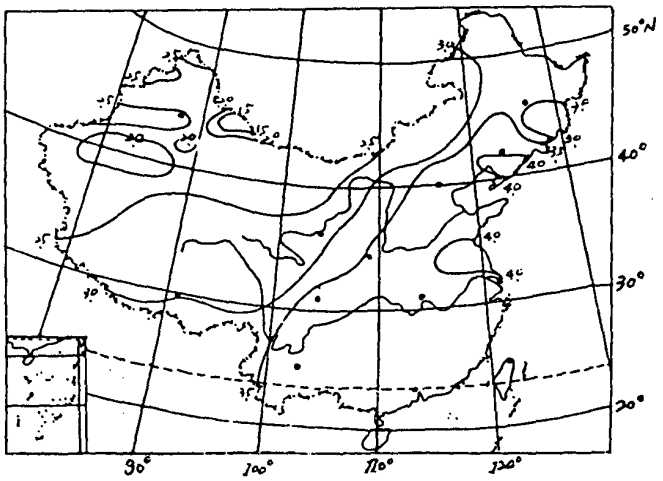


FIG.2c Distribution of entropy for daily precipitation of continental autumn in China(nat)

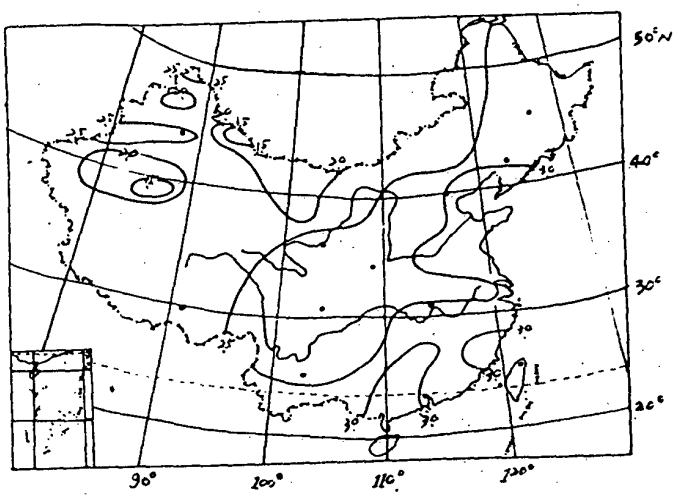


FIG.2d Distribution of entropy for daily precipitation of continental winter in China (nat)

2.3 Distribution of entropy for non-rain period length

The length of continuous non-rain period in a certain place is regarded as a random variable, thus its entropy can also be worked out. Figure 3 is the distribution of entropy value for non-rain period length in spring in Xinjiang region(Figure omitted).

3. Discussion

3.1 The work for analysing climate entropy started long ago. As early in 1960-70's, some Japanese researcher had done some works on it. Zhang Xuewen had also analysed it systematically in early 1980's.(Zhang, 1981)

3.2 In information theory, the measurement for entropy is laid on the uncertainty for various results of events occurred. In meteorology,when climate entropy for a certain place is larger, this also means that local weather changes grently, therefore climate entropy illustrates the magnitude of the change degree, it is also an index of scientific measurement for local meteorological prediction difficulty. If the entropy value of meteorological parameter is larger in a certain place than in another place, the weather prediction difficulty is also larger.

3.3 Entropy value $H(x)$ and the mean $\bar{x}$ are two physical elements which have different physical meaning. They describe the characteristic of random variable x from different aspects. In climatology, many climate maps have been published to illustrate the mean distribution, there is few climate map of entropy distribution, hence we should publish some maps of entropy distribution.

3.4 In the research on climate change at the present age, people only pay attention to the problem of the mean change, few people pay attention to the problem of entropy change for meteorological parameter, we should do some works in this field.

References

Reza,F.M., An introduction to information theory,Mcgraw-Hill Book Company, INC.NEW YORK,1961 (268)

Zhang Xuewen,The information analysis of weather forcasting problem,SCIENCE PRESS,BEIJING,1981,28-56

A STATISTICAL APPROACH TO THE EVALUATION OF A WEATHER MODIFICATION PROJECT

Taher A. Sharif
Department of meteorology, University of Alfateh
Tripoli, Libya
and
M. T. El-Alem
Cloud Seeding Project
Tripoli, Libya

1. INTRODUCTION

Little doubt exists that man can modify clouds. However, questions remain as to the extent to which clouds can be modified. Can the precipitation be changed in a predictable manner over a substantial area, in quantities valuable to agriculture and industry, without causing harm to surrounding areas; and can this be done in an economically beneficial way? World Meteorological organization (WMO) designed, executed and evaluated a precipitation enhancement project to provide scientifically convinced answers to these questions and others WMO (1982).
The objective of this study is to preliminary evaluate the cloud seeding operations which have been coducted in Libya since fall 1980 by using a well known statistical technique. The results of this study is intended primarily as a technical base for recommendations to the authorities regarding the continuation of the operations of the precipitation enhancement project. In view of the fact that the data used is preliminary and containing some missing values, the results and conclusions must be considered tentative. The future planned data collection may provide more adequate information that can be used for further statistical evaluation of these operations.

Weather modification experiments typically have relied on a mixture of statistical and physical evaluation approaches to discern effects of cloud seeding. Evaluation were sources of controversy of four major weather modification projects Hsu and Huff (1986). However, our study is statistical with no cloud physics or seeding logistics considered at this time; such physical evaluation was previously undertaken by us, Sharif et al. 1990. The effects produced by a typical cloud seeding program are usually smaller than the natural background variations. In data processing terms, the search for seeding effects is a search for a weak signal in the presence of random noise, and the effect can only be estimated. It was perhaps inevitable that evaluators should turn to statistics, which is defined as "systematic compilation of instances for the influence of general truths" Dennis (1980).

2. DATA

The libyan cloud seeding project is operating from three different sites, namely Jeffera Plain near Tripol, Sirte, and El-Marej. The data of Jeffera Plain site will only be analyzed in present study. It is proposed that the highest potential for cloud seeding is associated with cumulus cloud development which generally occurs in the cold sector of the cyclonic systems passing across North Africa during October through March which is referred to

in this paper as seeding season. In such synoptic situation the air current is usually westerly.

Unfortunately the cloud seeding operations in the site under present investigation was not experimentally designed for future statistical evaluation except for the first and second seasons , after which seeding took the non-experimental behavior i.e seeding when and wherever required clouds exist with no precious scientific guidance. For this reason we faced some difficulties in classifying the data, as we were trying to approximately bring it to a suitable format to suite our statistical model.

The aircraft seeding tracks were identified for all seeding cases for the period from 1980 to 1990, so that the area within the boundary which includes all those tracks is designated as target area. An area upwind of the target area is chosen to be the control area. The target and control areas include synoptic and climatic stations. A correlation analysis performed first on the stations in each area separately. Among the stations in the target area only four stations which have correlation greater than 0.65 were selected to represent this area. By the same way three stations were chosen to represent the control area.

It was decided to separate the duration of monthly average rainfall for both areas into two parts. The period prior to 1980 is considered as independent data to be used for the calculation of the long term monthly averages of 24 years, these averages are shown in Table 1. The period from 1980 to 1990 for which the cloud seeding operations were conducted will be used for the evaluation purposes.

Table 1. Long-term average of monthly total rainfall of target and control areas.

Area	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Season
Target	42.7	36.0	53.2	54.6	30.3	30.4	247.2
Control	37.4	24.8	35.6	33.0	19.8	23.2	173.4

It should be noted that no attempt was made to compensate for the missing observations, whenever a missing observation is encountered in a certain time, the same observation is excluded from the averages of all the stations in the specific area.

3. STATISTICAL MODEL

The most widely used approach to the evaluation of operational cloud seeding for rain enhancement projects is the comparison of average rainfall in the target area to that in one or more control areas for the same period Dennis (1980). This can be done by calculating the rainfall in each area in terms of percentage of normal, where "normal" is the long-term average rainfall during some predetermined historical period.

Suppose that, during a given period of operations, the target area received RT percentage of its normal rainfall for that period, while the control area received only RC percentage of normal. A ratio that estimate the increase of target area rainfall due to seeding over that which would have fallen without seeding is called target-control

ratio SE and can be calculated as:

$$SE=(RT/NT-RC/NC)/(RC/NC)$$

$$=(NC*RT/NT*RC)-1$$

where RT is the percentage of the target area rainfall from its normal, RC is the percentage of the control area rainfall from its normal, NT is the long-term average (normal) rainfall of target area and, NC is the long-term average (normal) rainfall of control area.

4. DISCUSSION

The issue of evaluating seeding effects for producing more rain using both physical measurements and statistical tests to establish proof of weather changes was another area of little knowledge Changnon and Lambright (1990). Many scientific controversies during the different stages of weather modification by cloud seeding revealed a lack of scientific knowledge in this area. It should be clear that using statistics and specially ratios in evaluating weather modification experiments where target-control ratios or seed-no-seed ratios drown from a single target area, carries a special risk of error Dennis (1980). Assuming that the target and control areas have similar rainfall patterns, the expected value of the target-control ratio is generally greater than 1. Its exact value depends on the distributions of rainfall events in the target and control areas.

In this investigation, the target-control ratio is calculated using the data prescribed in previous sections and is shown in Table 2. In this table the ratios are computed separately for each month of the season (Oct., Nov., Dec., Jan., Feb. and Mar.) and also for the seasonal total. The target area rainfall for then seasons (80/81 through 89/90) were examined. Although there are significant correlation between the rainfall of the target and control areas, there rainfall distributions are not very similar, for this reason the table indicates positive and negative ratios. The positive ratio signifies that there was increase in the rainfall due to seeding while the negative indicates that the seeding was ineffective. It is remarkably clear that the ratios of 80/81 and 81/82 seasons are positive for the individual months and for the seasonal totals, but for the ratios of the remaining seasons 82/83 through 89/90, the signs are

Table 2. The target-control area ratios.

Season	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Seasonal
80/81	0.9	8.4	0.2	5.6	2.5	54.4	1.7
81/82	*	*	4.2	0.9	11.0	1.9	2.2
82/83	-0.4	0.7	-0.0	1.0	-0.2	-0.1	-0.2
83/84	0.5	0.4	-0.1	-0.5	1.8	*	0.1
84/85	-0.1	-0.5	0.0	0.4	-0.5	*	-0.1
85/86	1.2	8.5	-0.4	0.1	*	*	-0.2
86/87	0.5	-0.2	1.2	2.7	0.6	*	0.2
87/88	*	-0.4	-0.7	0.6	*	1.5	-0.2
88/89	*	-0.6	0.5	0.9	-0.5	0.1	0.1
89/90	0.3	0.0	1.6	-0.5	*	*	-0.2

* is missing information.

ambiguous. A question that may be raised is that why for the first two seasons, the cloud seeding operations were significantly successful while there was no persistency for the remaining seasons? Part of the answer is that for the first two seasons, all the cloud seeding equipments were in good conditions, in addition to that, the operating companies for these two seasons are different from those for the remaining seasons i.e. different experience and techniques etc..

It will be more efficient if individual seeding cases data is used directly for comparison of target and control areas rainfall instead of monthly or seasonal averages. For the project under investigation, it seems that statistical evaluation can not be relied on if standard conditions are not met, such conditions like scientific guidance, management, standardization of equipments etc. are the primary requirement for any satisfactory evaluation. The results of this statistical evaluation should be cautiously used until an improved data is collected.

5. CONCLUSION

The evaluation of weather modification projects is a difficult task. One of the statistical techniques that can be used is the target -control ratio. In our present study, the target-control ratio has indicated some seeding effects in the first two seasons but ambiguous results for the remaining seasons, such ambiguousness needs to be investigated.

6. REFERENCES

Changnon,S. A. and W. H. Lambright, 1990: Experimentation involving controversial scientific and technological issues: Weather Modification as a case illustration. Bull. Amer. Meteor. Soc, 71, 334-344.

Dennis, A. S., 1980: Weather modification by cloud seeding. Academic Press, N.Y., 267pp.

Hsu, C. F. and F. A. Huff, 1986: Assessment of problems, successes and failure in past weather modification projects. J. Wea. Mod., 16, 14.

Sharif, T. A., P. S. Sehra and A. M. Ben-Ali, 1990: Some experimental studies on precipitation enhancement by using silver iodide nuclei. Proceedings 7th conference on cloud physics, Amer. Meteor. Soc., San Francisco Cal. USA.

WMO Report, 1982: Preliminary assessment of the site selection phase-3 of the precipitation enhancement project.

5th INTERNATIONAL MEETING
ON
STATISTICAL CLIMATOLOGY

12th CONFERENCE ON PROBABILITY AND STATISTICS
IN
ATMOSPHERIC SCIENCES

Joint Papers

THE INSTRUMENTAL RECORD OF SURFACE TEMPERATURE: HOW GOOD IS IT AND WHAT CAN IT TELL US ABOUT CLIMATE CHANGE AND VARIABILITY?

Chris K. Folland and David E. Parker

Hadley Centre, Meteorological Office, Bracknell, Berkshire, England

1. INTRODUCTION

Considerable progress has been made by many groups with the creation and analysis of historical instrumental temperature data. However, a full analysis of their systematic and random errors remains to be carried out. Recently Trenberth (1992) and Folland et al. (1992) have provided estimates of uncertainties in recent SST analyses on monthly and seasonal time scales and Madden et al. (1992) have assessed the sampling errors of global land and ocean data in January from 1869-1979. Discussions of random errors in historic temperature data for the Intergovernmental Panel on Climate Change (Folland et al. 1990) have been limited to the use of "frozen grids". These give useful but minimum estimates of the effects of incomplete data coverage on climate trends. Figs 1a and 1b show frozen grid analyses of combined land and ocean data for the Northern and Southern Hemispheres averaged over all seasons for 1861-1991. These diagrams use a recent Met Office historical sea surface temperature (SST) analysis (MOHSST5) and the Jones et al. (1991) land data. Trends are similar even when data are confined to locations adequately sampled in 1861-70. The root mean squared difference (RMSD) of annual values based on all data compared those

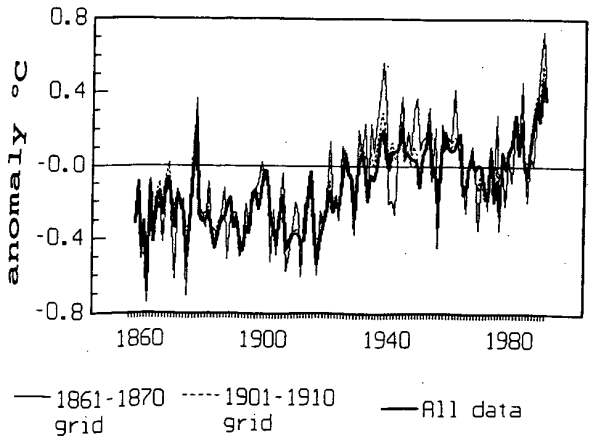


Fig. 1a "Frozen grid" tests of unsmoothed annual Northern Hemisphere land air and sea surface temperature data

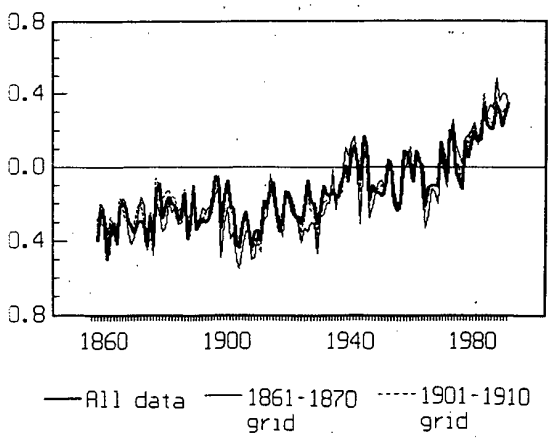


Fig. 1b As Fig. 1a but Southern Hemisphere

based on the 1861-70 data coverage is only 0.15°C (Northern Hemisphere) and 0.09°C (Southern Hemisphere). Seasonal data behave in a similar way; thus seasonal RMSD are only a little greater except in Northern Hemisphere winter (January to March RMSD is 0.35°C).

It is important for improvements to the historic temperature data base to be pursued vigorously. Thus for the oceans, Woodruff (1990) has described the considerable potential that exists for improving the data base of most marine parameters for the last century or more.

2. BIASES IN HISTORIC DATA

Several authors have pointed out the likely large size of systematic inhomogeneities in SST before the Second World War. Bottomley et al. (1990) attempt to correct in detail for these problems as a function of season and geographical position using the "bucket" method. Using a recently revised form of the "bucket" corrections, SST on a global average appears to be at least 0.1°C too cold in the 1860s and about 0.4°C too cold in the 1930s relative to the period 1951-80. Biases varied in the 1930s from near zero (much of the Southern Ocean and parts of

the mid-latitude Northern Hemisphere oceans in early summer) to at least 0.8°C (small areas east of USA and Japan near 35°-40°N in December). Figs 2a and 2b show uncorrected and corrected SST data, and smoothed night marine air temperature (NMAT) data (Bottomley et al. 1990) averaged for the North Pacific in October-December and the North Indian Ocean for the whole year. The NMAT data give good support to the SST corrections which reach 0.5°C in the North Indian Ocean in the 1930s. Large tropical biases at this time are mostly explained by evaporation from uninsulated canvas buckets, while sensible heat losses contribute much more to the large December biases mentioned above.

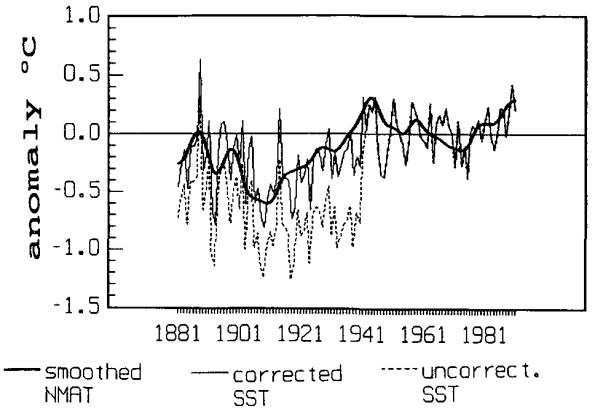


Fig. 2a Uncorrected and corrected Oct-Dec sea surface and night marine air temperature, North Pacific, smoothed with a 21 point binomial filter, for 1881-1991

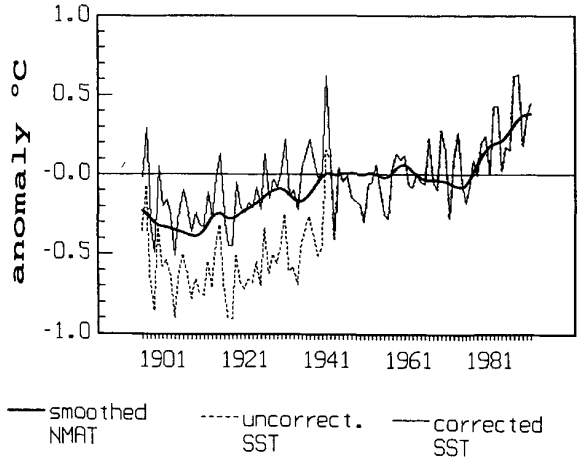


Fig. 2b As Fig. 2a but North Indian Ocean for 1901-1991 whole year

Fig. 3 shows the results of a recent experiment on board ship in the tropical North Atlantic where the observed cooling of sea water in an uninsulated canvas bucket is compared to theory (Folland, 1991). To make the results clearer, differences in observed and calculated cooling are shown for a 10 minute exposure time of the bucket on deck, considerably longer than that thought to have been typical. The average observed cooling rate is insignificantly different from that calculated from theory, with a highly significant correlation of 0.67 between observed and modelled cooling after 10 minutes. Tests in other ocean regions are in progress. Parker (1990) suggests that land data in parts of the tropics may be biased substantially too warm during the

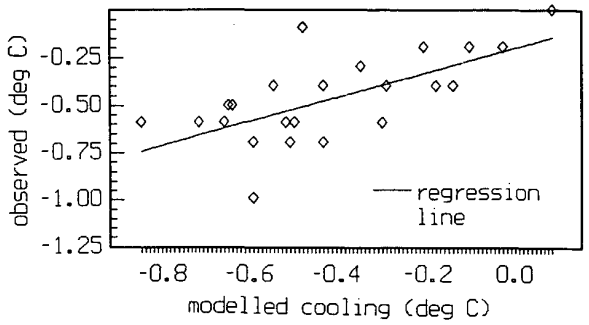


Fig. 3 Observed and calculated cooling of sea water in a canvas bucket on board ship Data courtesy of Professor R Newell, MIT, and the American Sea Education Association

first few decades of this century due to the use of "thatched shed" screens. Fig. 4a shows "colocated" differences between SST, NMAT and air temperature for the region 15°N-equator, 70°-85°E (coastal regions of India and Sri Lanka and Indian Ocean islands). Only data in the same or adjacent 5°x5° areas in a given season are used. Fig. 4a indicates a probable substantial bias in land data of about 0.5°C in the early decades of the twentieth century. Fig. 4b indicates that the warm bias significantly affects "global" land data but with a substantially reduced magnitude. There is clearly a need to investigate this problem further.

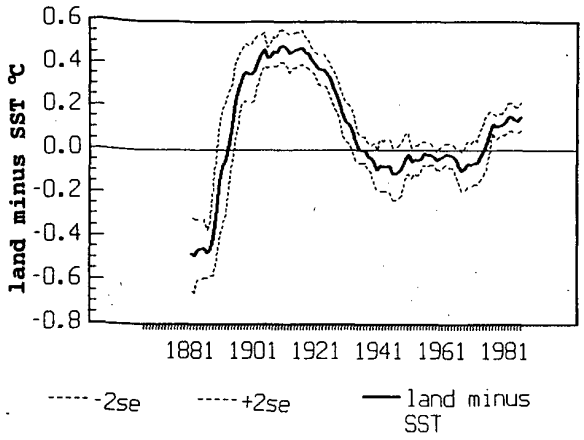


Fig. 4a 10 year running mean differences between colocated land air and sea surface temperature, 1881-1986 for 15°-0N, 70°-85°E. Dashed lines are twice the standard error. (corrected SST used)

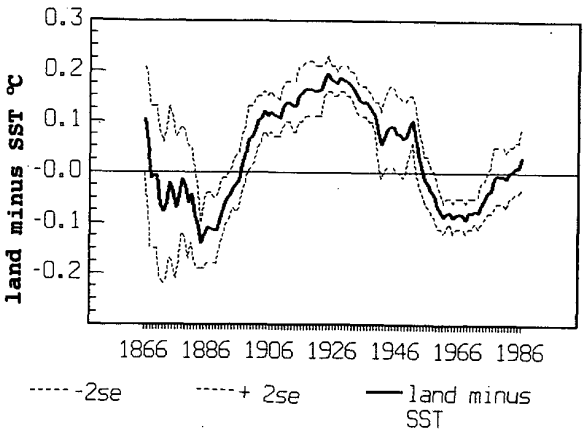


Fig. 4b As Fig 4a but for the Globe

3. SAMPLING ERRORS OF ANALYSES - SOME CONSIDERATIONS

Analyses of sampling errors that result from the inadequate temporal and spatial coverage of historical temperature data, and the intrinsic influence of random errors in individual observations (Trenberth, 1992), need to be carried out separately for each analysis. Thus the recent Met Office historical SST data set (MOHSST5) comes in several forms though all forms use 5° degree boxes analysed on the monthly time scale. Overall, more data is accepted locally as the unit time scale of analysis extends from a month to a decade as unacceptably noisy local data in an individual month may be acceptably sampled over a decade. Thus errors of decadal analyses are not straightforwardly related to those of constituent monthly analyses. Fig. 5, for decadal averages, shows how "data coverage" varies with the choice of three selection criteria for a given 5°x5° area (a) minimum of three seasons data per half decade (b) minimum of four seasons drawn from at least 2 separate years per half decade (c) minimum

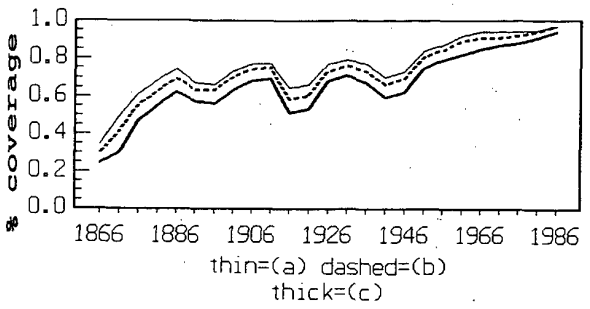


Fig. 5 Variation of apparent SST data coverage for overlapping decades expressed as a percentage of global ocean surface area for three data selection criteria.

of 6 seasons drawn from at least three years per half decade. These choices give (not shown) small variations in decadally averaged hemispheric temperature. Use of 2° monthly boxes, as in Madden (1992), would further alter perception of "data coverage" (reduce it). So the appropriate "radius of influence" of observations is also an issue needing more investigation. It is proposed that random errors could be estimated using a climate model forced with the globally complete blended satellite and in situ SST and ice analysis available since 1982 (Reynolds and Marsico, 1992) (RM) on sub-monthly time scales. Model data would be sampled in the same way in time and space as the analysis under test over land; over the oceans the RM SST data would be used directly. Where temporal or spatial interpolation is used in the analysis under test, modelled air temperature or blended SST data would be treated in as similar a way as possible. In the Met Office marine data sets the temporal distribution of original observations within the month can be extracted to help with this. NMAT could also be studied this way but using modelled near-surface air temperatures over the oceans. These would be constrained by the RM data to be spatially realistic. Climate models have already been used to assess sampling in land data (Hansen and Lebedeff, 1987). The globally complete Microwave Sounding Unit

tropospheric temperature data are also useful for sampling error studies (Madden et al. 1992), though they do not have identical characteristics to surface data (Trenberth et al.(1992), Folland et al. 1992).

4. WHAT CAN EXISTING TEMPERATURE DATA TELL US?

Much of the potential of historical surface temperature data remains to be tapped. Several contrasting applications are given to indicate their value.

4.1 Global Climate Change

Fig. 6a compares approximately decadally smoothed globally average air temperature (an average of the Hansen and Lebedeff (1987), Jones et al. (1991), and Vinnikov et al. (1990) analyses), SST (an average of Jones et al. (1991) data and MOHSST5) and the Met Office NMAT analysis (Bottomley et al. 1990). Multi-decadal and century time scale variations are broadly similar; the good agreement between the three nearly independent types of data set (less true before 1893) contributes much to a general acceptance that a global warming of near 0.5°C has happened over the last 100 years. However the Southern Oceans are mostly very poorly sampled, Antarctica is almost absent before 1957 and Africa is almost absent before 1900.

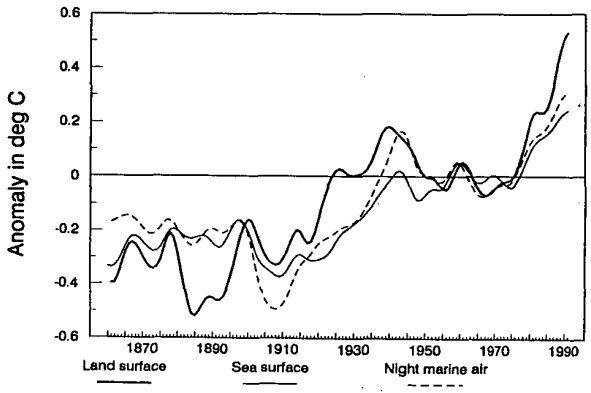


Fig. 6a Smoothed annual global air, sea surface and night marine air temperature, 1861-1991, (smoothed with a 21 point binomial filter) using major compilations.

Warmer land data relative to ocean data in the early twentieth century may result from the use of uncorrected "thatched shed" values. There is also a difference (not shown) between the two SST data sets used in Fig. 5a in much of the late nineteenth century of near 0.15°C largely due to disagreements about corrections for the use of wooden sea temperature buckets.

4.2 Regional climate change and variation, New Zealand

Fig. 6b compares a newly homogenised data set of annual New Zealand air temperatures (NZT) (Salinger, personal communication) with MOHSST5 and NMAT in the surrounding region 35°-50°S, 165°-180°E.

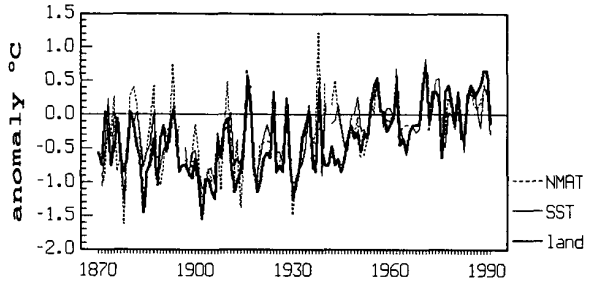


Fig. 6b Annual values of New Zealand air temperature, and sea surface and night marine air temperature in surrounding seas.

Correlations between annual values since 1870 (six marine values missing) are: SST-NZT, r=0.80; SST-NMAT, r=0.86 and NMAT-NZT, r=0.72. The correlations are reduced somewhat by very poor data in the Second World War (see centre of Fig. 6b) but otherwise annual, decadal and century time scale variations are similar. Recent data appear good enough, for instance, to study the individual effects of regional SST and atmospheric circulation variations on seasonal to decadal air temperature anomalies in different New Zealand regions.

4.3 Climate variation and long-range forecasting, UK

Fig. 6c compares five-year average temperatures measured in the Scilly Islands off south west England (50°N, 6°W) in a range of "westerly" atmospheric circulation types over the UK with SST locally, SST upwind in a region covering the North East North Atlantic and SST

averaged over the Atlantic north of 35°N. Scilly Islands data are selected daily so that only westerly "Lamb" atmospheric

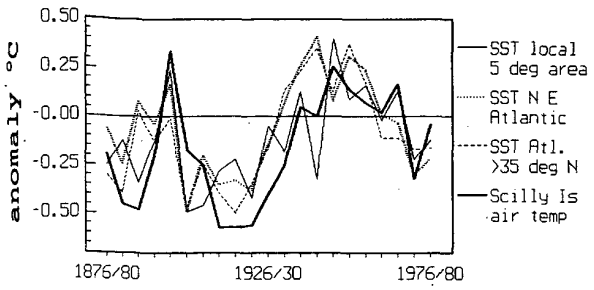


Fig. 6c Comparison of 5 year average air temperature anomalies relative to a 1951-80 average for the Scilly Islands with three North Atlantic sea surface temperature series.

circulation types between southwest and northwest are included. Scilly temperature is quite highly correlated on this time scale with SST upwind, correlations being (local SST) r=0.73 (north east Atlantic) r=0.66 and (North Atlantic north of 35°N) r=0.64. The three SST data sets are well correlated, hinting at climatic processes that have affected much of the North Atlantic over the last century.

Studies for the period 1951-80 for 10 different regions covering the UK show weaker but still useful correlations (typically r=0.3 to 0.6) between upwind SST and air temperature in individual months or half months, especially in summer and autumn (Legg, personal communication). Thus SST is an important component of "specification equations" used to convert long-range predictions of atmospheric circulation to regional temperature anomalies over the UK.

4.4 Use of SST in seasonal forecasting in the tropics

SST data provide the basis of real-time seasonal rainfall forecasting in the Sahel, e.g. Rowell et al.(1992), and for one of the schemes used to forecast rainfall for the wet season in the North Nordeste of Brazil (north east Brazil) (Ward and Folland, 1991). For these forecasts, large-scale SST anomaly patterns are used as statistical predictors, for the Sahel forecasts, also to force a climate model. For the statistical methods, SST provides training data extending over several past decades or more. The good skill of nine of the 11 real-time tropical forecasts issued between 1986 and 1991 (Fig. 7, next page) indicates strongly that historical SST data may be adequate to help develop seasonal forecasting in other tropical regions. However, the likely availability of more accurate, globally complete, satellite SST data in the future should contribute significantly to the further development of seasonal forecasting in the tropics.

REFERENCES

Bottomley, M., C.K. Folland, J. Hsiung, R.E. Newell and D.E. Parker, 1990: *Global Ocean Surface Temperature Atlas (GOSTA)*. Met. Office/MIT, Boston. Funded by US Dept of Energy, US National Science Foundation, US Office of Naval Research, UK Depts of Energy and Environment. pp20+iv and 313 Plates. HMSO, London.

Folland, C.K., 1991: Sea temperature bucket models used to correct historical SST data in the Meteorological Office. *CRTN 14*, 29pp.[Available from the National Meteorological Library, Bracknell, England].

----, T.R. Karl and K.Ya Vinnikov, 1990: Observed climate variations and change. *Climate Change, the IPCC Scientific Assessment*. J.T. Houghton, G.J. Jenkins and J.J. Ephraums (Eds), WMO/UNEP/IPCC, Cambridge University Press, 195-238.

----, ----, N. Nicholls, B.S. Nyenzi, D.E. Parker and K.Ya Vinnikov, 1992: Observed climate variability and change. *Climate Change 1992, the Supplementary Report to the IPCC Scientific Assessment*. J.T. Houghton, B.A. Callander and S.K. Varney (Eds), WMO/UNEP/IPCC, Cambridge University Press (in press).

----, R.W. Reynolds, M. Gordon and D.E. Parker, 1992: A study of six operational sea surface temperature analyses. *J. Clim.* (accepted)

Hansen, J., and S. Lebedeff, 1987: Global trends of measured surface air temperature. *J. Geophys. Res.*, **92**, 13345-13372.

Jones, P.D., T.M.L. Wigley and G. Farmer, 1991: Marine and land temperature data sets: a comparison and a look at recent trends. *Greenhouse-Gas-Induced*

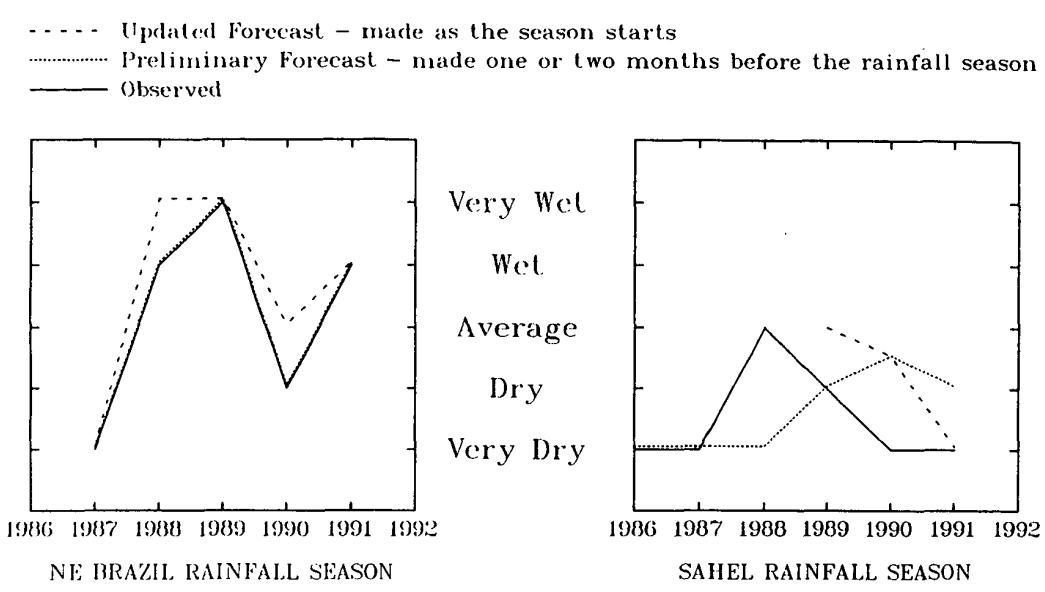


Fig. 7 Performance of real-time seasonal rainfall forecasts based on SST data for the Sahel and North Nordeste of Brazil

Climatic Change: a Critical Appraisal of Simulations and Observations. M.E. Schlesinger (Ed), Elsevier, 153-172.

Madden, R.A., D.J. Shea, G.W. Branstator, J.J. Tribbia and R. Weber, 1992: The effects of imperfect spatial and temporal sampling on estimates of the global mean temperature: Experiments with model and satellite data. NCAR/0401/90-12 [Available from the National Center for Atmospheric Research, Boulder, CO, USA].

Parker, D.E., 1990: Effects of changing exposure of thermometers at land stations. *Observed Climate Variations and Change: Contributions in Support of Section 7 of the 1990 IPCC Scientific Assessment.* D.E. Parker (Ed.), ppXVIII.I-XVIII.31. IPCC/WMO /UNEP.

Trenberth, K.E., J.R. Christy and J.W. Hurrell, 1992: Monitoring global monthly mean surface temperatures. *J. Clim.*, (in press).

Reynolds, R.W., and D.D. Marsico, 1992: An improved real-time sea surface temperature analysis. *J. Clim.*, (submitted)

Rowell, D.P., C.K. Folland, K. Maskell, J.A. Owen and M.N. Ward, 1992: Causes and predictability of Sahel rainfall variability. *Geophys. Res. Lett.*, (in press).

Vinnikov, K.Ya, P.Ya Groisman and K.M. Lugina, 1990: Empirical data on contemporary global climate changes (temperature and precipitation). *J. Clim.*, 3, 662-677.

Ward, M.N., and C.K. Folland, 1991: Prediction of seasonal rainfall in the North Nordeste of Brazil using eigenvectors of sea-surface temperature. *Int. J. Climatol.*, 11, 711-743.

Woodruff, S.D., 1990: Preliminary comparison of COADS (US) and MDB (UK) ship reports. *Observed Climate Variations and Change: Contributions in Support of Section 7 of the 1990 IPCC Scientific Assessment.* D.E. Parker (Ed.), ppXXVII.1-XXVII.36 IPCC/WMO/UNEP.

SIGNAL AND NOISE IN THE SURFACE TEMPERATURE RECORD

Kevin E. Trenberth

National Center for Atmospheric Research[1]
Boulder, Colorado

1. INTRODUCTION

How well do we know individual monthly mean surface temperatures over the entire globe? How well can we determine the surface conditions for monitoring the earth's climate and, in particular, climate change, such as that expected from increases in greenhouse gases? Monitoring changes is actually easier than determining absolute values because systematic errors, such as those arising from biases in methods of measurement or spatial coverage, are largely removed. These are the questions addressed in this paper.

There have been several prodigious efforts made to reconstruct the surface temperature record for the past hundred years or more from instrumental observations. Fairly complete observations of surface temperature over land now exist, although their continuity back in time is an issue. Substantial corrections need to be applied to the marine surface sea and air temperatures (Jones et al. 1991, Bottomley et al. 1990, IPCC 1990). The compilations of surface air temperature variations agree that, in global terms, a warming of ~0.5°C has occurred since 1880. Recent decades appear to have been somewhat warmer than any other time in the past 120 years or so, while the period from about 1890 to the 1920s was distinctly colder than average.

A primary objective of this paper is to assess how well the global surface air temperatures are known for a particular month over the past decade or so, essentially using all of the in situ observations available today. The sources of noise in the data, the numbers of observations, and the spatial coverage are assessed along with the size of the climate signal. Different analyzed results are compared to see how reproducible they are. A more complete report is given in Trenberth et al. (1992).

2. THE DATA SETS

We have used data from the Comprehensive Ocean-Atmosphere Data Set (COADS), SST analyses from the Climate Analysis Center (labeled CAC) (Reynolds 1988), and the recently updated IPCC (1990) surface data set (labeled SFC) which is a combination of Jones land data (Jones et al. 1986a, 1986b, Jones 1988) and the United Kingdom Meteorological Office (UKMO) SST data (Bottomley et al. 1990).

The surface data set is based upon station temperatures over land surfaces combined with in situ SSTs in the form of anomalies, relative to 1951-80. Satellite estimates of sea surface temperatures (SSTs) have not been included in compiling the historical record, although recent CAC global SST analyses have used information from satellites on SST patterns and gradients.

Large differences occur between the surface air temperature and SST in some places and at certain times of the year, notably off the east coast of Asia and North America and off Antarctica in winter where mean differences exceed 4°C, but it is generally expected that *anomalies* of air and sea surface temperatures will go hand-in-hand. Because sampling is a major problem over most of the ocean and SSTs have much greater persistence, it has generally been preferred to use SSTs for monitoring surface temperatures. The differences between SST and air temperature anomalies are quantified in section 4.

3. SIGNAL-TO-NOISE RATIO

We consider the common signal in two time series. The approach is related to a method called "common elements" (Snedecor and Cochran 1973, p181). It can be contrasted with regression analysis in which the two time series are not treated equally (one is the predictor and one is the predictand) and the variance of one series explained by the other, if both are normalized, is r^2, where r is the correlation coefficient.

Given two time series $x(t)$ and $y(t)$ with zero mean that ostensibly measure the same quantity $z(t)$ in the presence of noise. The signal z will be measured by its variance S. We put

$$x(t) = z(t) + \epsilon(t)$$

$$y(t) = z(t) + e(t)$$

where $\epsilon(t)$ and $e(t)$ are the noise, assumed random, and give rise to noise variances N_x and N_y. Then the correlation coefficient between x and y is

$$r = \frac{S}{\sigma_x \sigma_y} \tag{1}$$

where σ is the standard deviation. Because $\sigma_x^2 = S + N_x$ and $\sigma_y^2 = S + N_y$, a measure of the average noise is

$$N = 0.5(N_x + N_y) = 0.5(\sigma_x^2 + \sigma_y^2 - 2r\sigma_x\sigma_y)$$

using (1), and the signal-to-noise ratio is

$$S/N = \frac{2r}{\frac{\sigma_x}{\sigma_y} + \frac{\sigma_y}{\sigma_x} - 2r} \tag{2}$$

or, in the case where $\sigma_x \approx \sigma_y$, then

$$S/N = \frac{r}{1-r}. \tag{3}$$

Note also that the variance of the difference between x and y is simply $2N$, and this provides an additional reason for defining the noise in this way. Thus the root mean square difference between the two series is a direct measure of the square root of the noise.

The expression (3) can be contrasted with that expected from regression analysis in which the r would be replaced with r^2. In (3) S/N is zero for $r = 0$ and takes on values of 1 for $r = 0.5$, and 9 for $r = 0.9$.

[1] The National Center for Atmospheric Research is sponsored by the National Science Foundation.

4. THE SURFACE DATA SET

A measure of the signal can be seen from the map of standard deviations of monthly mean anomalies from the full surface temperature data set (Fig. 1). In fact this quantity squared depicts the sum of the actual signal plus the noise variance but it is assumed that the former dominates in most places. Largest values are found over the higher latitude continents over the Northern Hemisphere, with more moderate values over the subtropical continents of the Southern Hemisphere. Values exceed 1°C over the oceans only in the tropical eastern Pacific where the dominant signal is known to be the El Niño phenomenon.

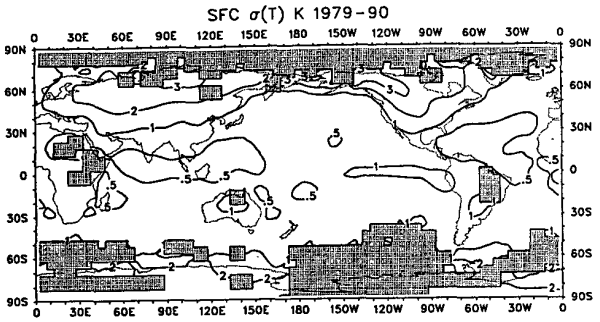


Fig. 1 Standard deviation over 144 months from 1979-1990 of SFC temperature anomalies. Contours are 0.5, 1, 2 and 3 K. Areas of missing data are indicated by stippling.

Consequently, over land the signal tends to be fairly strong and easier to detect than over the oceans. Therefore, we will focus on the oceans, which cover about 70% of the globe, in analyzing the signal-to-noise ratio. Over the oceans there are questions about the SFC data set because the coverage and sampling vary greatly in time and space. The focus is on in situ observations over the past decade or so.

a. Noise in monthly mean SSTs

The sources of errors in estimating monthly mean SSTs from ship data can be divided up as follows.

(i) *Errors in making individual observations.* There is a mix in the way observations are taken, from samples of water collected in buckets and from measurements in the engine room from the intake of water used to cool the engines. The latter temperatures average ~0.3°C higher but the bias varies considerably depending on the size of the water intake and depth below the surface. The thermometer calibration and exposure in the engine room also account for part of the inaccuracies. Based on several studies, and confirmed later, the standard error for individual observations is ~1°C.

(ii) *Incomplete sampling of the diurnal cycle.* At times when winds are light, diurnal heating can exceed 2°C, but mostly it is less than 0.5°C. North of Australia, Morrissey (1990, see his Fig. 10) finds a mean amplitude of ~0.15°C. The diurnal cycle is very small, but noticeable, at OWS (= Ocean Weather Station) P in the NE Pacific—maximum in spring with an amplitude of .095°C.

(iii) *Incomplete sampling of within-month variance other than the diurnal and seasonal cycles.* The within-month variance has been estimated from spectra at a number of OWSs in the North Atlantic and the within-month variance ranges from .02 to .11°C^2. The distribution of high frequency variance probably varies considerably spatially, however, and is much greater in western boundary currents and over the southern oceans.

(iv) *Within-month mean variance due to the seasonal cycle.* The annual cycle has typical amplitudes for SST of 3°C in the SH and 5°C in NH but only 1°C in the tropics (Shea et al. 1990). In some months, the seasonal trend within a month can be large—e.g., July–June, where it is over 3°C in the North Pacific. Assuming only the 12 month (first) harmonic of amplitude A reveals that the average contribution to within-month variance is .0112 A^2. This assumes a linear trend in each month for 1/12 of a cycle.

(v) *Incomplete sampling of the spatial gradients within a grid square.* This problem is mainly of importance in regions of strong gradients such as in the boundary currents off the east coast of Asia and in the Gulf Stream off the east coast of North America. For instance, at 60°W from 42 to 44°N the gradient in winter is 9°C over 2° latitude. Any minor change in ship tracks across that region can lead to substantial apparent anomalies in SST locally. For a uniform gradient of $2a$°C across 2° latitude the expected variance within a 2° box would be $\frac{1}{3}a^2$ °C^2, if well sampled.

These problems are all present in COADS. However, for the UKMO, Bottomley et al. (1990) attempted to minimize the effects of some of these biases by first generating a climatological background field smoothed in space and time. The individual monthly means were created by analyzing anomalies which were subsequently added to the background field. To reduce the effects of spatial gradients within 5° squares and temporal variations from the annual cycle, a working grid of 1° boxes was used with a temporal resolution of 5 days (pentads).

All of the above five items are sampling issues. If we estimate the sampling variance from the first four items above as 1.0, .01, .1, and .28°C^2 (corresponding to $A = 5$°C), respectively, the total error variance would be 1.39°C^2. Expressed as an anticipated standard error for a monthly mean, this would be 1.2°C for a single observation.

To get a better estimate of how the various sources of error are manifest in practice, the COADS have been used to estimate the total noise from all sources. To do this, the standard deviation of the SST observations within each box for each month of 1979 has been examined. The grid squares are all 2×2° (not 5×5°).

Rather than use the actual standard deviation of observations in each box, which may be contaminated by outliers, all observations from COADS were sorted within each box and month into sextiles with a slight modification to the first and fifth sextiles such that, for a Gaussian distribution, the difference between the first and fifth sextile values is equal to 2 times the standard deviation. Accordingly, a robust estimate of the standard deviation can be obtained (called e in COADS documentation).

An exploratory analysis of e revealed that the values were much as expected, provided there were a reasonable number of observations present and the box was not located in a region of strong SST gradient. Near the Gulf Stream e was inflated relative to other areas, with values in excess of 3°C. Spatial gradients of SST within a box were insufficient to explain the observed increase in variance in these regions, which implies that the temporal variability is also greatly enhanced in such regions and contributes significantly to the within box/month variance.

To summarize the results from all areas, the exploratory analysis was used to define criteria for sampling boxes. Boxes without strong SST gradients were chosen by excluding all boxes where the mean SST difference from any surrounding 2° box was greater than 3°C, so that the contribution from spatial variance within each box is

less than 0.75° C². All boxes with less than 16 observations were excluded. We then averaged over all qualifying boxes in large areas by using

$$< e >= \left[\frac{\sum Ne^2}{\sum N}\right]^{\frac{1}{2}}$$

where the areas chosen are the North Pacific and North Atlantic north of 20°N, South Pacific and South Atlantic south of 20°S, tropical Pacific and tropical Atlantic 20°N to 20°S, and the Indian Ocean which is dominated by observations in the tropics. Table 1 shows the results for some ocean basins for the annual mean, and also gives the mean number of observations used in each month. Largest standard deviations are found in the North Pacific where the annual mean $< e >= 1.42$°C, while the value in the North Atlantic is $< e > = 1.15$°C. Values are consistently larger in the Pacific for all regions. In the tropics, $< e >$ averages close to 1°C. At higher latitudes, modest SST gradients of 2°C across a box could increase that to 1.15°C, so it appears that there is also a small contribution in the extratropics from temporal variability within each month. Other studies have also found larger ship-to-ship differences in the North Pacific than in the North Atlantic and it appears that this is real and arises from a greater diversity of ships from many countries with different calibration standards and quality control checks.

Table 1. Robust estimates of the within 2° latitude × 2° longitude boxes and within-month SST standard deviations in °C for the areas given and average number of observations per month used to make the estimate, for the year 1979. Boxes were excluded if they contained less than 16 observations per month and if the gradient in SST between any adjacent box was greater than 3°C. Also given is the number of 2° boxes used and the average number of observations per box for those that qualified.

Area	N.Atlantic	Tr.Atlantic	S.Atlantic	N.Pacific	Tr.Pacific
Year	1.15	0.95	1.03	1.42	1.05
Obs/Mo	19180	5513	460	19384	2428
No. boxes	353	124	13	435	67
Obs/box	54	44	35	45	36

A measure of noise from individual observations is shown in Table 1. However, to a first approximation, the error in the mean has a variance which decreases by 1/N, where N is the number of observations, so that the standard error of the mean decreases by a factor of $N^{\frac{1}{2}}$. The COADS is used to provide estimates of N over 5° by 5° boxes, the same size as used by Bottomley et al. (1990). Fig. 2 presents $N^{\frac{1}{2}}$ for July 1979. In the North Pacific, $N^{\frac{1}{2}}$ is typically 11 to 13 compared with 15 to 19 in the North Atlantic. The exception is near the coasts of all the northern ocean basins, where $N^{\frac{1}{2}}$ is over 20. On the other hand, $N^{\frac{1}{2}}$ drops to less than 5 over the Pacific south of 10°N and south of the equator in the Atlantic.

Combining values from Fig. 2 with Table 1 for areas without strong SST gradients, an overall estimate of the noise in monthly mean SSTs over 5° boxes can be made. Lowest values of ~0.06°C are found in the North Atlantic. However, because of the strong SST gradients near the Gulf Stream, values there will be a factor of two or more larger. Noise estimates are ~0.1°C in the North Pacific and tropical Indian and Atlantic Oceans, but lie between 0.2°C and 0.3°C for the tropical and South Pacific and South Atlantic north of about 30°S. Farther south, except in ship tracks, the noise generally exceeds 0.5°C.

b. Differences between air and sea surface temperatures

An important factor which must be assessed is the extent to which the surface air temperature anomalies over

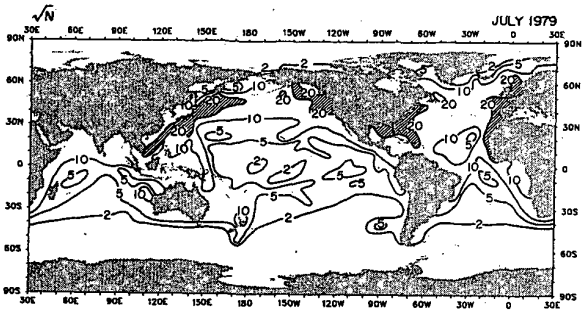


Fig. 2 Map of the square root of the number of SST observations in each 5° by 5° box from COADS for July 1979.

the oceans are represented by SST anomalies. Differences arise because the latter have a larger diurnal cycle than SSTs, and SSTs do not respond strongly to individual cold air outbreaks or to low frequency synoptic disturbances, such as atmospheric blocking, whereas substantial anomalies in monthly air temperature means are associated with such events.

Results from the same analysis of the COADS in Table 1, but with the surface marine air temperature instead of the SST, are given in Table 2. The same SST gradient criterion for acceptance of a box was used but now at least 16 air temperature observations are required. Generally, there are more air temperature than SST observations so more boxes qualify. However, the within-month variance for air temperatures is substantially larger, indicating that there is a need for more observations to reduce the standard error of the monthly mean to acceptable levels. The standard deviation of the within-month air temperature is 20 to 30% larger in the tropics and SH and 43 to 52% larger in the NH. In the North Atlantic on average, for example, a factor of 2.3 more observations of air temperature would be required to produce the same standard error of the monthly mean as for SST, and this factor rises to over 3 in February.

Table 2. Robust estimates of the within 2° latitude × 2° longitude boxes and within-month marine air temperature standard deviations in °C for the areas given and average number of observations per month used to make the estimate, for the year 1979. Boxes were excluded if they contained less than 16 observations per month and if the gradient in SST between any adjacent box was greater than 3°C. Also given is the number of 2° boxes used and the average number of observations per box for those that qualified.

Area	N.Atlantic	Tr.Atlantic	S.Atlantic	N.Pacific	Tr.Pacific
Year	1.75	1.16	1.27	2.03	1.36
Obs/Mo	21284	6289	582	21153	2782
No. boxes	369	136	16	461	74
Obs/box	58	46	36	46	38

We have compared monthly air temperature with SST anomalies from COADS in two regions, the North Atlantic and the North Pacific, where sampling is not a problem. Correlations of SST with surface air temperature anomalies are $r \approx 0.89$ in both oceans, showing that there is a distinct difference between SST and surface air temperature but 80% of the variance of one is captured by the other.

c. Comparison of UKMO and CAC SSTs

As a test of the validity of the above inferences concerning the noise in SST data, the reproducibility of SST anomalies between the CAC and UKMO SST analyzed values has been checked. The CAC analyses are available from 1982 and, because they use satellite data, contain no missing data. The CAC data are therefore included

only if the UKMO data were not missing. Separate annual cycles for the same period were removed using the monthly means, thereby removing possible systematic biases. The CAC anomalies, originally on a 2° grid, were averaged onto the same 5° grid of the UKMO. Area averaging, rather than interpolating, is compatible with the way the UKMO data were computed (Bottomley et al. 1990). The comparison covered 108 months from 1982 to 1990.

The results (Fig. 3) reveal correlations over 0.9 for most of the northern oceans and in the eastern tropical Pacific. Values drop to less than 0.75 south of about 10°N elsewhere and to much less than 0.5 in the central tropical Pacific and over the southern oceans, all regions where the numbers of months of data drop off. Correlations are also low and rms differences high near the coastal regions, evidently because of the land anomaly influence in the SFC data set. Elsewhere, the rms differences between the analyses (Fig. 3) increase from less than 0.2°C over the central North Atlantic to over 0.6°C in the central tropical Pacific, in the eastern Pacific south of 10°S, and generally south of 35°S, except near New Zealand; all areas where the correlation drops to less than ~0.6.

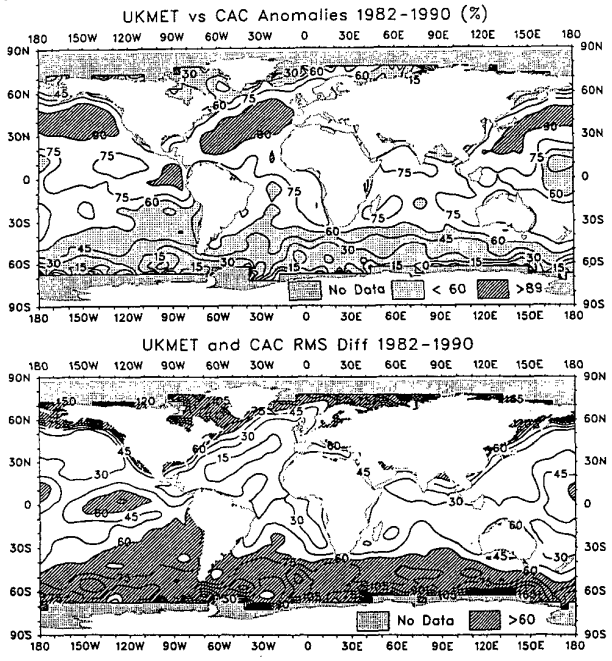


Fig. 3 Top: Correlation coefficient (× 100) between SST anomalies from the UKMO and CAC analyses for 1982 through 1990 (108 months), values greater than 0.9 are cross hatched, values less than 0.6 are stippled, and areas of no data are fine stippled. Bottom: Root mean square differences between the monthly anomalies in hundredths °C, values exceeding 0.6°C are cross hatched.

The results are remarkably compatible with the conclusions drawn earlier. Eq. (3) can be readily applied to the two SST records as they are both ostensibly measuring the identical quantity. For the areas where $r = 0.6$, for example, the standard deviation is ~ 0.7 K so the total variance is ~ 0.5 K^2 of which the signal is $r \times 0.5 = 0.3$ K^2, the noise is $(1 - r) \times 0.5 = 0.2$ K^2, and the latter is compatible with a noise computed from $N \approx 25$ observations and an rms difference of $(2(1 - r) \times 0.5)^{\frac{1}{2}} \approx 0.6$ K, as found in Fig. 3.

5. IMPLICATIONS

The results indicate that the inherent noise level in an SST observation is ~1.0°C and this is compounded when the observation is made in regions of large temperature gradient. Spatial coverage is an issue for SSTs and there is no global coverage. In the historical record, the number of observations decreases back in time, thereby increasing the noise level.

Most of the marine observations have been taken by ships of opportunity and over three-fourths have been taken since World War II. The ships of opportunity follow preferred routes and consequently leave vast areas of the oceans inadequately sampled (e.g., Fig. 4 shows coverage for the month of January for 1930 and 1950). Over parts of the southern oceans, coverage was actually better last century, prior to the opening of the Suez and Panama Canals.

A certain number of observations are needed to carry out gross quality control of SSTs. There are many SSTs reported with values outside physically possible limits and location errors are also common. In "untrimmed" COADS from 1970 to 1979 in 10° boxes, there were over 1000 reports in some landlocked areas and almost all landlocked boxes contained over 100 "observations". Presumably most, but not all, location errors are removed by tests on the data ("trimming").

Because the number of observations and the coverage decreases (Fig. 4), especially prior to 1950, the tendency has been to use as many observations as possible and require only 3 SST (Bottomley et al. 1990; Jones et al. 1991) observations per box to define an anomaly. Moreover, when computing decadal seasonal mean 5° anomalies, Bottomley et al. (1990) required only one month to define a season and 3 seasons (out of 10) to define a decade in computing their decadal seasonal average fields, i.e. perhaps as few as 9 observations and an anticipated noise level of at least 0.35°C.

The use of 5° boxes accumulates observations over a broad area and may cause problems associated with sampling gradients across a box. For January 1930 and 1950, Fig. 4 shows the drop in coverage when 5 observations per month per 2° box are required. Although three times as many observations might be expected over a season, it is clear from Fig. 4 that the observations are neither uniformly nor randomly distributed throughout each 5° box. This gives a more representative estimate of what the reliable historical coverage has been.

There are many other different problems in reconstructing the climate record other than are experienced for the decade from 1979 to 1988. Combined with the problems in correcting for systematic errors, the random error component in SSTs and limited number of observations show how difficult it is to obtain a reliable estimate of the past surface temperatures over the global oceans.

The above comments are all directed toward the level of noise in the surface record. However, the character of the signal may also change. In particular, because of increases in greenhouse gases in the atmosphere, the climate is expected to warm and result in "global warming". If the warming is global then just a few locations should capture that trend. The IPCC (1990) used a "frozen grid" analysis of the temperature trends using only those grid boxes with data from decades in the late 1800s or early 1900s. The results showed that the long-term trends

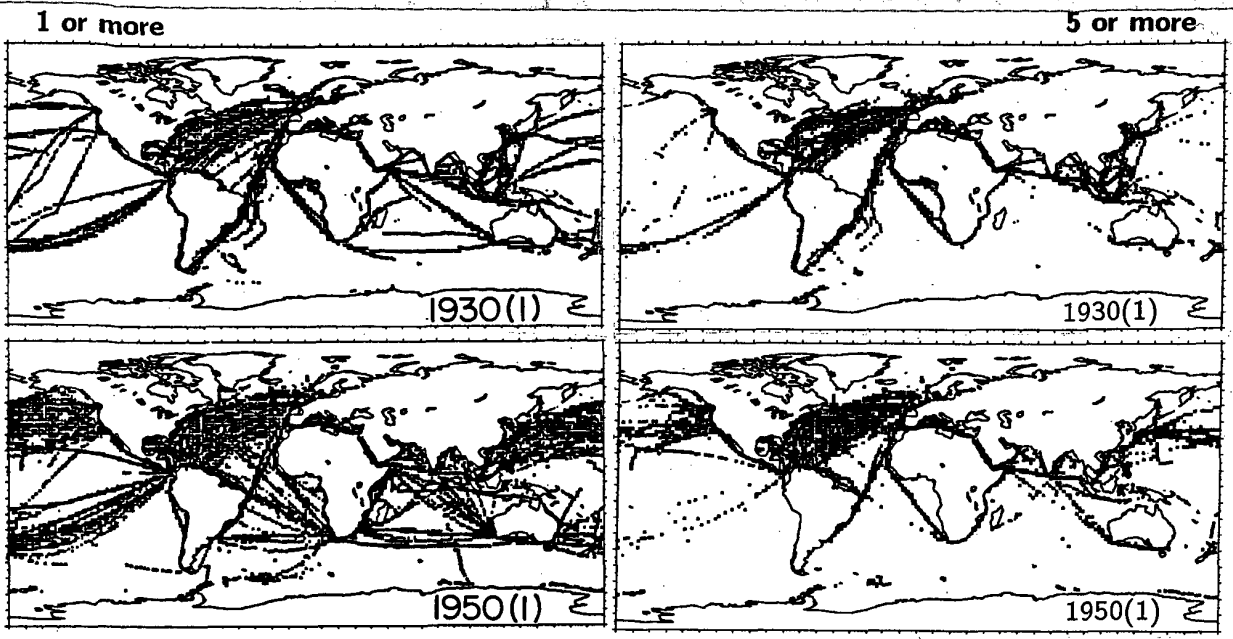


Fig. 4 Ship tracks of SST observations from COADS. In each panel a small x is plotted at the center of gravity of observations of month of January of 1930 and 1950 for (left) one or more, and (right) five or more observations in each 2° box.

in the full data set were well captured. However, in neither the frozen grid nor the full data set are the southern oceans sampled adequately, yet model experiments in which transient increases in carbon dioxide are included show a very slow and different response over the southern oceans. Therefore we cannot be sure that the signal on century time scales is truly global, and the frozen grid analysis says nothing about the more regional and local signals, such as found in the 1980s.

6. CONCLUDING REMARKS

The results have implications for how SST data should be processed into means. The COADS SSTs are assembled into 2° boxes and monthly means, and are therefore subject to significant sampling errors, including errors of type (iv) (not allowing for annual cycle variations during the month), and type (v) (not allowing for gradients across a box). The UKMO SSTs reduce but do not eliminate errors from these sources by using a working grid of 1° and 5-day time periods. Nevertheless, these two kinds of errors could be avoided altogether by defining a good background climatological mean field for each month and calculating the anomaly value for each observation relative to the exact interpolated mean value for that location and that day of the year. Errors of type (ii) from the diurnal cycle might be reduced by ensuring, wherever possible, that uniform sampling of the time of day occurs. Other sampling errors (type (i), measurement errors, and type (iii), real within-month variability) can only be reduced by sufficient numbers of observations to beat down the random component.

With the marked exception of the eastern tropical Pacific, where the large El Niño signal is easily detected, there are insufficient numbers of SST observations to reliably define SST or surface air temperature monthly mean anomalies over most of the oceans south of about 10°N. The use of seasons rather than months can improve the signal-to-noise ratio if careful treatment of the annual cycle is included. For seasonal means, SST anomalies cannot be reliably defined south of 20°S in the eastern Pacific and south of ~35°S elsewhere except near New Zealand.

It must be recalled that the different analyses of past data use the same observations and similar methods, so that just because they exhibit agreement does not ensure that the results are correct. On the contrary, the analysis here suggests that previous claims about our knowledge of past temperatures have been overstated.

Acknowledgments. This work was carried out in collaboration with Jim Hurrell and John Christy. Amy Solomon and Dennis Shea helped to process the data sets. This research was partially sponsored by NASA grants order numbers W-17661 and W-17214.

References

Bottomley, M., C. K. Folland, J. Hsiung, R. E. Newell and D. E. Parker, 1990: Global ocean surface temperature atlas. The U. K. Met. Office. 20 pp and 313 plates.

IPCC, 1990; Scientific Assessment of Climate Change. IPCC WG I, WMO, UNEP. J. T. Houghton, G. J. Jenkins and J. J. Ephraums, Editors. U. Cambridge Press, 365 pp.

Jones, P. D., 1988: Hemispheric surface air temperature variations: Recent trends and an update to 1987. *J. Climate*, **1**, 654–660.

Jones, P. D., T. M. L. Wigley and G. Farmer, 1991: Marine and land temperature data sets: A comparison and a look at recent trends. In"Greenhouse-gas-induced climatic change: A critical appraisal of simulations and observations". M. E. Schlesinger (Editor). Elsevier, Amsterdam. 153-172.

Morrissey, M. L., 1990: An evaluation of ship data in the equatorial Pacific. *J. Climate*, **3**, 99–112.

Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. *J. Climate*, **1**, 75–86.

Snedecor, G. W., and W. G. Cochran, 1973: *Statistical Methods*, Sixth Ed., The Iowa State Univ. Press. 593 pp.

Shea, D. J., K. E. Trenberth and R. W. Reynolds, 1990: A global monthly sea surface temperature climatology. NCAR Tech. Note NCAR/TN-345+STR, 167 pp.

Trenberth, K. E., J. R. Christy and J. W. Hurrell, 1992: Monitoring global monthly mean surface temperatures. *J. Climate*, (in press).

GLOBAL AND HEMISPHERIC TEMPERATURE TRENDS: UNCERTAINTIES RELATED TO INADEQUATE SPATIAL SAMPLING

Thomas R. Karl

Global Climate Laboratory
National Climatic Data Center/NESDIS/NOAA
Federal Building, Asheville, NC 28801.

Richard W. Knight

Global Climate Laboratory
National Climatic Data Center/NESDIS/NOAA
Federal Building, Asheville, NC 28801.

John R.Christy

Atmospheric Sciences
University of Alabama
Huntsville, AL.

April 1992

1.0 INTRODUCTION

Documentation of contemporary variations and changes of the surface global and hemispheric temperature record has been the topic of numerous studies over recent years (Spencer and Christy, 1990; Jones et al., 1986a, b; Jones, 1988; Hansen and Lebedeff, 1987, 1988; Vinnikov et al., 1990). Interest in these variations has been enhanced by efforts to detect the impact of anthropic increases in atmospheric greenhouse gases on the earth's climate (Wigley, 1990; Hansen, 1988; Lindzen, 1990; Schneider, 1990). A thorough understanding of the uncertainties and biases of the thermometric record is a prerequisite for any definitive conclusion regarding greenhouse detection or the identification of other physical forcings that may be affecting the record. This has been recognized by the IPCC (1990) which assessed various sources of errors and biases associated with hemispheric and global temperature records compiled by several research groups for the nineteenth and twentieth centuries. A number of independent sources of error have been identified including: 1) urban heat island bias, 2) changes in observing times, 3) changes in instrumentation, 4) station relocations, and 5) inadequate spatial and temporal sampling. Our analysis focuses on this last item because up until the last few decades most of the globe was not sampled (Fig. 1).

Previous analyses have considered various aspects of the spatial sampling problem (Parker, 1987). For example, Folland and Coleman (1988) used eigenvector patterns to show that the first eigenvector accounts for most of the variance in the sea-surface temperature record. They conclude that sampling inadequacies may not be as detrimental as one might be led to believe, but they urge caution against over-interpretation of their results since they did not have the benefit of complete global coverage to test their conclusion. Jones et al., 1986a performed a number of so-called "frozen-grid" experiments to demonstrate that large systematic biases were not apparent in their analysis when they limited the analysis to only those stations available during the early period of record. Recent work by Trenberth et al. (1992) and Folland (1992) address the issue of intra-monthly and intra-seasonal sampling errors with respect to marine

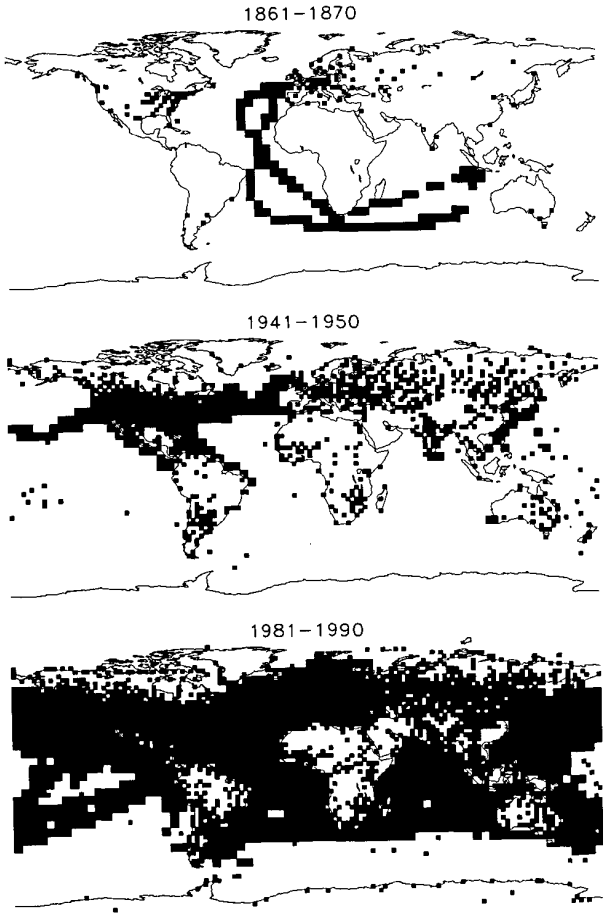


Figure 1 Global coverage of land and ocean temperature measurements for selected decades.

air temperatures and sea-surface temperatures. They accomplish this by comparing the analyses of several marine temperature data sets and surface-based marine observations with space-based measurements from the Microwave Sounding Unit (MSU) and the Advanced Very High Resolution Radiometer (AVHRR) aboard NOAA satellites. They calculate the sampling errors of monthly anomalies of sea-surface ocean temperatures. For the spatial coverage of the most recent decade, Trenberth et al. (1992) find average monthly standard errors of estimate which range from less than 0.1°C for well-sampled 5° X 5° grid cells in the North Atlantic to errors exceeding 0.5°C south of 30°S. Folland et al. (1992) find hemispheric root mean square differences near 0.05°C for two of the most similar types of operational sea-surface temperature analyses, but differences ten times as large for the most dissimilar data sets. Madden (1991) and Madden et al. (1992) used the global temperatures from the NCAR Community Climate Model to estimate the effect of inadequate spatial coverage on estimates of monthly global average temperature anomalies. By using the observed historical sampling of land and ocean temperatures they found standard errors of estimate (root mean square differences) for the month of January which range between 0.05 to 0.22°C, with the largest errors occurring during the Nineteenth Century. In this analysis we attempt to quantify the random and the systematic errors as they relate to our ability to calculate accurate short (10 to 30 years) and long-term (100 to 150 years) trends.

The goal of this article then, is to quantify the effects of incomplete spatial sampling as they affect our ability to calculate long- (100 to 150 years) or short-term (10 to 30 years) trends of hemispheric and global temperatures. We do not attempt to define an optimal observing network, but rather, we mimic the data and techniques used in previous analyses. Despite the acknowledged simplicity of linear trend estimation it is frequently used as an indicator of global and large-scale temperature change (IPCC, 1990;1992; Spencer and Christy, 1990). For these reasons we focus on errors of trend estimation. We consider two types of errors, the errors that arise due to an absence of any observations (TYPE I errors of incomplete geographic coverage), and the errors that arise due to imperfect sampling within grid cells (or averaging areas) which do have observations (TYPE II within grid errors). TYPE I errors are further decomposed by consideration of the random and systematic errors that stem from incomplete geographic sampling. Both TYPE I and TYPE II errors should be independent of the numerous other inhomogeneities which affect the temperature record. Clearly, they are only a portion of all the errors and biases affecting the calculation of hemispheric and global temperature trends which must be considered in any comprehensive error analysis.

2. Data

The data bases which we use to estimate the errors due to spatial sampling inadequacies include; the historical in-situ land based data sets, two marine sea-surface temperature data sets (the Comprehensive Ocean Atmosphere Data Set and the United Kingdom's Marine Data Base), a blended in-situ and satellite sea-surface temperature data set, global data sets of satellite derived mid- and lower-tropospheric temperatures, and the data derived from an ocean-atmosphere general circulation transient model simulation with enhanced levels of carbon dioxide. Table 1 provides pertained information about these data and the abbreviations used to reference the data sets.

Figure 1 depicts the time-varying coverage of the historical surface-based temperature observations. Over the ocean it is based on the GOSTA data base used in IPCC (1990, 1992)[1] and over the land the GHCN Data Set[2] is used. The

Table 1. Data used in the analyses.

MODERN DATA	Period of Record	Authors
Polar orbiting NOAA Microwave Sounding Unit (MSU2R)	1979-1990	Spencer & Christy (1992a, b)
CO_2 Perturbed Transient Model Simulation (OAGCM)	100 Years	Manabe et al. (1991)
Reynolds Blended Satellite/in-situ Sea-Surface Data (BLEND)	1982-1990	Reynolds & Marsico (1991)
HISTORICAL DATA	**Period of Record**	**Authors**
Global Ocean Surface Temperature Atlas (GOSTA)	1851-1990	Bottomley et al. (1990)
Comprehensive Ocean-Atmosphere Data Set (COADS)	1851-1990	Woodruff et al. (1987)
Global Historical Climate Network (GHCN)	1851-1990	Vose et al. (1992)

GHCN is an aggregate of numerous land-surface data bases. The COADS was used to calculate the number of sea surface temperature (SST) observations in each grid cell. A number of simple rules were used to determine the presence or absence of grid cell anomalies for any decade. Anytime a land-based 2.5° X 2.5° grid cell failed to contain a single station with at least 80% of its decadal monthly mean temperatures, it was not used in the global domain. Similarly, for those grid cells which lose continuity of coverage (proceeding backwards in time), they were not allowed to re-enter the sampling grid once they failed to maintain a single station for any decade. The rationale is related to our interest in calculating long-terms trends. These are not exactly the same procedures used by Jones et al. (1986a, b, 1988), Hansen and Lebedeff (1987, 1988), or Vinnikov et al. (1990) in their analysis of global land surface temperature trends. For example, Jones et al. (1988) and Vinnikov et al. (1990) require that all of their stations have data through a common base period in order to calculate an anomaly series. Also, Jones et al. (1986a, b) use a 5° X 10° grid cell while Vinnikov et al. (1990) use optimal averaging methods to aggregate their data into six zones, 0-30°, 30°-60°, and 60°-90° both north and south of the equator. Hansen and Lebedeff (1988) use the longest stations of record within relatively large (e.g. 20° X 30°in mid-latitudes) grid cells to standardize all stations. In this manner Hansen and Lebedeff (1987) make use of all available land stations within any grid cell, but constrain their analysis to a constant difference of temperature between stations in the grid cell. Our selection criteria tend to maximize the use of all available stations, similar to the work of Hansen and Lebedeff (1987), but unlike in Hansen and Lebedeff (1987) we need not rely on a single reference station because our interest is limited to trend estimation. Our initial selection of grid cell size over land was partly driven by the preference for a similar grid size with our global data sets (MSU2R, OAGCM) nonetheless, we test the sensitivity of the results to the grid cell size by introducing biased area-weights.

[1]Somewhat better spatial coverage over the oceans will soon be available as the COADS is merged with the GOSTA data base.

[2]The Global Historical Climate Network Data Base is an ongoing effort of the Carbon Dioxide Information Analysis Center and the National Climatic Data Center. At this date 15 distinct surface temperature data sets have been merged. Several of these are global surface temperature data sets (Jones et al. 1986, Hansen and Lebedeff, 1988), the remainder are national data sets.

Over the oceans the grid cell was increased to 5° X 5°, exactly matching the size used in the IPCC (1990, 1992) and the GOSTA analysis. The procedures used to establish whether a 5° X 5° grid cell was present or absent in any month were exactly the same as to those used by the IPCC (1990, 1992). That is, on a monthly basis a grid cell was considered present even if only a single observation was available for the whole month. In practice, this was a very infrequent occurrence. Additionally, however, a grid cell could not be considered present in any decade whenever it had fewer than 80% of the months present. Unlike the land grid cells, ocean grid cells could swap in and out depending on the decade. This is consistent with the procedures used by Bottomley et al. (1990) and Farmer et al. (1989). The looser criterion is related to the greater homogeneity of the spatial fields of mean SSTs versus land surface temperatures. Using these rules the disparity of the coverage in the Northern versus the Southern Hemisphere is quite apparent (Fig. 2).

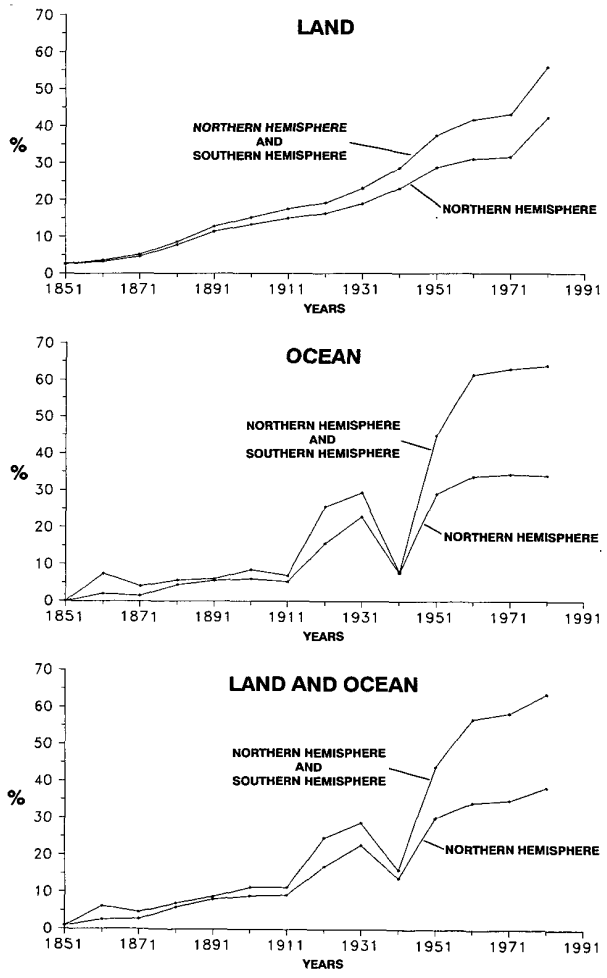


Figure 2 Percent of land, ocean and combined land and ocean with temperature observation.

The data derived from the MSU (Spencer and Christy, 1992a, b), were used to help define the spatial sampling errors. In our analysis we used the so-called MSU2R data set described by Spencer & Christy (1992a, b). These data represent integrated lower- and mid-tropospheric temperatures. The MSU2R data cannot serve as a proxy for the effects of incomplete spatial sampling on the surface temperature. Instead they represent a pattern of change observed over the past decade or more. The pattern of change represented by this data base cannot be extrapolated back or forward in time, but we can test the robustness of the historical coverage to the type of pattern of temperature change represented by the MSU data. It will be important to keep this in mind when we interpret the results. Nonetheless, Elms et al. (1990) have found that the correlation between the MSU tropospheric temperatures and the surface temperatures becomes stronger as the frequency decreases. Moreover, the recent controversy (IPCC, 1990, 1992) regarding the smaller rate of temperature increase in the MSU2R data compared with the IPCC surface-based record make the analysis of these data even more compelling. Other data sets were also used to determine the sensitivity of our analyses on the choice of data used to represent the dependency of geographic coverage on the evolution of the pattern of change. The BLND data set was used to represent a blended data base of space-based and in-situ SST data which had complete global ocean coverage since 1982. Additionally, the global surface-layer temperatures (approximately 70 m above ground level) were used from the Manabe et al. (1991) 100-year transient ocean/atmosphere general circulation model (OAGCM). This model was perturbed with 1% per year increase of carbon dioxide concentration. The output was chosen for analysis because the 100-year northern hemisphere warming trend was 1.3°C larger than the southern hemisphere trend (4.2°C vs. 2.9°C). This asymmetry provided the means to test the affect of the non-uniform spatial sampling of the Southern Hemisphere compared with the Northern Hemisphere, especially the deficiency of observations south of 30°S.

3. PROCEDURES

a) Definitions

The true global (hemispheric) mean temperature anomaly, T_t can be obtained by:

$$T_t = \sum_{g=1}^{NG} \omega_g T_{g,t} \quad , (1)$$

of which ω is the appropriate grid cell weight and NG is the total number of grid cells. An estimate of T_t, with less than complete spatial coverage is calculated by,

$$\hat{T}_t = \sum_{g=1}^{NG-M} \omega_g (T_{g,t} + \epsilon) \quad , (2)$$

where M is the number of missing grid cells and ϵ is a standard normal deviate with zero mean and variance s_ϵ^2.

If the true slope of the temperature time series is represented by b, and an estimate for b is denoted by $\hat{b}$ then a random standard error of estimate for b can be defined by:

$$r.s.e._b = \left(\sum_{i=1}^{NS} (b_i - \hat{b}_i)^2 / NS-1 \right)^{0.5} \quad , (3)$$

where NS is the number of simulations (or estimates) of b. In our analysis, trends are calculated over several different time intervals, τ, where τ takes on values of 10, 30, 50, 100 or 150 data points (years). The quantities, b_i and $\hat{b}_i$ are derived from a simulated time series (Fig. 3) of global (or some other large region, e.g. hemispheric) mean temperature anomalies, $\tilde{T}$. The quantity $\tilde{T}$ at time, t, is defined by:

$$\tilde{T}_t = a_t + B(\alpha_{t,k} + (\delta_k + E_{t,k})) \quad . (4)$$

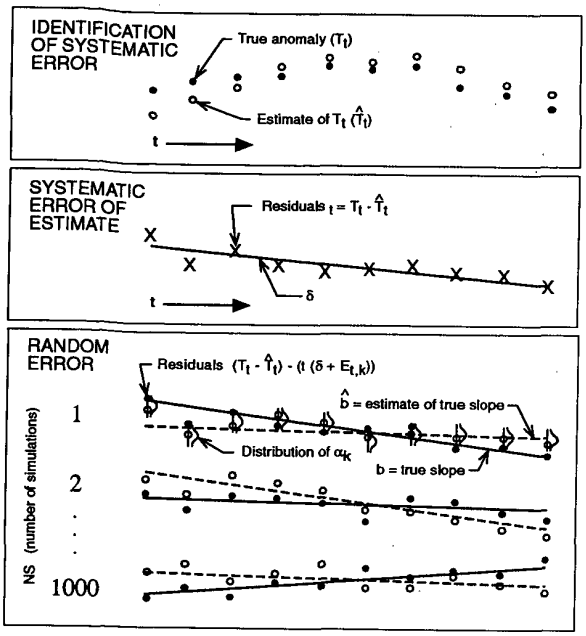


Figure 3 Schematic of the technique used to calculate the systematic and random symbols are: standard error of estimate of the true slope of a trend.

In Eq. (2) a_t is a random standard normal deviate with zero mean and variance equal to the observed annual global (hemispheric) mean temperature anomalies, $\alpha_{t,k}$ is the random standard error of estimate defined by a standard normal deviate with zero mean and variance equal to s.e.^{2}, t is the time increment (1 year), k represents a specific decade, e.g., 1850s, 1860s,...1980s, δ is the systematic error of estimate, E is a standard normal deviate with mean zero and variance equal to the standard error of the slope, δ, and B is a binary variable equal to zero when $\hat{T}_t$ is used to calculate b_i or unity when $\hat{T}_t$ is used to calculate $\hat{b}_i$. The systematic error of estimate, δ_k, is defined as the trend of the residuals, $T_t - \hat{T}_t$, where $\hat{T}_t$ is an estimate of the true global mean temperature anomaly based on the observed spatial coverage for decade, k. The random standard error of estimate of the estimated global mean temperature anomalies is defined by,

$$s.e._T = \left(\sum_{t=1}^{n} \left((T_t - \hat{T}_t) - t(\delta_k + E_{t,k}) \right)^2 / n - 1 \right)^{0.5} \quad , \quad (5)$$

where n is the time series length.

The systematic standard error of estimate for b, $s.s.e._b$ is given by:

$$s.s.e._b = \delta + E \quad . \quad (6)$$

Upper and lower bounds of the cumulative systematic and random standard errors of estimate of b are defined by:

$$u.s.e._b = r.s.e._b + (-s.s.e.)_b \quad \text{and (7a)}$$

$$l.s.e._b = -r.s.e._b + (-s.s.e.)_b \quad . \quad (7b)$$

The negative sign is used for convenience in Eqs. 7a and 7b associated with the systematic standard error of estimate. In this manner an overestimate (underestimate) of the slope occurs when $u.s.e._b$ and $l.s.e._b$ span positive (negative) slopes.

In Eq. 2, for Type I error analyses $s_\epsilon^2 = 0$, but for the cumulative effect of Type I and Type II errors $s_\epsilon^2 \neq 0$. The standard deviation of ϵ was calculated separately for ocean and land temperature anomalies. For land grid cells,

$$s_{\epsilon,mo,g} = \mu_{mo,g} / \sqrt{STNS} \quad , \quad (8)$$

In Equation (8) STNS is the number of stations in a grid cell, mo is the month of interest, and μ is defined by:

$$\mu_{mo,g} = \left(\sum_{j=1}^{STNS} (T_{g,mo} - \hat{T}_{g,j,mo})^2 / STNS - 1 \right)^{0.5} \quad , \quad (9)$$

where $T_{g,mo}$ is the grid cell temperature anomaly for month, mo, and $\hat{T}_{g,j,mo}$ is an estimate of the true grid cell temperature anomaly. In Eq. 9 $\hat{T}_{g,mo}$ is calculated using a finite, but large number (between 4 and 8 stations) of stations in grid cell, g. Over the oceans, Trenberth et al. (1992) provide an estimate for $\mu_{mo,g}$, and we use the COADS to derive the historical values of STNS in Eq. 8. This is simply the number of year-month SST observations in each 5° X 5° grid cell.

b) Type I errors (incomplete geographic coverage)

Numerous experiments were performed to understand the impact of incomplete geographical coverage on the trend. Table 2 summarizes all of these experiments. The experiments can be divided into three distinct categories based on the data set used to determine b_i and $\hat{b}_i$. The MSU, GCM, and BLND experiments represent different estimates of surface temperature anomaly estimation. Each data set has a unique set of temperature

Table 2. A description of the experiments performed to estimate the standard error of estimate for the slope of 10, 30, 50, 100, and 150-year trend estimates. Abbreviations are: MSU2R implies the Spencer and Christy (1992 a,b) space-based data set, OAGCM implies the ocean-atmosphere general circulation model transient run of Manabe et al., 1991, BLEND implies the in-situ/space-based blended data set of Reynolds and Marsico, 1991, Lnd implies land, Ocn implies ocean, NH implies Northern Hemisphere, SH implies Southern Hemisphere, Glbl implies Global, Types imply the Type of error estimate (incomplete coverage---I or inaccurate grid cell anomalies---II), and Exp Num implies experiment number, and B represents a biased area weighted experiment using the method of Vinnikov et al. (V) or Hansen and Lebedeff (H).

HISTORICAL						VS	TRUE					Types		Exp Num
Data Set	Lnd	Ocn	NH	SH	Glbl		Lnd	Ocn	NH	SH	Glbl	I	II	
MSU2R	x				x		x	x			x	x		1MSU
MSU2R		x			x		x	x			x	x		2MSU
MSU2R	x				x		x				x	x		3MSU
MSU2R	x				x		x				x	x	x	4MSU
MSU2R		x			x			x			x	x		5MSU
MSU2R		x	x					x	x			x		6MSU
MSU2R		x		x				x		x		x		7MSU
MSU2R	x	x			x		x	x			x	x		8MSU
MSU2R	x	x			x		x	x			x	x	x	9MSU
MSU2R	x	x			x		x	x			x	x	x	9MSUBV
MSU2R	x	x			x		x	x			x	x	x	9MSUBH
MSU2R	x	x	x				x	x	x			x		10MSU
MSU2R	x	x		x			x	x		x		x		11MSU
OAGCM	x				x		x	x			x	x		1GCM
OAGCM		x			x		x	x			x	x		2GCM
OAGCM	x				x		x				x	x		3GCM
OAGCM	x				x		x				x	x	x	4GCM
OAGCM		x			x			x			x	x		5GCM
OAGCM		x	x					x	x			x		6GCM
OAGCM		x		x				x		x		x		7GCM
OAGCM	x	x			x		x	x			x	x		8GCM
OAGCM	x	x			x		x	x			x	x	x	9GCM
OAGCM	x	x			x		x	x			x	x	x	9GCMBV
OAGCM	x	x			x		x	x			x	x	x	9GCMBH
OAGCM	x	x	x				x	x	x			x		10GCM
OAGCM	x	x		x			x	x		x		x		11GCM
BLEND		x			x			x			x	x		5BLND
BLEND		x	x					x	x			x		6BLND
BLEND		x		x				x		x		x		7BLND

anomalies and trends. Moreover, the MSU data do not exactly represent the surface temperature, the blended space based/in-situ data set pertains to SST temperatures only, and although the GCM model output provides an interesting test of the robustness of trend estimation over the past two Centuries with asymmetric hemispheric trends, it represents an exaggerated upper bound. For the reasons we pay particular attention to the range of the results among the MSU, GCM, and BLND experiments.

Aside from the categorization of the experiments based on data set type, we performed a number of experiments to help understand the how various partitions of geographic coverage effect the results. For example, we perform experiments using only the historical hemispheric or global coverage over the oceans or land separately, and compare the results to complete hemispheric or global coverage over the combined ocean and land areas. This partition aids in understanding the historical effects of incomplete observational coverage over the oceans versus the land and over the Northern versus the Southern Hemisphere. These experiments are important with respect to the magnitude of bias corrections that are required in the historical ocean and land temperature data sets (Jones et al., 1986a; Bottomley et al., 1990) due to changing observing methods relative to the error introduced by completing ignoring the land or ocean temperatures.

In our experiments we summarize the effect of incomplete sampling by reporting the range of μ.s.e.$_b$ and l.s.e.$_b$ value of s.e.$_b$ from 1000 simulations over several time periods ranging from 10 to 150 years. The changes in the observed historical coverage of observations from decade-to-decade drive the sign and magnitude of the Type I errors relative to the evolution of temperature change in each of the three data sets. When the trend is calculated beyond 1990, the spatial coverage of the 1980s is used for all subsequent decades.

c) Type II errors (within grid errors)

Four different experiments were specifically performed to determine the cumulative effect of both Type I and Type II errors. These experiments are listed in Table 2, and they pertain to global land and ocean and global land-only experiments (4MSU, 9MSU, 4GCM, and 9GGM). The cumulative effect of both the time varying geographical coverage of the globe and the time varying standard errors of estimate for grid cell anomalies determine the magnitude of the standard errors of estimate associated with s.e.$_b$.

Trenberth et al. (1992) calculate the standard errors of estimate for monthly mean SST anomalies for a single SST observation in a 5° X 5° ocean grid cell. The regionally dependent, monthly varying estimates of $\mu_{mo,g}$ provided by Trenberth et al. (1992) (Fig. 4) were interpolated and extrapolated to cover all ocean basins. Strictly speaking, this will be an underestimate of the standard error of estimate in some regions because Trenberth et al. (1992) purposely avoid the specification of standard errors of estimate in areas where there are strong gradients of SST. Although initially troubling, we shall see that this does not appear to be an important contribution to the u.s.e.$_b$ or the l.s.e.$_b$ in the historical record.

Over the land $\mu_{mo,g}$ was calculated by sampling a few grid cells which had a sufficient number of observations over a 2.5° X 2.5° grid cell to represent an integrated estimate of the monthly temperature anomaly. We identified three grid cells across North America where we had 4 to 8 stations available each month for a period of 20 years (Fig. 4). Monthly average standard errors of estimate were calculated for each grid cell by

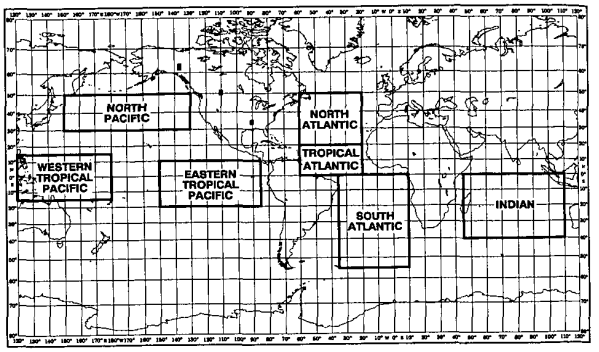


Figure 4 Land grid cells and ocean regions over which standard errors of estimate ($\mu_{mo,g}$) were calculated (Oceans from Trenberth et al., 1992).

Table 3. Standard errors of estimate, $\mu_{mo,g}$, (°C) of the monthly temperature anomaly over land and ocean using a single station (land) or a single observation (ocean) for selected regions. Figure 4 depicts the regions and grid cells.

Month	Land			Ocean		
	South	Central	North	N. Atlantic	N. Pacific	S. Pacific
Jan	0.57	1.16	3.57	1.14	1.34	1.26
Feb	0.56	1.12	3.17	1.10	1.31	1.34
Mar	0.46	0.89	1.78	1.03	1.28	1.28
Apr	0.62	0.60	0.67	1.05	1.34	1.29
May	0.62	0.54	0.91	1.27	1.45	1.23
Jun	0.58	0.53	0.65	1.29	1.59	1.10
Jul	0.56	0.52	0.60	1.21	1.51	1.17
Aug	0.58	0.55	0.62	1.03	1.35	0.95
Sep	0.57	0.60	0.54	1.05	1.34	1.05
Oct	0.60	0.59	1.46	1.23	1.56	1.29
Nov	0.50	0.87	2.91	1.13	1.48	1.32
Dec	0.56	1.15	3.44	1.19	1.41	1.36

systematically using only 1 station to estimate the monthly anomaly. The magnitude of $\mu_{mo,g}$ is listed in Table 3. Given the similarity of the warm season estimates of $\mu_{mo,g}$ among all three grid cells, and the absence of an annual cycle in the southern most grid cell, we used the grid cell values to represent zonal averages. The northern grid cell represented all land areas above 55°N(S), the southern grid cell those areas below 35°N(S), and the central grid cell represented the remaining land areas.

d) Weighted global and large-scale anomalies

The historical coverage of temperature sampling has not been spatially random (Figs. 1 and 2). Because of this fact, previous analyses of global and hemispheric temperature change have directly used biased area weights to calculate global and hemispheric temperature anomalies (Vinnikov et al., 1990), and others have indirectly used biased area weights (Hansen and Lebedeff, 1987;1988). In order to ascertain the sensitivity of our results to unbiased and biased area weights we replicate some of our experiments with biased area weights. This is accomplished by modifying Eq. 2 as given by:

$$\hat{T}_t = \sum_{i=1}^{NZ} w \left(\sum_{g=1}^{Ng-m} \omega_g (T_{g,t} + \epsilon) \right) \quad , (10)$$

where NZ is the number of zones, Ng is the number of grid cells in each zone, and w is the appropriate area weight of the zone.

In two of our simulations (9MSUBV and 9GCMBV, Table 2) we let Nz=6 where each zone corresponds to a zonal average. These simulations are

distinguished from the unbiased simulations listed in Table 2 by adding the letter "B" and the first inial of the research team at the end of the experiment number. The zones used were the same as those used by Vinnikov et al. (1990) for their time series. This includes the zones 0°-30°, 30-60°, and 60°-90° north and south of the equator. The rationale of using biased area weights relates to the fact that unbiased area weights inevitably give little chance for the poorly sampled areas of the globe (e.g., Southern Hemisphere) to have an influence on the global average when they are considered with the well sampled middle latitudes of the northern hemisphere. In two other simulations (9MSUBH and 9GCMBH, Table 2) we let NZ=80 where the zones are equal areas, similar to those used by Hansen and Lebedeff (1987). In their analysis, all grid cells or zones occupy a relatively large area e.g., 20° X 30° in the middle latitudes regardless of the number of stations in each large grid cell. The approach used by Hansen and Lebedeff (1987, 1988) effectively biases the area-weight of those grid cells which have only a few, or even one station, in a large area relative to those grid cells which have adequate spatial coverage.

4. Results

The confidence intervals of the standard errors of estimate are positively biased for all slopes calculated over each of the five time intervals (10-year to 150-year trends) using the patterns of change derived from the MSU data (Figs. 5 through 9). The magnitude of the bias is particularly large (over 1°C/100 years) for trends of 50 years or less (Figs. 7 through 9) calculated from the beginning of the Nineteenth Century. This positive bias occurs in simulations which include both Type I and Type II errors as well as in simulations which include only Type I errors. The difference between the confidence intervals in the two simulations with and without Type II errors (8SMU and 9MSU) is practically insignificant. This implies that the random errors introduced into the historical record by imperfect sampling within 2.5° X 2.5° land grid cells or 5° X 5° ocean grid cells tend to cancel when many grid cells are used and trends are calculated over at least a 10-year period. On the other hand, the magnitude of the systematic standard error of estimate due to incomplete spatial sampling is relatively large, and remains significant even over the most recent period of record.

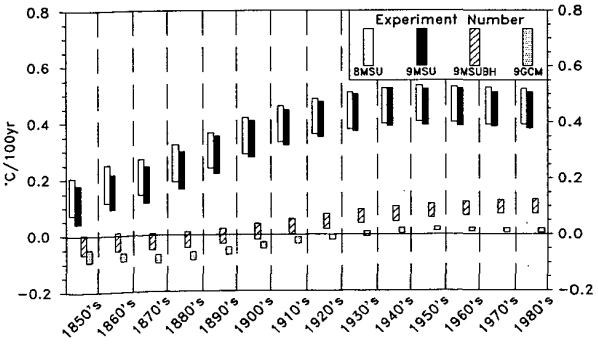


Figure 5 Confidence intervals of the standard errors of estimate (C°/100 years) for the slope of a trend calculated from the beginning of a decade through the subsequent 150 years. Results represent the observed historical global coverage over both the land and ocean versus complete land and ocean global coverage.

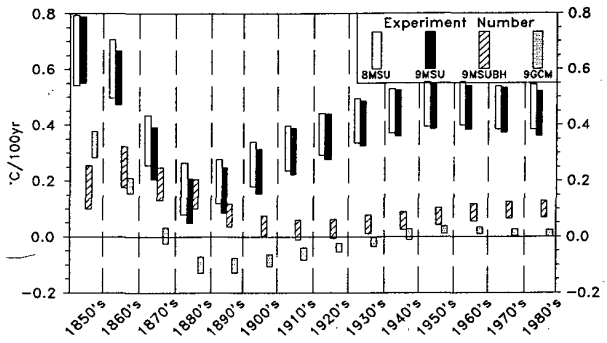


Figure 6 Same as Fig. 5 except for 100-year trends.

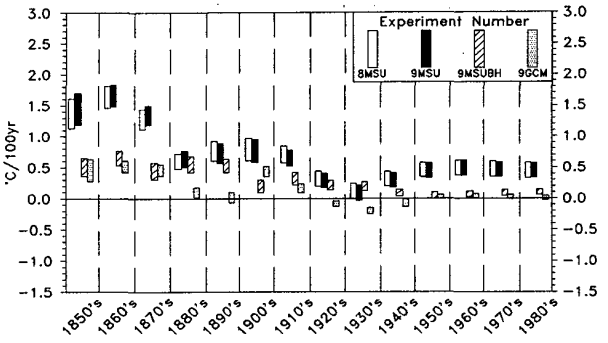


Figure 7 Same as Fig. 5 except for 50-year trends and the ordinate has a scale that spans 5 times the slopes.

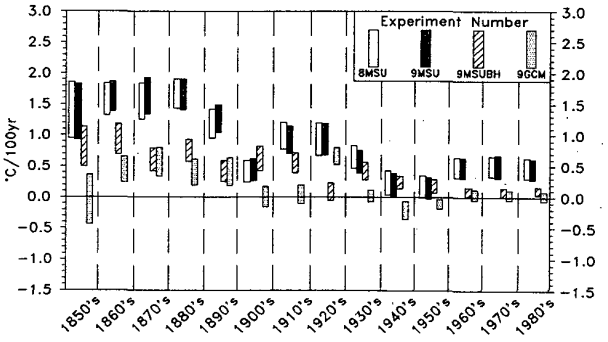


Figure 8 Same as Fig. 7 except for 30-year trends.

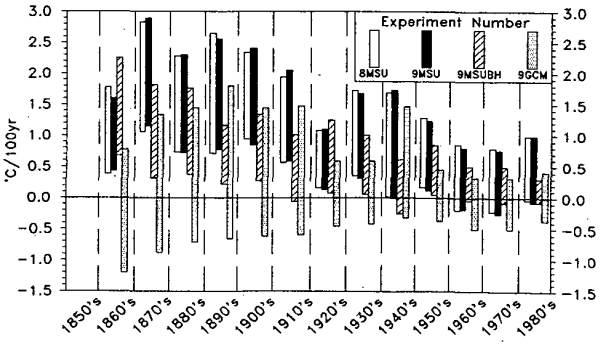


Figure 9 Same as Fig. 7 except for 10-year trends.

It is not difficult to understand why the relatively large positive bias of the global experiments, 8MSU and 9MSU, persist even through the time interval of the most comprehensive geographical coverage (1980s). Much of the warming in the MSU2R data is confined to mid-latitudes, especially the Northern Hemisphere (Table 4). A significant cooling is observed over the Tropics (25°N to 25°S) and the southern polar regions. The oversampling of the Northern Hemisphere mid-latitudes and the undersampling of the tropics and South Polar region accounts for the positive bias. By contrast when biased area average weights are used, similar to the procedure adapted by Hansen and Lebedeff (1987,1988), the oversampling of the mid-latitude warming in the Northern Hemisphere is significantly mitigated. It is noteworthy that the positive bias we find, even in the most comprehensive period of spatial coverage, is consistent with the hypothesis that at least some of the difference between the smaller temperature increase in the MSU data versus the surface temperatures, as reported by IPCC (1992), may be attributed to inadequate geographic coverage. Perhaps just as important however, the experiments suggest that in future IPCC assessments an effort should be made to bias the global area averages to account for non-random sampling.

Table 4 True zonal trends (°C/100 years) over the full period of record of the three data sets used in this study.

Zone	% Area of Globe	MSU	GCM	BLND
55°N-90°N	8.5	0	6.4	2.9
25°N-55°N	19.5	2.4	4.2	3.4
25°S-25°N	44	-1.6	3.1	0.4
55°S-25°S	19.5	1.4	2.9	2.0
90°S-55°S	8.5	-1.4	2.1	-0.1

Contrary to the results from the MSU simulations and perhaps somewhat surprisingly, the AOGCM 100-year patterns of temperature change (9GCM) was relatively well captured (standard errors of estimate of only a few hundreds of degrees and straddling near zero), even for the geographical coverage dating back to the turn of the Century (Figs. 5 and 6). This is in contrast to the rather large biases associated with the pattern of the MSU2R temperature change. The undersampling of the strong warming North of 55°N and the mild warming south of 25°S, relative to the other parts of the globe (Table 4), resulted in reasonably good estimates of global temperature trends. Nonetheless, the asymmetric interhemispheric temperature trend of the AOGCM does result in relatively large standard errors (as high as 1°C/100 years) of estimate and significant positive bias when estimating trends of 100-years or less and beginning in the Nineteenth Century.

5. Conclusions

Additional simulations will be presented and discussed at the Conference which confirm the potentially serious impact that incomplete spatial sampling can have on the calculation of global temperature trends. One of the most important findings of our work pertains to the large error that can be introduced into trend estimation by non-random sampling. This type of error dominates over the random errors when estimating trends. Our results also help explain the positive bias of the recent surface temperature trends, calculated from the historical records, relative to the recent MSU2R temperature trends reported by IPCC (1992). The positive bias could be mitigated however, by using biased area average estimates over both the land and the ocean, similar to that used by Hansen and Lebedeff (1987, 1988). Nonetheless, for 100 year trends the standard error of estimate for the slope of the trend is often as large as 0.1°C/100 years, and prior to this time the standard error of estimate for 100 to 150 year trends was often in the range of 0.2 to 0.4°C/100 years. Considering the magnitude of the global trend from the AOGCM (3.6°C/100 years), these errors are still an order of magnitude smaller than the observed trend, but for the size of the global trends in the MSU data (0.3°C/100 years) they are quite significant. If the observed pattern of the MSU2R global temperature change (1979-90) were projected onto to the surface and the observed spatial coverage of 1980s persisted for the next 100 years the slope of the global temperature trend would be seriously overestimated, unless steps were taken to mitigate the non-random sampling.

Unfortunately, it will never be possible to be certain about the magnitude of the errors that may have been introduced into the historical record due to incomplete spatial sampling. We will never know the true evolution of the spatial patterns of temperature change. Nonetheless, the relatively small errors associated with estimating the temperature change from the AOGCM data set is encouraging, despite the large asymmetric hemispheric temperature change. Moreover, biased area estimates of the global temperature anomalies can help compensate for non-random sampling effects. Trends calculated over short periods, 10 years or less, have large standard errors of estimate regardless of the period chosen.

Our results point to the critical need of a well conceived global network of surface-based observing stations if we ever hope to obtain highly reliable estimates of surface boundary layer temperature trends. The urgency of this effort is underscored by the less than satisfactory attempts at space-based monitoring of the surface-layer diurnal temperatures.

6. Acknowledgements

This work has been supported by NOAA's Office of Global Programs and a United States Dept. of Energy/Dept. of Commerce Interagency Agreement.

Thanks goes due Dr. Suki Manabe for providing his model results in this study.

References

Bottomley, M., C.K. Folland, J. Hsiung, R.E. Newell and D.E. Parker, 1990: Global Ocean Surface Temperature Atlas (GOSTA). Joint Meteorological Office/Massachusetts Institute of Technology Project. Project supported by US Dept of Energy, US National Science Foundation and US Office of Naval Research. Publication funded by UK Dept of Energy of Environment. 20+iv pp and 313 Plates. HMSO, London.

Elms, J., R.R. Karl, A.L. McNab, J.R. Christy, and R.W. Spencer, 1990: A preliminary analysis of the covariation of surface and mid tropospheric temperatures. Prepared for 41st International Astronautical Congress, Dresden, Germany, Oct. 1990.

Farmer, G., T.M.L. Wigley, P.D. Jones and M. Salmon, 1989: Documenting and explaining recent global-mean temperature changes. Climatic Research Unit. Norwich, Final Report to NERC, UK, Contract GR3/6565.

Folland, C.K., and A.W. Coleman, 1988: An interim analysis of the leading covariance eigenvectors of worldwide sea surface temperature anomalies for 1951-80. LRFC 20, available from National Meteorological Library, Meteorological Office, Bracknell, UK.

Folland, C.K., R.W. Reynolds, M. Gordon, and D.E. Parker, 1992: A study of six operational sea surface temperature analyses. J. Clim. IN REVIEW

Hansen, J. and S. Lebedeff, 1987: Global trends of measured surface air temperature. J. Geophys. Res., 92, 13345-13372.

Hansen, J. and S. Lebedeff, 1988: Global surface temperatures: update through 1987. Geophys. Res. Lett., 15, 323-326.

Hansen, J.E. and A.A. Lacis, 1990: Sun and dust versus the greenhouse. Nature, 346, 713-719.

Intergovernmental Pamel on Climate Change, IPCC, 1990: Climate Change, The IPCC Scientific Assessment. Eds. J.T. Houghton, C.J. Jenkins, and J.J. Ephraums. Cambridge University Press, 365 pp incl. Appendices.

Intergovernmental Pamel on Climate Change, IPCC, 1992: Climate Change, The IPCC Scientific Assessment. Eds. J.T. Houghton, B.A. Callendar, and S.K. Varney. Cambridge University Press, 198 pp incl. Appendices.

Jones, P.D., 1988: Hemispheric surface air temperature variations: recent trends update to 1987. J.Clim., 1, 654-660.

Jones, P.D., S.C.B. Raper, R.S. Bradley, H.F. Diaz, P.M. Kelly, and T.M.L. Wigley, 1986a: Northern hemisphere surface air temperature variations, 1851-1984. J. Clim. Appl. Meteor., 25, 161-179.

Jones, P.D., S.C.B. Raper, R.S. Bradley, H.F. Diaz, P.M Kelly, and T.M.L. Wigley, 1986b: Southern hemisphere surface air temperature variations, 1851-1984. J. Clim. Appl. Meteor., 25, 1213-1230.

Lindzen, R.S., 1990: Some coolness concerning global warming. Bull. Amer. Meteor. Soc., 71, 288-299.

Madden, R.A., D.J. Shea, G.W. Branstator, J.J. Tribbia, and R. Weber, 1992: The effects of imperfect spatial and temporal sampling on estimates of the global mean temperature: Experiments with model and satellite data. IN REVIEW J. Climate.

Manabe, S., R.J. Stouffer, M.J. Spelman, and K. Bryan, 1991: Transient responses of a coupled-ocean atmosphere model to gradual changes of atmospheric CO_2. Part I: Annual mean response. J. Clim.,4, 785-818.

Parker, D.E., 1987: The sensitivity of estimates of trends of global and hemispheric marine temperatures to limitations in geographical coverage. LRFC 12, UK Meteorological Office, London Road, Brachnell, Berkshive RG12 257.

Reynolds, R.W. and D.C. Marsico, 1992: An improved real-time global SST analysis. IN PRESS J.Clim.

Schneider, S.H., 1990: The global warming debate heats up: Analysis and perspective. Bull. Amer. Meteor. Soc., 71, 1292-1304.

Spencer, R.W. and J.R. Christy, 1990: Precise monitoring of global temperature trends from satellites. Science, 247, 1558-1562.

Spencer, R.W. and J.R. Christy, 1992: Precision and modiosonde validation of satellite grid point temperature anomalies, Part I: MSU Channel 2 J. Clim., IN PRESS.

Spencer, R.W. and J.R. Christy, 1992, Precision and modiosonde validation of satellite grid point temperature anomalies, Part II: tropospheric retrieval and trends during 1979-90. J. Clim., IN PRESS.

Trenberth, K.E., J.R. Christy, and J.W. Hurrell, 1992: Monitoring global monthly mean surface temperature. J. Clim., IN PRESS.

Vinnikov, K. Ya,, P. Ya Groisman, and K. M. Lugina, 1990: The empirical data on modern global climate changes (temperature and precipitation). J. Clim., 3, 662-677.

Vose, R.S., Richard L. Schmoyer, Peter M. Steurer, Thomas C. Peterson, Richard Heim, Thomas R. Karl, and J. Eischeid. 1992. The Global Historical Climatology Network: Long-term Monthly Temperature, Precipitation, Sea Level Pressure, and Station Pressure Data. ORNL/CDIAC-53, NDP-041. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Wigley, T.M.L., 1991: Could reducing fossil-fuel emissions cause global warming? Nature, 349, 503-506.

Woodruff, S.D., R.J. Slutz, R.L. Jenne, P.M. Steurer, 1987: A comprehensive ocean-atmosphere data set. Bull. Amer. Meteor. Soc., 68, 1239-1250.

APPLICATION OF STATISTICAL TECHNIQUES FOR THE REGIONALIZATION OF MARINE SURFACE TEMPERATURE

Henry F. Diaz

NOAA/ERL, 325 Broadway, Boulder, CO 80303

Timothy J. Brown

CIRES, University of Colorado, Boulder, CO 80309

1. INTRODUCTION

We present the results of our efforts in applying a number of objective statistical techniques to define homogeneous climatic regions over the world oceans. The principal variable used is monthly and seasonal sea surface temperature (SST) for the period 1950-89. The reasons for exploring the nature of "climatically homogeneous" oceanic regions are two-fold. First, from the point of view of climate monitoring and climate change detection efforts, it is useful to divide the ocean into smaller units; second, there are a number of air-sea interaction processes which might be advantageous to study on regional scales. It is also of interest to compare SST changes in the different ocean basins.

2. METHODOLOGY

Using four-degree latitude/longitude area boxes from the COADS data set (Woodruff, et. al., 1987), a T-mode principal component analysis (PCA) was performed using the calendar month 1950-89 long-term mean SSTs in each box as the variables, with a minimum requirement of five years of data within the 40-year period in each box. The T-mode PCA technique was first used by Stidd (1967) to classify the precipitation climate of Nevada, and also by Skaggs (1975) to study the 1930s drought in the U.S.

K-means cluster analysis (Hartigan, 1975) was then used to classify the gridbox SST data, using as input the loadings from the first two eigenvectors of the monthly mean SST. Initially, a total of 1,939 four-degree boxes were used as input to the clustering algorithm, with a prescribed seed of 15 groups. Boxes in certain inland seas, and areas with relatively sparse data coverage were manually edited out of the final regional configuration, eliminating 123 boxes for a final total of 1,816 four-degree boxes yielding 34 regional clusters. In general, distinct regions were chosen for each hemisphere and ocean basin. The resulting cluster patterns were visually examined to determine if the physical boundaries were climatologically consistent, and various tests were performed to ascertain the temporal coherence of the constituent boxes for each region. This involved, for instance, comparing the distribution of the departure values of the individual four-degree boxes and calculating the mean inter-box correlation within each region. Figure 1 shows a world map of the 34 cluster regions that were classified in this manner, and Table 1 presents a few basic summary statistics for each region.

Table 1. Selected summary statistics based on annual SST means for each region. The ordinary least squares trend is given in °C/decade.

Region	Mean	S.D.	OLS trend	t-value
1	6.06	.50	-.37	-10.05
2	7.56	.58	-.37	-6.89
3	9.45	.26	-.16	-6.29
4	8.14	.63	-.36	-5.57
5	9.19	.57	-.25	-3.59
6	16.73	.23	-.04	-1.41
7	19.09	.28	-.08	-2.14
8	15.65	.54	-.27	-4.49
9	18.92	.41	-.23	-5.45
10	17.72	.26	-.01	-.40
11	19.79	.32	-.11	-2.78
12	23.21	.24	-.08	-2.69
13	23.49	.20	-.08	-3.15
14	25.36	.16	-.02	-.80
15	24.61	.17	.02	.87
16	27.78	.24	.13	4.97
17	27.51	.21	.06	2.45
18	27.06	.23	-.02	-.55
19	27.86	.23	.10	3.90
20	24.32	.49	.12	1.79
21	27.19	.20	.09	3.65
22	25.38	.28	.14	4.31
23	25.75	.28	.17	6.25
24	25.53	.21	.06	2.19
25	20.61	.57	.16	2.20
26	20.74	.33	.20	6.19
27	20.82	.22	.03	.85
28	21.05	.31	.18	5.99
29	15.74	.27	.11	3.25
30	15.55	.24	-.05	-1.62
31	16.00	.35	.05	1.10
32	16.24	.31	-.12	-3.13
33	9.70	.68	-.16	-1.81
34	8.27	.48	.15	2.38

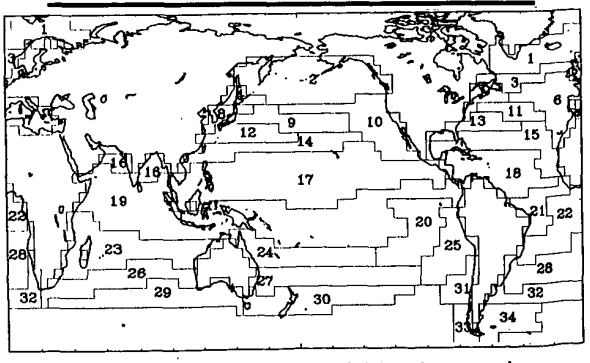


Figure 1. World map of COADS SST cluster regions.

3. RESULTS

In general, as expected, the largest regions are found in the tropics. Principal component analysis using the 34 regional time series of 160 seasonal values (SST anomalies) as input indicates that the first four to five significant eigenvector patterns account for 50-55% of the regional seasonal SST variance. Maps illustrating regions with significant loadings of $r > |0.3|$ from the first two eigenvectors are shown in Figure 2. The temporal coefficients of the first eigenvector closely follows the variations in global mean SST anomalies during this time (r = .79) as shown in Figure 3.

The spatial patterns of the first two eigenvectors illustrated here, are similar to the second and third eigenvectors of SST anomaly calculated by Parker and Folland (1991) for the period since 1901 (their first eigenvector representing the long-term trend in the (adjusted) SST record).

For comparison purposes, the variance spectrum of each regional anomaly time series was calculated. Most of the significant spectral peaks occur near the biennial time scale (2.1-2.2 yr) and in the ENSO frequency band (~3-6 yr).

Acknowledgments

This work is part of an effort by several NOAA units, with funding from NOAA's Climate and Global Change Program, to develop a global climate perspectives system for climate diagnostics research and climate change detection efforts.

References Cited

Hartigan, J. A., 1975: *Clustering Algorithms*. J. Wiley & Sons, New York, 351 pp.

Parker, D.E. and C. K. Folland, 1991: Worldwide surface temperature trends. *In* M. E. Schelesinger (ed.) *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations*. Amsterdam, Elsevier, pp. 173-193.

Skaggs, R. H., 1975: Drought in the United States, 1931-40. *Annals of the Assoc. of Amer. Geogr.*, **65**, 391-402.

Stidd, C. K., 1967: The use of eigenvectors for climatic estimates. *J. Appl. Meteor.*, **6**, 255-264.

Woodruff, S.D., R.J. Slutz, R.L. Jenne, and P.M. Steurer, 1987: A comprehensive ocean-atmosphere data set. *Bull. Amer. Meteor. Soc.*, **68**, 1239-1250.

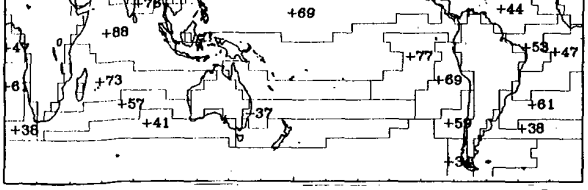


Figure 2a. Loadings from the first eigenvector of the PCA analysis. Only values exceeding |0.3| are shown.

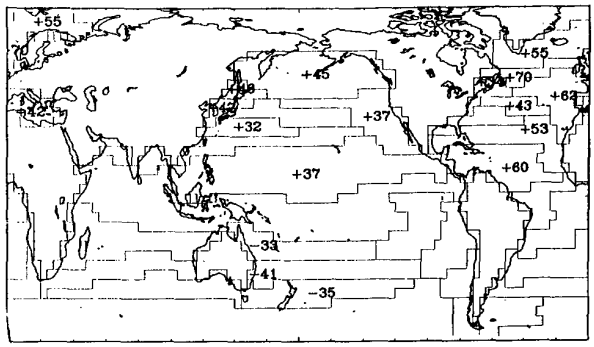
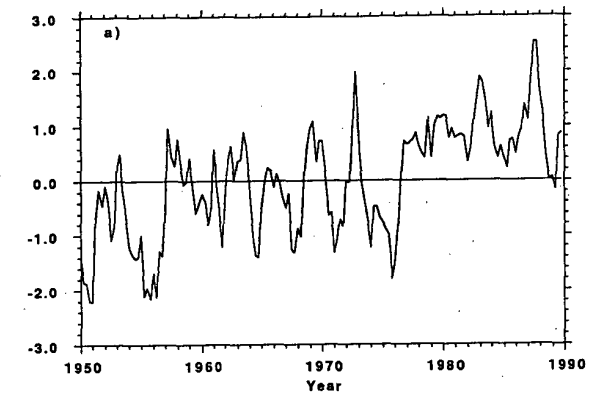


Figure 2b. Loadings from the second eigenvector of the PCA analysis. Only values exceeding |0.3| are shown.



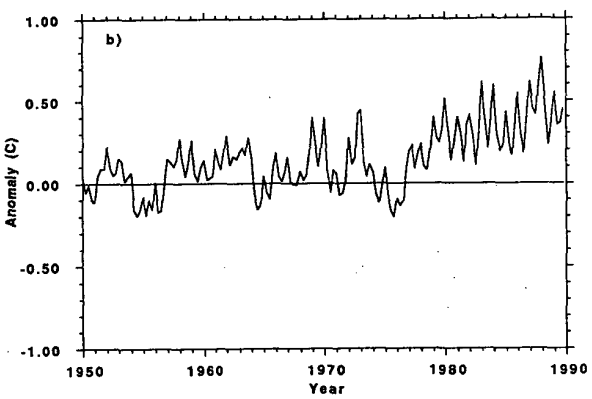


Figure 3. a) temporal coefficients from the first eigenvector and b) global SST seasonal anomalies.

CLIMATE VARIABILITY DURING THE DECADE OF THE 1980's AND ITS INFLUENCE ON THE 1961-1990 BASE PERIOD MEANS

Michael S. Halpert and Chester F. Ropelewski

W/NMC52
Climate Analysis Center/NMC/NWS/NOAA
Washington, D.C. 20233

1. INTRODUCTION

During the decade of the 1980's, global land temperatures were the highest in the historical record (Halpert and Ropelewski, 1991). Almost all land areas of the Northern Hemisphere had temperatures which averaged above normal for the decade (Fig. 1). In addition, surface temperatures for the Southern Hemisphere, except for central South America, were above normal. However, although most areas of the globe experienced above normal temperatures during the 1980's, the anomalies were inconsistent from season to season. Also, precipitation over some areas was well below the 1951-1980 base period means over some locations. These unusual surface anomalies will impart significant differences to the 1961-1990 base period means as compared to the 1951-1980 base period and will cause changes in the perception of anomalously warm and cold spells as well as wet and dry period.

2. CLIMATE VARIABILITY DURING THE 1980'S

Over the Northern Hemisphere, much of the decadal warmth during the 1980's occurred during the December - May period. The large differences in temperature anomalies between the first and second half of the year is shown in Fig. 2. The December - May period had average temperatures greater than 1.5°C above normal covering a large area over Alaska and western Canada. Another large area with positive anomalies exceeding 1.0°C was found over Siberia. In contrast, anomalies over Alaska and western Canada were negative, albeit small, during the June - November period.

Precipitation during the decade was also highly variable. The most significant precipitation anomalies were over the African Sahel, where drought conditions occurred throughout most of the decade. Precipitation during the June - September period was above normal during only one year (1988), while very dry conditions occurred in many years, especially in the early part of the decade. Rainfall was also below normal across northern Australia and southeastern Asia. Precipitation was generally abundant across the United States, although, on average, the Southeast was drier than normal during the March - May season.

3. 1961-1990 BASE PERIOD MEANS

These highly variable patterns of temperature and precipitation will have fairly large influences on the calculation and appearances of anomalies in the coming years. Surface temperature and precipitation anomalies or percentiles are usually computed with respect to a climatological record based on the most recent complete three decades of data. This practice ensures that the anomalies are being computed with respect to contemporary climate conditions. Anomalies have been computed with respect to the 1951-1980 base period throughout the 1980s, but now should be based on the 1961-1990 time frame.

Regions in which changes in the base period will exert a significant influence on the surface temperatures and precipitation anomalies are identified in Figs. 3 and 4. These figures depict the difference in precipitation totals and mean surface temperature between the two base periods by season for the Northern Hemisphere. Since the two base periods include a common 20-year period (1961-1980), the analyses actually reflect mean differences between the 1950's and the 1980's.

Since the highest decadal mean surface temperatures in the historical record occurred during the 1980's at many locations in the Northern Hemisphere, temperature anomalies will be computed with respect to higher base period means than had been used

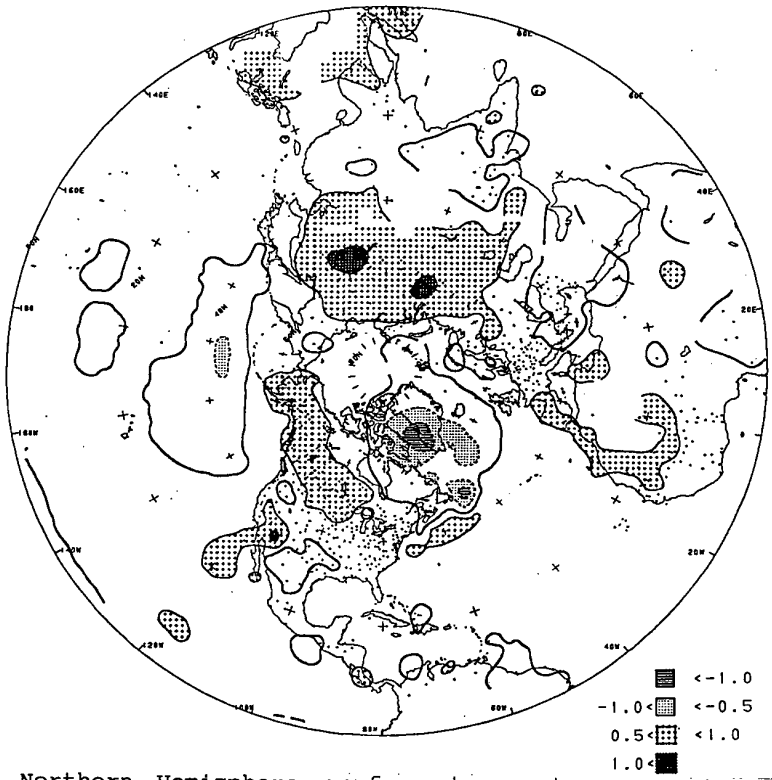


Figure 1. Northern Hemisphere surface temperature anomalies for the period 1981-1990. Analysis is based on station data over land and on sea surface temperature (SST) data over the water. Anomalies for station data are from the 1951-1980 base period, while SST anomalies are computed as departures from the COADS/ICE climatology (Reynolds 1988). Contour interval is 0.5°C, with negative anomalies dashed.

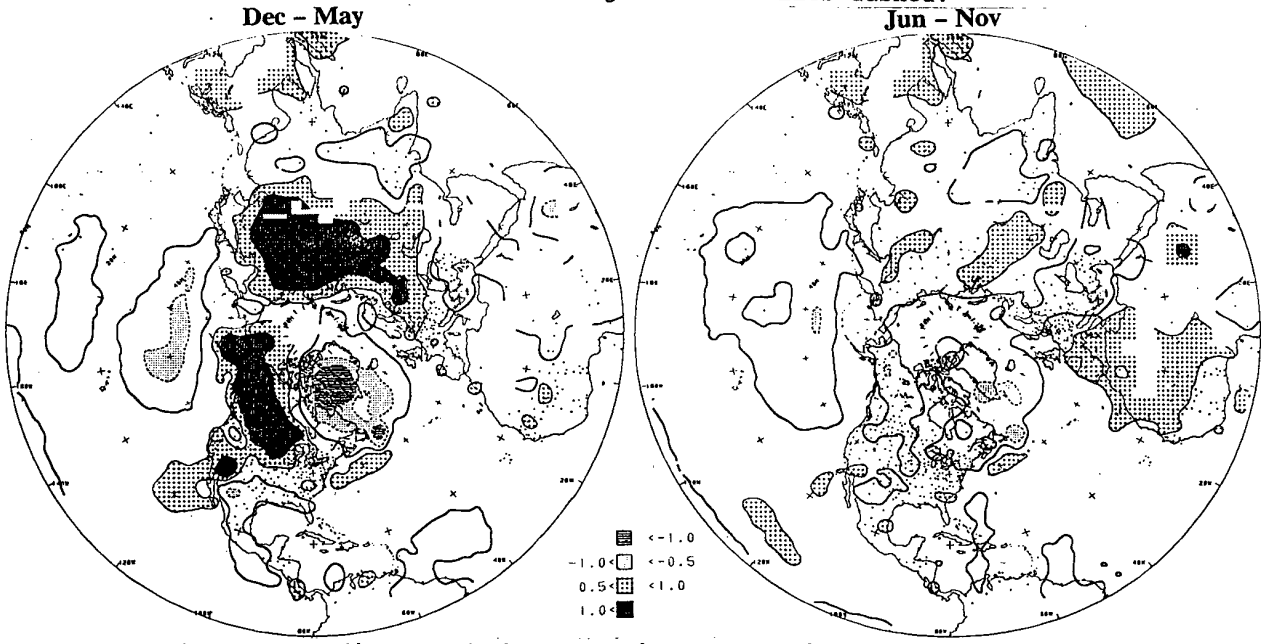


Figure 2. Northern Hemisphere December - May (left) and the June - November (right) surface temperature anomalies for period 1981-1990. Analysis is based on station data over land and on sea surface temperature (SST) data over the water. Anomalies for station data are from the 1951-1980 base period, while SST anomalies are computed as departures from the COADS/ICE climatology (Reynolds 1988). Contour interval is 0.5°C, with negative anomalies dashed.

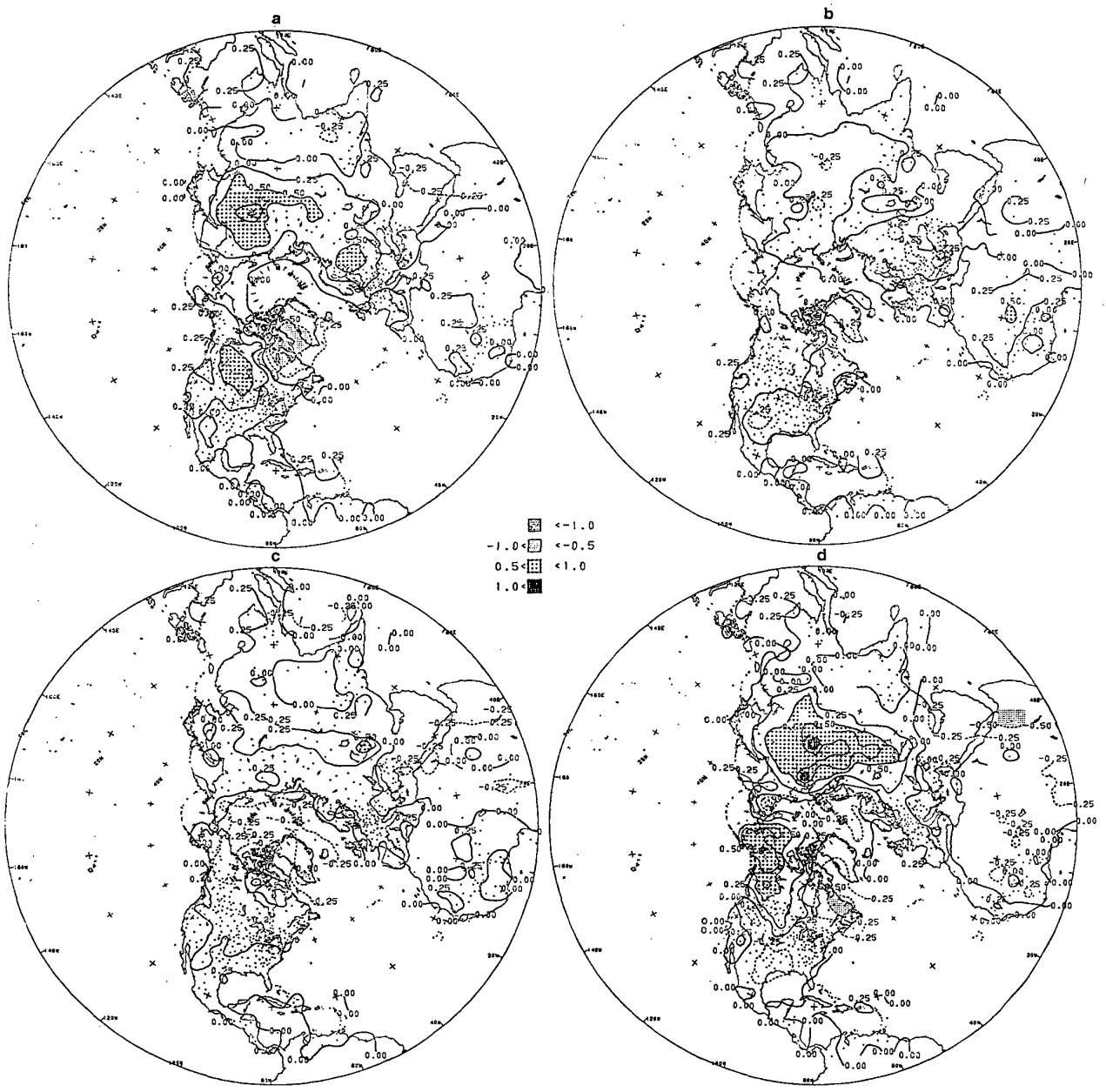


Figure 3. Northern Hemisphere temperature differences between the 1961-1990 base period mean and the 1951-1980 base period mean for a) MAM, b) JJA, c) SON, and d) DJF. Contour interval is 0.25°C. Station locations are denoted by small "+"; no analysis is done in areas with insufficient data.

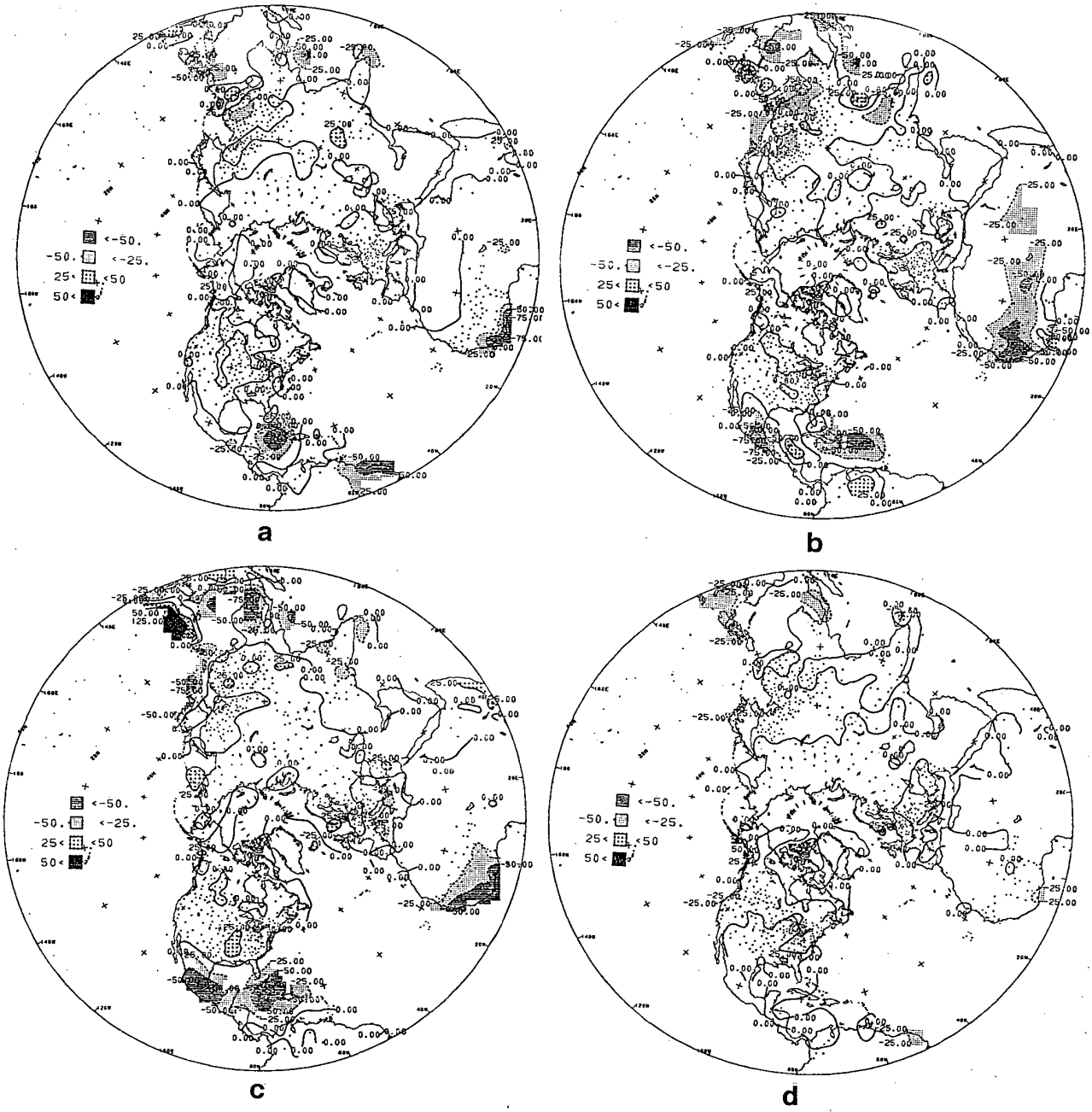


Figure 4. Northern Hemisphere precipitation differences between the 1961-1990 base period mean and the 1951-1980 base period mean for a) MAM, b) JJA, c) SON, and d) DJF. Contour interval is 25 mm. Station locations are denoted by small "+"; no analysis is done in areas with insufficient data.

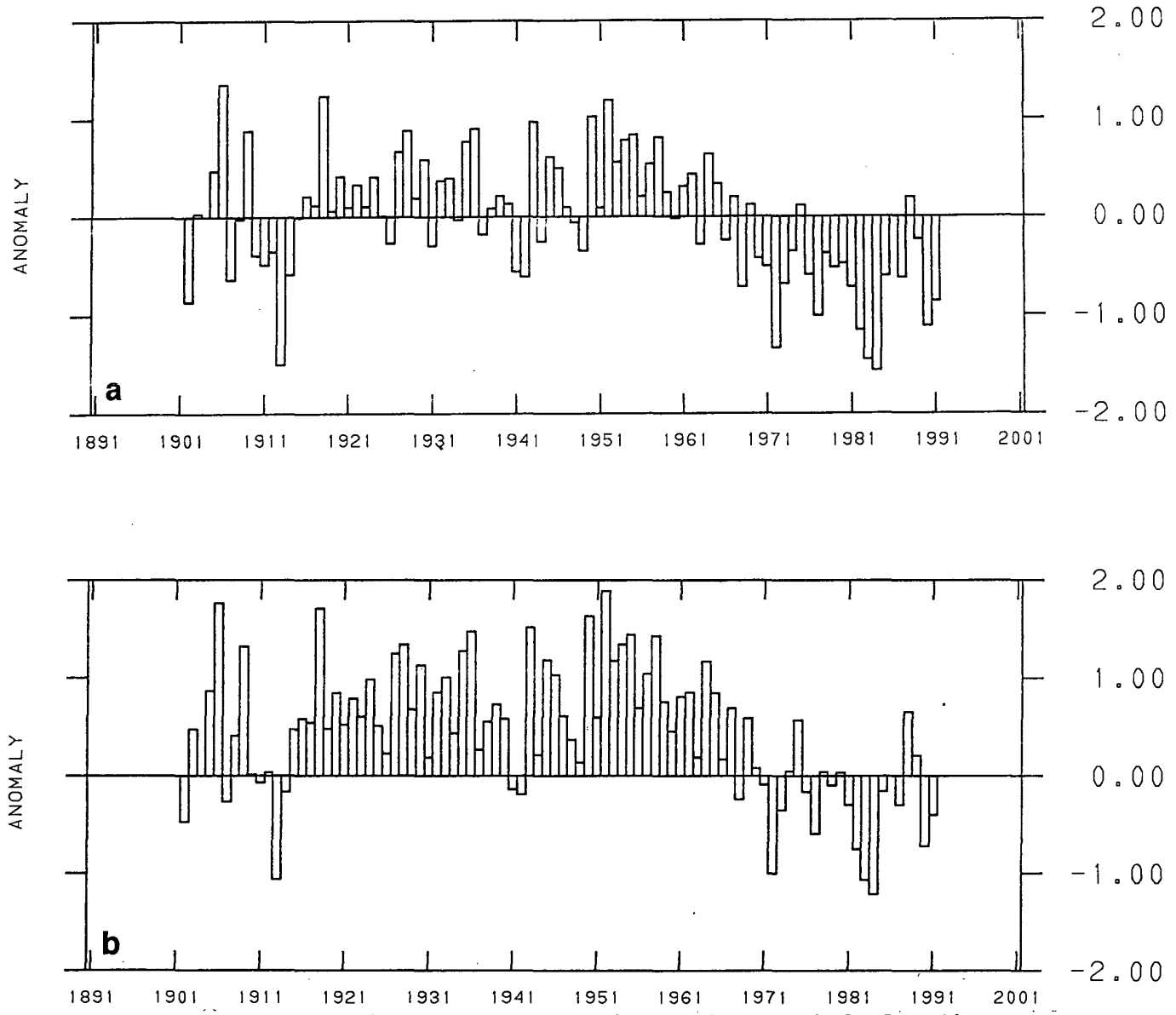


Figure 5. Precipitation index for the western Sahel for the June - August period normalized by the a) 1951-1980 base period and the b) 1961-1990 base period.

previously. The largest base period mean differences are found during the December - February and the March - May seasons (Fig. 3). During each of the periods, positive temperature differences greater than 0.5°C are found in both northern and central Asia and in northwestern North America. In contrast, negative differences are found over northwestern North America during the September - November season. Most stations in the Southern Hemisphere will have new base period means that are within 0.25°C of their old means.

Differences in the seasonal mean precipitation totals between the new and old base periods for the Northern Hemisphere are shown in Fig. 4. Precipitation means are similar for both base periods over most regions of the mid- and high latitudes for all seasons. In the tropics and sub-tropics, the base period means for December - February will also remain about the same, but the means exhibit large differences in the other seasons. The most significant changes are found over the African Sahel, where the new means will be as much as 100 mm lower than the old means during the JJA season (Fig. 4b). As a consequence, the current prolonged drought conditions appear much less severe in a time series of normalized precipitation anomalies computed from the 1961-1990 base period (Fig. 5a) than in a time series computed from the 1951-1980 base period (Fig. 5b). Precipitation means in the Caribbean are substantially lower for the 1961-1990 base period during the March - May, the June - August, and the September - November seasons. This reduction may reflect less frequent tropical storm activity over the region during the 1980's compared to the 1950's.

Largest differences in the precipitation base period means for the Southern Hemisphere occur during the December - February season. Large positive differences (indicating greater means for the 1961-1990 base period) are found over Madagascar and central South America. In contrast, the new means for northeastern Australia will be significantly lower than the old means. This difference is partially the result of the very dry conditions that affected this region in association with the ENSO episodes during 1982/83 and 1986/87 (Ropelewski and Halpert 1987). During the other seasons, the differences over Australia and southern Africa are generally less than 25 mm. Over central South America, precipitation differences are larger during MAM, with the base period means for 1961-1990 being more than 50 mm larger than for the 1951-1980 period.

4. SUMMARY

Unusual warmth and dryness over large areas of the Northern Hemisphere during the decade of the 1980's will result in significant changes to the new base period means. Temperature anomalies over many regions will often be computed relative to higher base period means, resulting in smaller positive anomalies or negative anomalies as compared to the previous base period. However, these differences are regional and seasonal in nature. Also, the appearances of droughts will be greatly modified in some regions of the globe, such as in the African Sahel. The lower base period means over this region will result in decreasing the appearance of the severity of drought conditions during the past 20 years. Changes in the temperature base period means are generally small in the Southern Hemisphere and therefore have less impact on the appearances of hot or cold spells.

REFERENCES

Halpert, M. S. and C. F. Ropelewski [Eds.], 1991, Climate Assessment: A Decadal Review, 1981-1990. [U. S. Govt. Printing Office: 1991 - 281-557/40426, 109 pp.

Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. J. Climate, 1, 75-86.

Ropelewski, C. F. and M. S. Halpert, 1987: Global and Regional Scale Precipitation Patterns Associated with the El Niño/Southern Oscillation. Mon. Wea. Rev., 115, 1606-1626.

TECHNIQUES FOR DETECTING AND ADJUSTING FOR ARTIFICIAL DISCONTINUITIES IN CLIMATOLOGICAL TIME SERIES: A REVIEW

David R. Easterling and Thomas C. Peterson

Global Climate Laboratory
National Climatic Data Center
Asheville, NC 28801

1. Introduction.

Many of the research problems facing climatologists today, such as global climate change, require long homogeneous time series: time series where variations are caused only by variations of weather and climate (Conrad and Pollock 1950). Using data that have not been adequately adjusted for inhomogeneities can lead to significant errors since the climate change signal can easily be obscured or erroneously enhanced by discontinuities in the data. For example, Hansen and Lebedeff's (1988) analysis of global surface temperature trends indicate considerable warming in the 1980s around St. Helena Island (16°S, 6°W). However, much of this "warming" was actually due to an inhomogeneity in St. Helena Island time series caused by the station moving to a considerably lower elevation (Michaels 1992).

Numerous tests have been devised to determine the homogeneity of a time series relative to time series of adjacent stations. All of these tests are based on the assumption that inhomogeneities will become apparent when examining a difference (or ratio) series between the time series in question and those from surrounding stations. However, care must be taken since entire station networks can undergo instrument changes simultaneously (Heino, 1991) causing the reference series and the station time series to have identical discontinuities.

Inhomogeneous time series arise either due to abrupt discontinuities or to effects that gradually increase over time. Abrupt discontinuities occur due to instrument and exposure changes, station moves, and changes in calculation procedures for time-averaged values. Inhomogeneities that occur as a gradually increasing effect generally develop due to a change in the environment surrounding a station, or due to a gradual change in either the instrument calibration or the physical characteristics of the instrument apparatus (e.g. yellowing of a white plastic instrument shelter, Quayle, *et al.* 1991).

Techniques for detecting and correcting for inhomogeneities fall into 3 general categories. The first category relies on metadata and parallel observations. For example, if the type of thermometer at a station is changed, and observations from both the old and new thermometers are taken for a specified period, an adjustment factor can be determined and applied to create a homogeneous time series. The second type depends on metadata detailing station moves, instrument changes, etc., and uses this information to determine if the change caused a statistically significant discontinuity. The third type is statistical (Potter 1981, Alexandersson 1986) or graphical (Jones, *et al.* 1986) in nature and relies on no metadata to identify discontinuities and adjustment factors.

For this paper we evaluated only techniques that do not rely on metadata and are easily automated. We attempted to determine how well the various techniques can identify the year (we use the term "year" to refer to a place in a time series) of a discontinuity and the magnitude of the adjustment factor. All of the techniques we evaluated use a reference time series that is a composite of surrounding stations to identify discontinuities at a candidate station. While the quality of reference time series can be critically important, for this paper we did not evaluate the various ways a reference time series can be created.

2. Techniques.

Each of the techniques we examined start with a time series for the candidate station and a reference series developed from nearby highly correlated stations. In most instances a difference series (or for precipitation a ratio series) was created. For a given technique the test statistic was calculated for each year of the candidate time series starting with year 5 and ending with year N-5. Extremes in the test statistic values were examined for statistically significant values (95% confidence level) indicating a discontinuity.

Alexandersson's technique: Alexandersson (1986) developed a test starting from a standardized difference (or ratio) series, with the test statistic as follows:

$$T_o = \max_{5 \le v \le n-5} \{T_v\} = \max_{5 \le v \le n-5} [v\bar{z}_1^2 + (n-v)\bar{z}_2^2] .$$

where T_o is the test statistic, $\bar{z}_1$ is the arithmetic mean of the difference series from year 1 to v, and $\bar{z}_2$ is the arithmetic mean of the difference series from year v+1 to n.

Potter's technique: Potter (1981) developed a test that starts by standardizing both the candidate (Y) and reference (X) time series and calculating the following:

$$X_i = 1/i \sum_{j=1}^{i} x_j, \; Y_i = 1/i \sum_{j=1}^{i} y_j$$

$$\bar{X} = X_n, \; \bar{Y} = Y_n$$

$$S_x = \sum_{j=1}^{n} (x_j - \bar{X})^2, \; S_y = \sum_{j=1}^{n} (y_j - \bar{Y})^2,$$

$$S_{xy} = \sum_{j=1}^{n} (x_j - \bar{X})(y_j - \bar{Y}),$$

$$F_i = S_x - (X_i - \bar{X})^2 ni/(n-i), \; i<n,$$

$$D_i = [S_x(\bar{Y} - Y_i) - S_{xy}(\bar{X} - X_i)]\, n/[(n-i) F_i]$$

$$T_i = [i(n-i) D_i^2 F_i] / (S_x S_y - S_{xy}^2) .$$

with the test statistic given by:

$$T_o = \max_{5 \le i \le n-5} \{T_i\}.$$

the value of i for T_o is the maximum likelihood estimate of i_o and D_i is the maximum likelihood estimate of the offset.

Regression: We developed a linear regression approach (Reg Y-hat in Table I) that is applied to the difference series (candidate minus reference series). For a given year two regression equations are found with time as the predictor and the difference as the predictand. One equation is found for the period before the year in question and one for the period after the year. Using the two regression equations, the regression estimates of the predictand are determined for the year and differenced. The 95% confidence intervals are also calculated for the year and compared to the difference in the two regression estimates. If the difference is bigger than the confidence intervals this year is identified as a potential discontinuity. After iterating through the entire time series, the largest potential discontinuity is identified, with the difference being the offset. A variant on this approach (Reg avg Y-hat in Table I) is to calculate the average regression estimate for the periods before and after the year in question.

Double-Mass analysis: Usually considered a graphical technique, double-mass analysis (Kohler 1949) is a plot of the cumulative sum of the station time series versus the cumulative sum of the reference time series. Discontinuities are indicated by a change in the slope.

CUSUM: The cumulative sum technique (van Dobben de Bruyn 1968) is a graphical technique that plots the cumulative sum of the difference series (station minus reference) versus time and looks for change points in the slope of the line. Rhoades and Salinger (1991) successfully used this technique to identify inhomogeneities in time series. Our automated versions of both double-mass and CUSUM looks for the largest change in slope.

Two-Phase Regression: Solow (1987) described a technique for detecting a change in the trend of a time series by looking for the change point in a two-phase regression of (for our application) the difference series versus time.

Other techniques: Several other techniques were conceived and examined (e.g. cluster analyses, and variations on regression) but had such poor results that there is no need to describe them here. The lesson we learned from these failures is that the signal deduced by examining very detailed aspects of the data can easily be obscured by noise and therefore "fancier" tests may reveal less. In addition, subjective graphical techniques such as that used by Jones, *et al.* (1986) were not evaluated.

3. Methods.

There are a number of approaches that can be followed to evaluate the ability of a technique to identify and adjust for discontinuities. We chose to use two approaches: to create simulated temperature time series with imposed discontinuities on them, and to use temperature time series for stations with potential discontinuities documented.

For the simulated data we created 1000 candidate time series and 1000 reference series using the following procedure for each candidate-reference time series pair. Two Z score series of 100 time elements (years) taken

from a random number generator were forced to have a serial correlation of 0.1 using an AR(1) process. The two Z score series were then cross correlated at 0.75-0.80 by multiplying the value of the first series for a given year by 1.5 and adding it to the value for that year in the second series. After cross correlation, they were re-standardized and formed into simulated annual temperature series, one with a mean of 23 and one with a mean of 25, both with standard deviations of 1.0 (annual temperature time series we examined from the U.S. Historical Climatology Network (U.S. HCN) typically had standard deviations of 0.75-1.0° C, and were approximately normally distributed).

An artificial inhomogeneity was then imposed at year 65 by adding a specified offset value to each annual value from year 65 through 100. Four different inhomogeneities were tested, ranging from offsets of one-half (0.5° C) to twice (2.0° C) the standard deviation. We ran each technique on the 1000 time series to evaluate how often it could find the correct year of the inhomogeneity and how well it could determine the magnitude of the offset. The techniques with the best results were then tested on artificial time series with multiple inhomogeneities.

Real time series of climate data that have inhomogeneities are easy to locate, but they present the difficulty of *a priori* knowing what the homogeneous time series should be to allow proper evaluation of the different techniques. To mitigate this problem, we chose four candidate stations that had 0 to 4 station moves (here a station move refers to some change, such as a move or instrument change, that could cause a discontinuity) documented in their station history files along with six moderate to highly correlated companion stations for each candidate from the U.S. HCN (Karl, *et al.* 1990). We selected the four candidate stations to reflect very different climate regimes: Red River, New Mexico; Eastman, Georgia; Elmira, New York; and Grapeview, Washington. The U.S. HCN provides data that are adjusted for known discontinuities using extensive station history files. We used the adjusted data for the companion stations to develop a reference time series for each of the four stations. The candidate series were the unadjusted data for each of the four stations. The dates of the known moves were used to evaluate the results.

4. Results.

Testing the techniques on 1000 simulated time series with one discontinuity revealed that most of the techniques worked well in evaluating discontinuities of twice the standard deviation (see Table 1). The only one that did not was the two-phase regression which apparently is much better suited for trend analysis rather than locating a marked discontinuity. The real test of the techniques came at smaller discontinuities. With a single discontinuity of half the standard deviation, Alexandersson, Potter, and the t-test worked considerably better than the other tests. However, Monte Carlo simulations showed that the t-test's critical values drastically increased near the ends of the time series making it less appropriate for examining the shorter intervals needed to look for multiple discontinuities. Therefore, we retained only the Alexandersson and Potter techniques to look at multiple discontinuities.

For evaluating multiple discontinuities, the Alexandersson and Potter techniques were designed to look for the most significant discontinuity, determine the offset value, apply the offset, and iterated again until no discontinuities were identified at the 95% confidence level. The offset value for both tests were calculated by taking the difference between the average of the difference series before the discontinuity and the average of the difference after the discontinuity. The offset was then added to the candidate time series from year 1 through the year of the discontinuity, then for the Alexandersson technique the difference series was re-formed, and re-standardized. The final adjusted time series were then correlated with the initial simulated time series prior to the imposition of the discontinuities. The results of this analysis are shown in Table II. For two discontinuities, Alexandersson's and Potter's techniques

Table I. Percentage of simulated time series with the date of the inhomogeneity identified correctly to within 1 year for the various techniques, by the magnitude of the discontinuity.

Discontinuity	0.5	1.0	1.5	2.0
Alexandersson	50.8	87.6	97.2	99.8
Potter	47.3	84.1	96.3	99.4
t-test	45.7	86.1	95.9	99.0
Reg Y-hat	24.7	75.9	94.5	98.9
Reg avg Y-hat	31.2	74.4	92.6	97.6
Double-mass	30.1	73.8	92.3	98.0
CUSUM	30.7	73.0	92.0	97.6
Two-Phase Reg	0.0	0.0	0.0	0.0

worked equally well. However, for three discontinuities, Alexandersson's technique produced somewhat better results.

Applying these two techniques to real data produced some interesting results (see Table III). When the correlation between the candidate and reference series was very high, both techniques found the same discontinuities in the time series, approximately corresponding to dates of known station moves. For instance, the 1927 move at Grapeview, WA, that was identified as a discontinuity at 1929 in the maximum and 1935 in the minimum, was actually a one foot increase in the height of the instrument shelter, and most likely a change in the equipment. Furthermore, since the technique used for calculating the offsets were the same, the magnitude of the adjustments were identical when the year was the same. The main difference between the two techniques occurred in the maximum temperature for Elmira, NY. Alexandersson's technique found three discontinuities, yet Potter's technique indicated no discontinuity even down to the 90% confidence level (normally we used the 95% confidence level in our analyses). Another point is that these techniques can also identify discontinuities when the metadata do not indicate a station move. For example, the metadata for Red River, NM showed no station moves during the time period we examined, yet both techniques found discontinuities for the same years. This may be a problem with the development of the reference series (e.g candidate and reference are not strongly correlated), or there may be genuine, undocumented discontinuities.

5. Summary and Conclusions.

Many techniques exist to detect inhomogeneities in time series data. We have attempted to evaluate those techniques most commonly cited in the literature, as well as a few we attempted to develop. Some of the techniques we evaluated, though quite effective for their original intent, are clearly inappropriate for identifying discontinuities. The techniques developed specifically for identifying discontinuities all appear to work well when discontinuities are large. However, the reliability of many of these techniques decreased markedly for discontinuities smaller than the standard deviation of the data. Also, the reliability of these techniques decreased

Table II. Multiple discontinuities. The correlation of 1000 adjusted time series with the original time series before the discontinuities were imposed.

Discontinuities Magnitude @ year	Alexandersson	Potter
+1 @ 25, +1 @ 75	0.96	0.96
+1 @ 25, -1 @ 75	0.98	0.98
-2 @ 25, +3 @ 75	0.93	0.93
+1 @ 30, +1 @ 50, +1 @ 80	0.92	0.91
+1 @ 30, -1 @ 50, +1 @ 80	0.94	0.92
+2 @ 25, -2 @ 50, +3 @ 75	0.89	0.84

Table III. Real Data. The date and the magnitude of the adjustment that Alexandersson's and Potter's techniques would make on maximum and minimum temperature time series from four US stations. Also shown are the dates of known station moves and the correlation between the time series and a reference series created from nearby highly (or as high as possible) correlated stations.

	Year known moves	Alexandersson		Potter		Correlation Cand. w/ Ref.	
U.S. HCN Station		Max T Yr (offset)	Min T Yr (offset)	Max T Yr (offset)	Min T Yr (offset)	Max T	Min T
Eastman, GA	1951, 1955, 1956, 1979	None	1956 (0.29) 1979 (-0.41)	None	1956 (0.29) 1979 (-0.41)	0.66	0.91
Elmira, NY	1949, 1961, 1963, 1977	1946 (-0.72) 1960 (0.31) 1967 (0.40)	1946 (0.85)	None	1946 (0.85)	0.85	0.83
Grapeview, WA	1927, 1971	1929 (0.85) 1970 (0.42)	1935 (-0.43) 1967 (0.37)	1929 (0.85) 1970 (0.42)	1935 (-0.43) 1967 (0.37)	0.76	0.82
Red River, NM	None	1932 (0.61) 1972 (-0.49)	1943 (0.49) 1966 (-0.55)	1932 (0.61) 1972 (-0.49)	1943 (0.49) 1966 (-0.55)	0.66	0.78

as the correlation between the candidate time series and the reference time series decreased; therefore effort spent in creating a highly correlated reference series is well worthwhile. We found Alexandersson's (1986) and Potter's (1981) likelihood ratio techniques the most sensitive to small discontinuities. These two techniques were more rigorously tested with Alexandersson's technique proving to be most sensitive.

These techniques generally are intended to focus on sharp discontinuities in the time series. What if the inhomogeneity is a slow trend instead of a sharp break? For instance, what if the candidate series has slow, but steady urban warming while the reference series does not? We took our 1000 simulated time series and applied an increase of 0.005° C yr^{-1} (0.5° C per 100 years) increase to each candidate series in place of any discontinuities. Our evaluation of this problem revealed that both Alexandersson's and Potter's techniques will respond to the trend, but will identify the trend as distinct discontinuities approximately in the middle of the time series. Therefore, if trends are present in the data, they must be reflected approximately equally in both the candidate and reference time series, or spurious results may occur.

Acknowledgements. This work was supported by U.S. Department of Energy Interagency Agreement DE-A105-90ER60952.

References

Alexandersson, H., 1986: A homogeneity test applied to precipitation data., *J. Climatol.*, **6**, pp. 661-675

Conrad, V., and C. Pollock, 1950: **Methods in Climatology**, Harvard University Press, Cambridge, Mass. 459 pp.

Hansen, J., and S. Lebedeff, 1988: Global surface air temperatures: update through 1987. *Geophy. Res. Letters*, **15**, 323-326.

Heino, Raino, 1991: Personal communication.

Jones, P., S. Raper, R. Bradley, H. Diaz, P. Kelly, and T. Wigley, 1986: Northern Hemisphere surface air temperature variations: 1851-1984. *J. Climate Appl. Meteor.*, **25**, pp. 161-179

Karl, T.R., C.N. Williams Jr., F.T. Quinlin, and T.A. Boden,1990: **United States Historical Climatology Network (HCN) Serial Temperature and Precipitation Data,** Environment Sciences Division, Publication No. 3404, Carbon Dioxide Information and Analysis Center, Oak Ridge National Laboratory, 83 pp.

Kohler, M.A., 1949: Double-mass analysis for testing the consistency of records and for making adjustments. *Bull. Amer. Meteor. Soc.*, **30**, pp. 188-189

Michaels, Patrick, 1992: Personal Communication.

Potter, K.W., 1981: Illustration of a new test for detecting a shift in mean in precipitation series. *Mon. Wea. Rev.*, **109**, pp. 2040-2045

Quayle, R.G., D.R. Easterling, T.R. Karl, and P.Y. Hughes, 1991: Effects of recent thermometer changes on the cooperative network. *Bull. Amer. Meteor. Soc.*, **72**, pp. 1718-1724

Rhoades, D.A. and M.J. Salinger, 1991: Methods for adjusting temperature and rainfall records for site changes. DSIR Physical Sciences DSIRPS-C-41, Wellington, New Zealand.

Solow, Andrew R., 1987: Testing for Climate Change: An Application of the Two-Phase Regression Model. *J. Climate Appl. Meteor.*, **26**, pp. 1401-1405.

van Dobben de Bruyn, C.S., 1968: *Cumulative Sum Tests: Theory and Practice*, Griffin's Statistical Monographs and Courses, No. 24, London: Griffin.

J2.5

UNCERTAINTIES IN SEA SURFACE TEMPERATURE ANALYSIS DUE TO CHANGES IN THE SAMPLING NETWORK

D. Marsico[1], R.W. Reynolds[1], and C.F. Ropelewski[2]

[1]Coupled Model Project/NMC/NWS/NOAA

[2]Climate Analysis Center/NMC/NWS/NOAA,
Washington DC 20233

1. Introduction

Historical analyses of sea surface temperature (SST) fields depend on the data coverage provided by ships and, to a lesser extent, buoys. This data coverage has not been constant over time and the number of observations available for analysis has increased dramatically since the mid-1950's. Furthermore, in the past hundred years, there have been large and potentially very significant shifts in the data distribution over the oceans due to changes in commercial shipping lanes, conversion from wind powered to steam driven ships, and the opening of the Panama and Suez canals. In this study, we provide one set of preliminary SST uncertainty estimates that can be ascribed to under-sampling of the SST fields. We also examine the temporal changes in these uncertainties as a function of changes in the sampling array over the past 110 years using data from the Comprehensive Ocean Atmosphere Data Set (COADS) (Woodruff et al., 1987).

2. Analysis

Uncertainties due to temporal changes in the array

In this initial analysis, SST for three decades representing time periods with significantly different ship track patterns (the 1880's, 1920's and 1970's) are examined for a winter and summer month (February and August). The data are extracted from the COADS Monthly Summary Trimmed Groups.

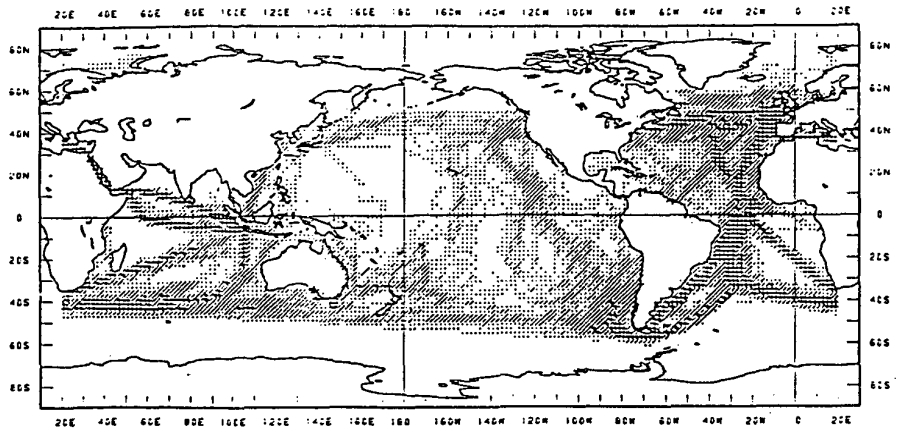
The data distributions for August in each decade are illustrated in Fig.1. During the 1880's the most heavily sampled regions were in the North Atlantic, in the South Atlantic, along the ship tracks rounding Cape Horn and Cape of Good Hope and in the northern Indian Ocean. During this decade, the Pacific was, in general, much less well sampled, with almost no observations in the north central and western regions. The southeastern Pacific, however, is an exception, being better sampled in the 1880's than in contemporary analysis.

During the 1920's there was an important increase in the number of ship observations in the North Atlantic and dramatic shift in the sampling patterns. In particular, the opening of the Panama Canal is reflected in the significant changes in ships tracks in the Pacific east of the dateline. By the 1970's most of the oceans are relatively well sampled with the exceptions of areas south of 40° S, the eastern Pacific, parts of the central Pacific, and the western South Atlantic.

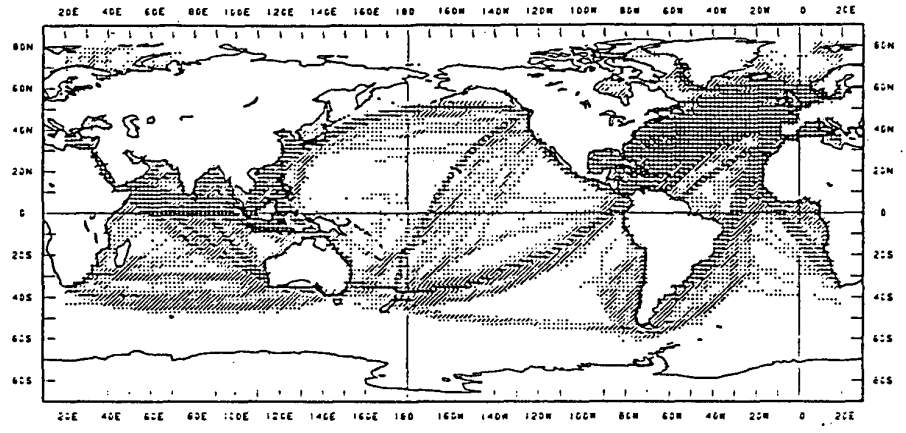
To effectively convert data with these uneven distributions into SST fields, or normalized SST fields, we specify a 1°C SST anomaly for all observations. We then analyze the resulting fields using and optimum interpolation (OI), Lorenc (1981), with a first guess of 0°C for the anomaly. In this way we separate the effect of natural spatial variations in SST from those differences due solely to changes the number and distribution of observations in the sampling arrays. The resulting analyzed fields e.g., Fig. 2 for August, closely reflect patterns represented by the respective sampling distributions but have the advantage of providing a means to quantify the effects of sampling. The February analyzed fields, Fig. 3, show essentially the same pattern expect for high latitudes in the Southern Hemisphere. In a continuation of the work presented here we will analyze each month in the COADS records from January 1880 to December 1987.

The effects of sampling on basin-scale SST averages for 5 ocean basins are summarized (Fig. 4); North Pacific, Southern Pacific, North Atlantic, South Atlantic, Indian and a global mean (80°N to 80°S). A "perfect" analysis would result in basin-scale average anomalies of 1°C. Thus the analysis for the 1970's in the North Pacific is close to "perfect" with a mean anomaly of 0.93°C versus an anomaly of only 0.21°C if the North Pacific is sampled with the array available in the 1880's. The difference between these two computed anomalies (0.72°C) is an estimate of sampling

1880's



1920's



1970's

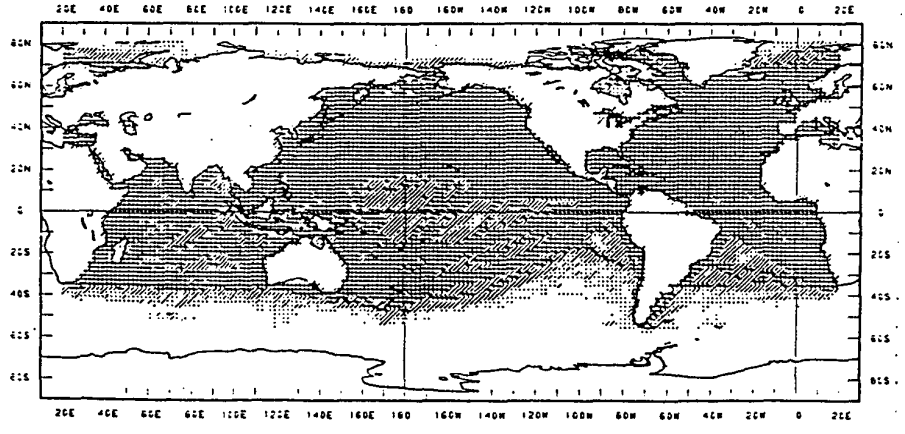
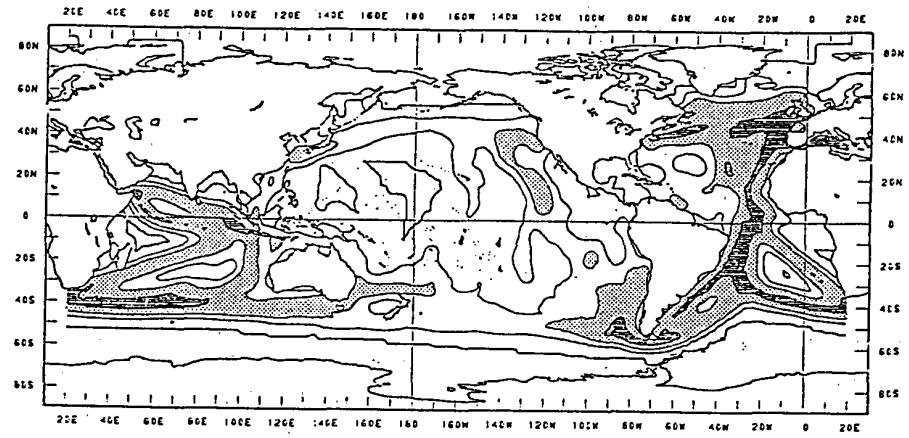
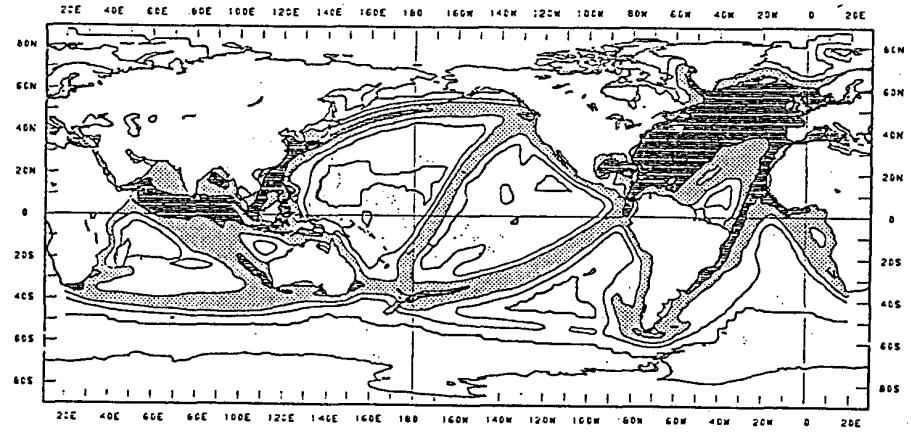


Figure 1: The number of years in the decade with observations in August. Dots represent 2° boxes with obs in 1 to 3 years. Slashes represent 4 to 7 years. Plus signs represent 8 to 10 years.

1880's



1920's



1970's

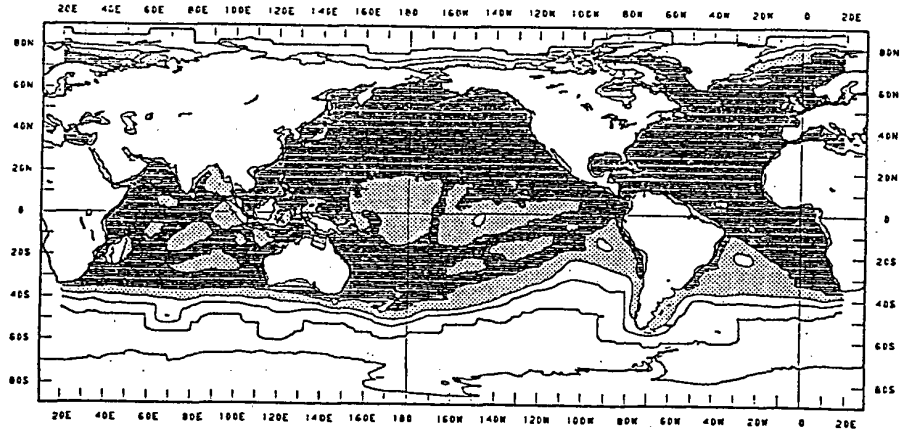
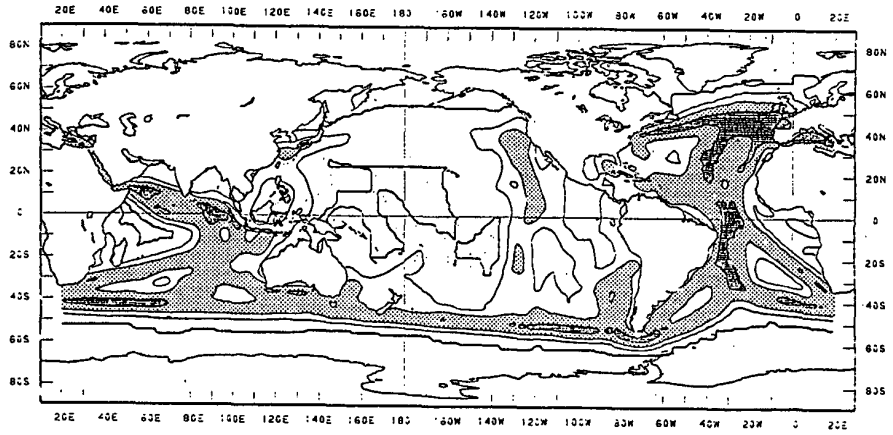
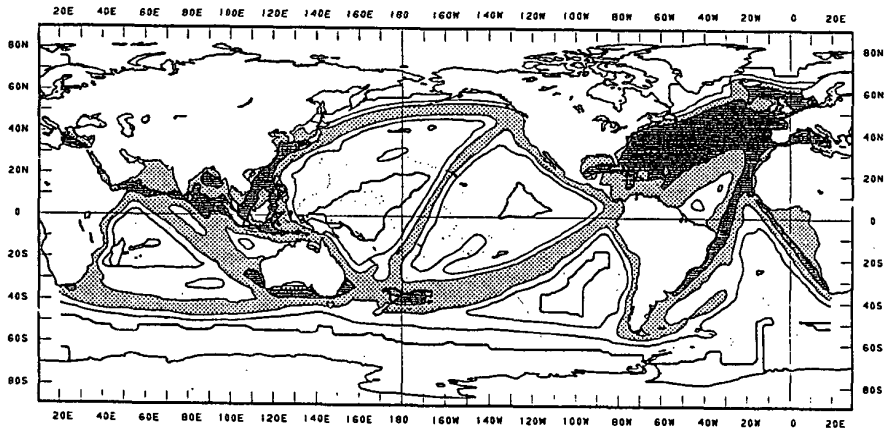


Figure 2: Decadal means of normalized SST for August. Heavy contours are at 0°C. Light contours are at 0.3°C, 0.6°C, and 0.9°C. Regions between 0.6°C and 0.9°C are lightly shaded. Regions greater than 0.9°C are heavily shaded.

1880's



1920's



1970's

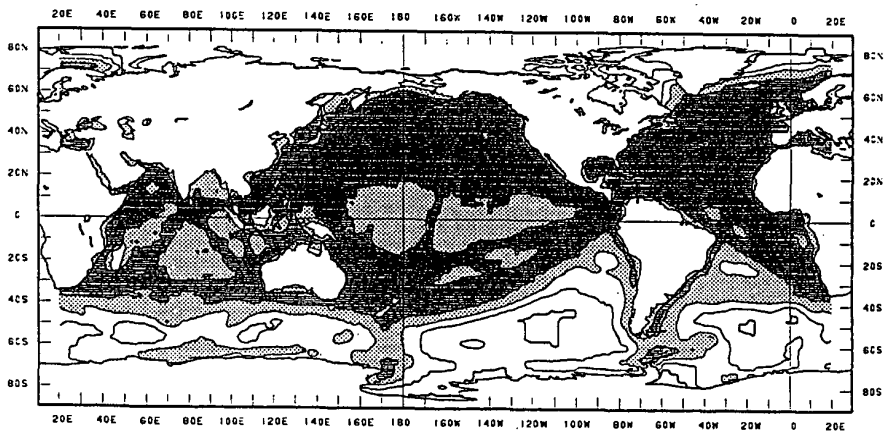


Figure 3: Decadal means of normalized SST for February. Heavy contours are at 0°C. Light contours are at 0.3°C, 0.6°C, and 0.9°C. Regions between 0.6°C and 0.9°C are lightly shaded. Regions greater than 0.9°C are heavily shaded.

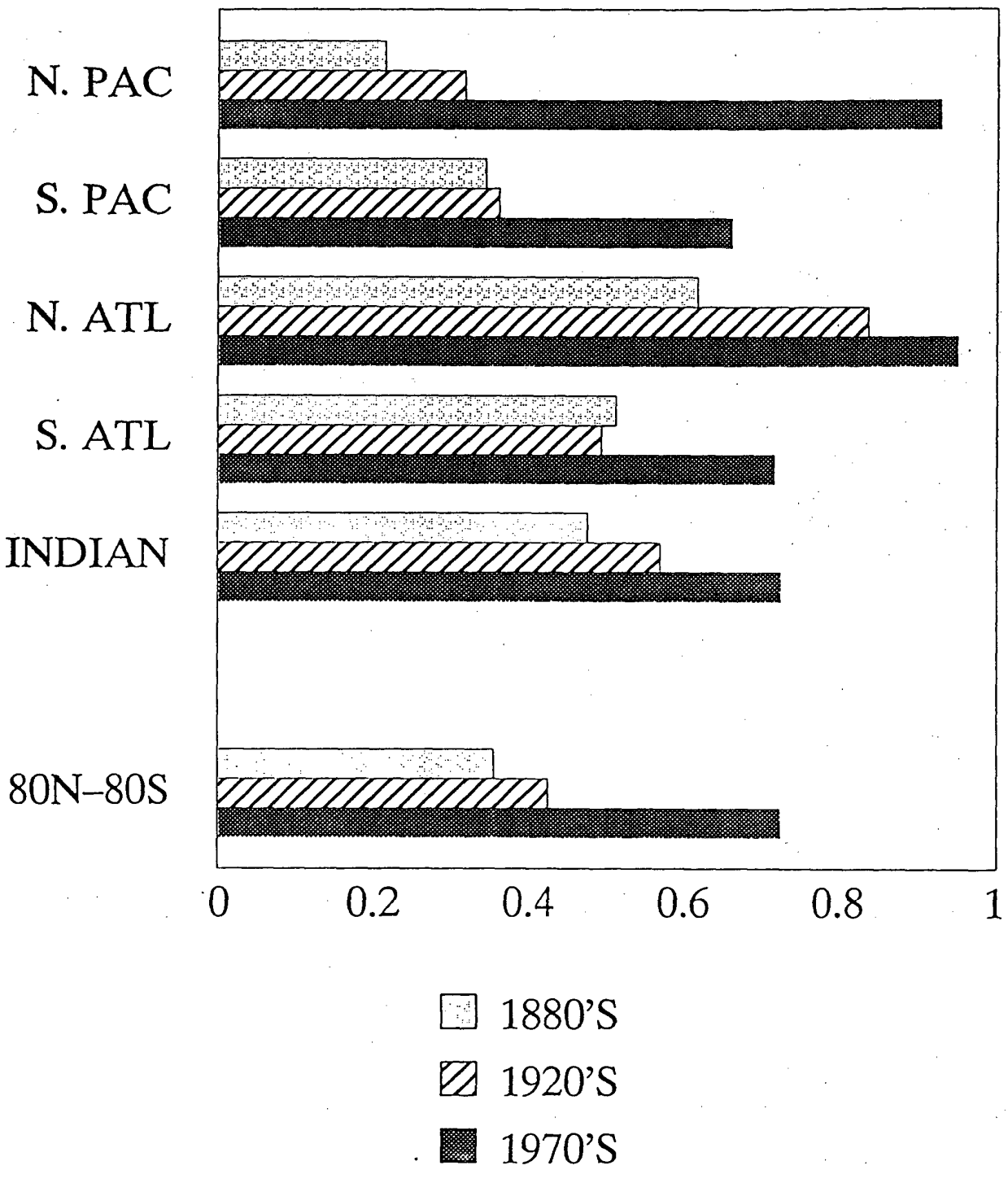


Figure 4: Decadal means of normalized SST (°C) for August averaged over each ocean basin. Ocean areas only. Each box is area-weighted.

uncertainty due to the changes in ship tracks and also to changes in the total number of observations. We see similar sampling differences for the other ocean basins but they are largest for the North Pacific which has gone from the poorest sampled ocean basin in the 1880's to being on a par with the North Atlantic in the 1970s. Although the mean anomaly estimates generally improve with time it's interesting to note, Fig. 4, that the South Atlantic was actually better sampled in the 1880's compared to the 1920's.

The differences between the 1970's value in each basin and a full field analysis value of 1.0°C can also be thought of as the sampling differences between in-situ only and "full field" estimate that are obtained from blended in-situ/satellite analyses.

While these results are preliminary they suggest that sampling uncertainties are non-negligible and need to be quantified in order to interpret any analysis of long SST time series. The method briefly outlined may provide one way to quantify uncertainties in such SST analyses.

REFERENCES

Lorenc, A.C., 1981: A global three-dimensional multivariate statistical interpolation scheme. Mon. Wea. Rev., **109**, 701-721.

Reynolds, R. W.,1988: A real-time global sea surface temperature analysis. J. of Clim., **1**, 75-86.

Ropelewski, C. F., D. C. Marsico, M. Chelliah, and T. Smith, 1992: What Can COADS Tell US? Potential and Limitations. Proc. of the 2nd International COADS Workshop, Jan 13-15, 1992, Boulder, CO.[Available from CRD/CMDL/ERL, Boulder CO 80303-3328]

Woodruff S. D., R.J. Slutz, R. L. Jenne, and P. M. Steurer, 1987: A comprehensive ocean-atmosphere data set. Bull. Amer. Meteor. Soc., **68**, 1239-1250.

AN EXAMINATION OF SPATIAL STATISTICAL TECHNIQUES FOR INTERPOLATION OF GRIDDED CLIMATE DATA

Timothy J. Brown

and

Jon K. Eischeid

CIRES, University of Colorado, Boulder, CO 80309

1. INTRODUCTION

Often in climate studies it is desirable to work with a gridded data set which has been derived from irregularly spaced point data. Generation of the gridded data may be accomplished by rather sophisticated techniques or simply by suitable averaging all of the stations within appropriate grid boxes. However, in the case of simple averaging (whether variable distance weights are used or not), this may leave gaps as stations are often not uniformly distributed in time and space. It would then be of interest to estimate or interpolate a value for the missing boxes and thus create a more spatially complete data set.

Here we present some initial work of estimating values for missing grid boxes in a precipitation and sea surface temperature (SST) data set. Results are compared using both optimal interpolation and kriging algorithms. The precipitation data is a subset of the Department of Energy comprehensive precipitation data set for global land areas (Eischeid et al., 1991). This subset, which covers most of the South American continent, is a 10x17 grid from 80°W to 35°W and 10°N to 54°S for the period 1951-1970. The SST data is a subset from the Comprehensive Ocean and Atmosphere Data Set (COADS) (Woodruff et al., 1987) encompassing the eastern equatorial Pacific region near the western South America coast from 106°W to 80°W and 0° to 26°S, which is a 13x13 grid, and spans the period 1950-89. In both data sets we examined results using only January and July monthly values.

2. OPTIMAL INTERPOLATION

Early uses of optimal interpolation (OI) in meteorology may be traced back to Gandin (1963). Since that time it has had wide usage in climatology and meteorology, including operational forecast models. Here we offer a brief overview of OI.

As with most objective analyses, OI is a grid point method which assigns weights to the observed difference values (observed minus first guess),

$$\hat{z} = z_f + \sum_{i=1}^{n} W_i (z_{oi} - z_{fi}) \ , \qquad (2.1)$$

where z_{oi} and z_{fi} are the observed and first guess values respectively, and z_f is the first guess value at the grid point to be interpolated. The weighting coefficients, W_i are determined in an objective manner from a statistical regression procedure such that the rms error of the analyzed difference values at each grid point is minimized over a particular spatial domain. The weights are dependent upon the spatial autocorrelations among the observation difference values, and typically modeled mathematically as a function of distance separating observations and grid points.

Figure 1 shows an example of a correlation-distance scatter diagram derived from the 20-year historical precipitation data for the South America grid. The data were fitted using a curve of the form $u(d) = A\ exp(-d^2 / R^2)$, where A was found to be .46 and R equal to 1300, when the distance is expressed in km. It should be noted that the precipitation spatial autocorrelations differed only slightly from month to month; thus, only one model was needed. The SST fields, however, revealed disparities in the spatial autocorrelations from one month to the next which were large enough to warrant modeling each month separately.

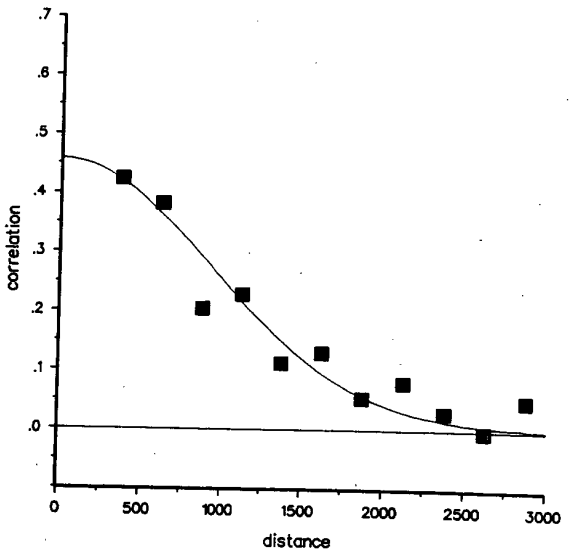


Figure 1. Correlation-distance scatter diagram with fitted curve for the South American precipitation data.

Once the data are prescribed to fit some mathematical function, the weights needed to solve equation (2.1) are given by the solution of a system of linear algebraic equations. In matrix form this can be written as

$$\mathbf{W_i C_{ir}} = \mathbf{G_r} \ , \qquad (2.2)$$

where $r = 1, 2, ..., N$ (number of observations) whose coefficients are the correlation coefficients that are easily

determined from observation separation distances (C, n x n matrix) and separation distances from each grid point to the observations (G, 1 x n vector). An analysis (interpolation) error is also determined at each grid point i by

$$\varepsilon_i^2 = 1 - \sum_{j=1}^{n} W_j u(d)_{ij} , \qquad (2.3)$$

where ε_i^2 is the error normalized by the variance.

3. KRIGING

Kriging is optimal spatial linear prediction based upon a minimum-mean-square-error method (Matheron, 1963). It has been widely used in earth sciences, but rarely in the atmospheric sciences. A very brief overview will be given here, while much more detailed discussions may be found elsewhere (e.g., Cressie, 1991).

As in the OI case, the extent of spatial continuity must be determined, but in kriging the semivariogram is used to find an optimal set of weights determined from observed sample locations which are then used in the estimation of unsampled locations. The semivariogram is expressed as

$$\gamma_h = \frac{1}{2n} \sum_{i=1}^{n-h} (z_i - z_{i+h})^2 , \qquad (3.1)$$

where z_i is the observed variable value at location i, and z_{i+h} is another observed value taken h distance away.

Figure 2 shows an example semivariogram for the 40-year mean of January SST, where variance is plotted against distance in degrees. The function is linear out to a range of 12 degrees (approximately 1300 km). The overall field variance is 1.18, but the linear function ceases around a variance of .95 (in the kriging nomenclature this is referred to as the "sill"). July SST as well as January and July precipitation also indicated linear functions, though of course the field variances were quite different.

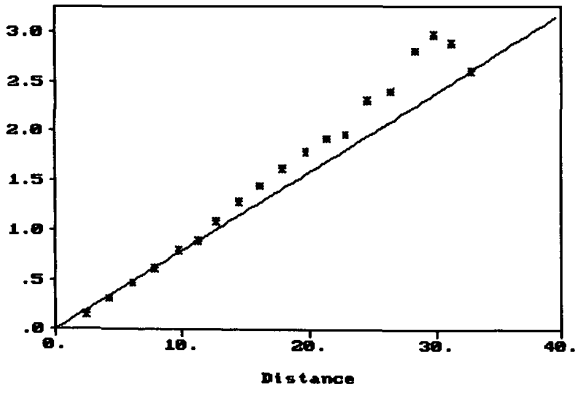


Figure 2. Semivariogram for January SST showing distance in degrees versus variance.

A nice feature of the semivariogram is that it is usually apparent where to cut off the range value, whereas in the corresponding distance function of OI the range value may initially be less obvious. As with all exploratory modeling, several iterations are necessary to achieve a "best" model.

In kriging, the estimate of an unknown point using a weighted linear combination of available samples is given by

$$\hat{z}_p = \sum_{i=1}^{n} W_i z_i . \qquad (3.2)$$

This type of kriging is known as ordinary kriging and is perhaps the most straightforward of the several versions. It also yields a standard error estimate with each kriged value. Our kriging efforts were accomplished using the GEOEAS software package (Englund and Sparks, 1991).

4. RESULTS

We begin our comparison of OI and kriging results by examining the South American precipitation data. Figures 3a and 3b gives the January 20-year standard deviation at each grid box location for OI and kriging respectively. Values in bold indicate those locations for which estimates were made for each of the 20 years. A handful of boxes required an estimation for just a few years (such as the upper left most box), hence the slight difference in values at these locations noted in the two figures.

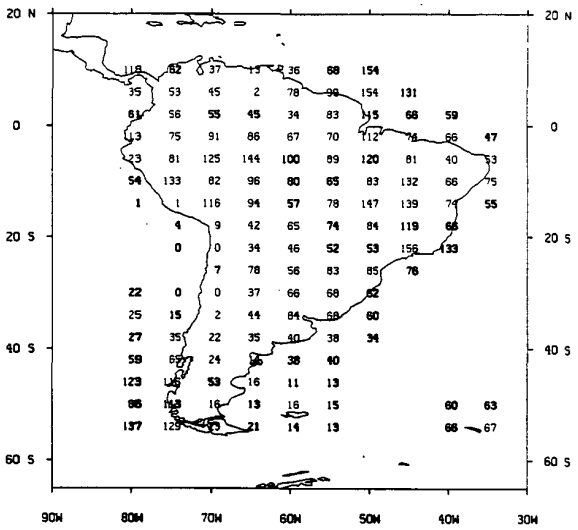


Figure 3a. January standard deviations from OI for the South American precipitation grid. Bold values indicate boxes where estimates were generated.

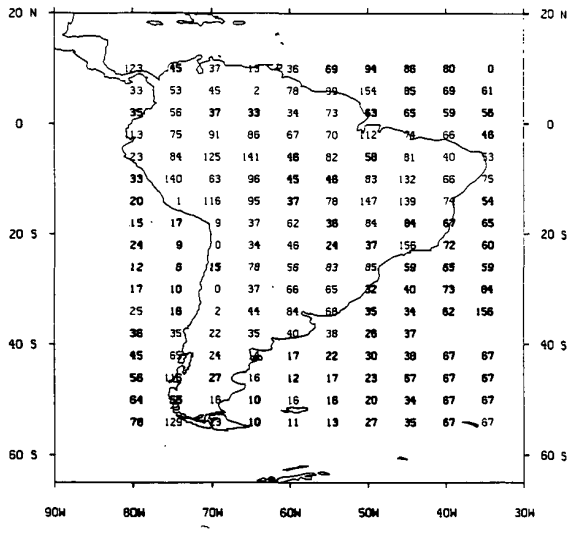


Figure 3b. Same as 3a except for kriging.

In general, the standard deviations are larger from OI estimates than from kriging. This suggests that kriging has estimated a somewhat smoother field than OI. Similar results occurred in the July case. It is known that

precipitation exhibits large variance both spatially and temporally over this area, so perhaps the kriging estimates are too smooth. Thus, one might argue at this point that OI has generated somewhat better estimates.

Further evidence that OI may have yielded better estimates is offered by examining the differences of the actual estimates over the entire period. Figure 4 shows the 20-year mean estimate difference for OI minus kriging, which indicates that for the most part kriging is consistently underestimating values compared to OI, and in some instances quite substantially. Again, similar results were found for July.

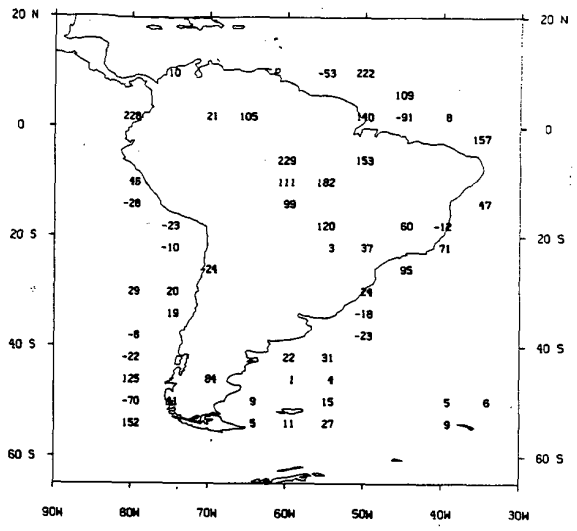


Figure 4. OI minus kriging differences for estimated grid locations for January South American precipitation data.

Area weighted averages of the entire grid through time provides another way of comparing the two techniques. In Figure 5, area weighted averages based upon the OI estimates and kriged estimates are plotted through time for both January and July. Here we can see that although the phase is similar throughout the period, kriging has clearly underestimated values compared to OI. However, when compared to the averages from the original gridded field, the results now suggests that OI may have overestimated the precipitation values and that the kriging estimates are much closer to the original values, despite the smaller kriging standard deviations.

Comparing estimates based on SST values, the two techniques appear overall to have yielded similar results. Figures 6a and 6b give the January 40-year standard deviations for OI and kriging, respectively. Unlike most of the missing precipitation grid boxes, the SST boxes are not consistently missing throughout the entire period. Hence, the bold values represent boxes in which at least 10 years out of the 40 possible were estimated for. The standard deviations appear in general to be smaller in the kriging case, and perhaps one could argue that the OI standard deviations are slightly too large. These results were similar in the July case.

Figure 7 shows the 40-year mean estimate difference for OI minus kriging, which indicates that OI is generally underestimating values compared to kriging, just the opposite of the precipitation case. These results were also similar for July.

The differences are also illustrated in Figure 8, which shows the area weighted averages based upon the OI and kriged estimates. The averages are very similar, especially in January where the averages derived from kriging are only slightly larger than OI. Compared to the observed averages, both OI and kriging have generated values which are slightly lower than the original grid.

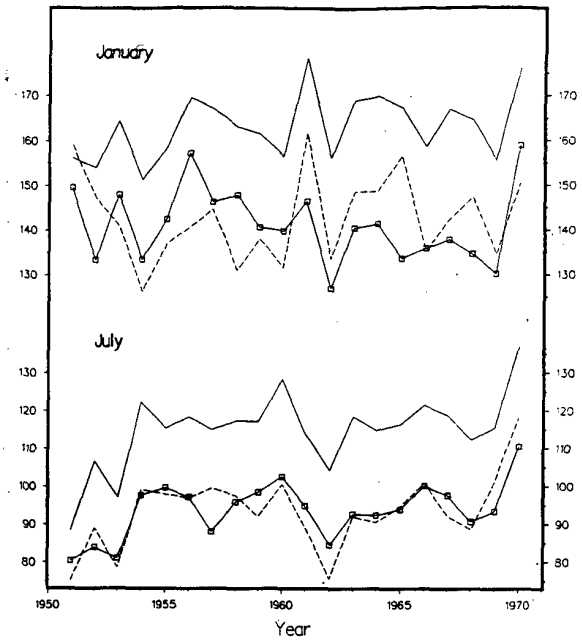


Figure 5. January and July area weighted time series of South American precipitation data for OI (solid line), kriging (dash line), and original gridded values (solid line with open boxes).

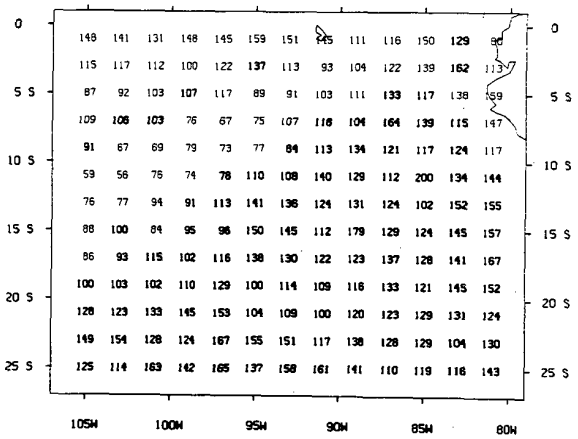


Figure 6a. January standard deviations from OI for the SST grid. Bold values indicate boxes where estimates were generated.

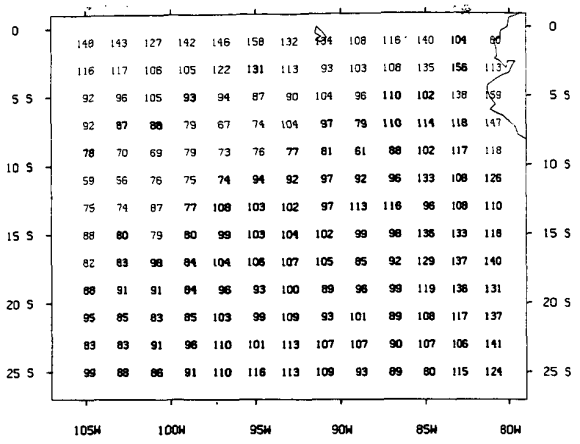


Figure 6b. Same as 6a except for kriging.

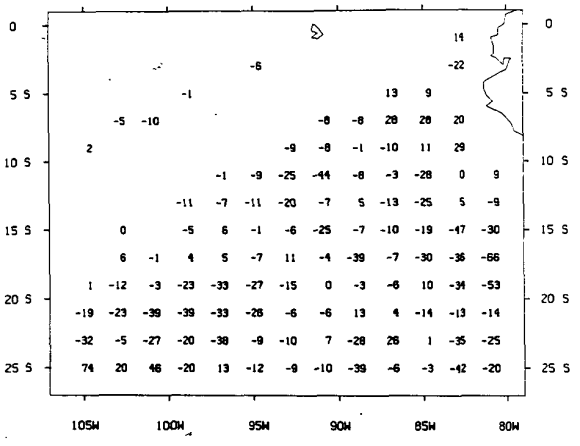


Figure 7. OI minus kriging differences for estimated grid locations for January SST data.

5. SUMMARY

We have attempted here to present some initial results of comparing OI and kriging methods on gridded precipitation and sea surface temperature data. Since there are no "truth" values at the estimated locations for direct evaluation, it is difficult to say conclusively which technique has yielded the better estimates. Some of the evidence suggests that OI worked better for precipitation values (which has a large field variance), but perhaps kriging worked better on the smaller variance SST values. However, kriging yielded values closer to the original grid in both data sets. It is clear that much further work is needed. For example, further experimentation in developing the OI distance function and semivariogram is desirable. This should include some exploration in variable transformations for example. Further work in estimating known values through cross-validation will also provide valuable insight as to the model adaquecy.

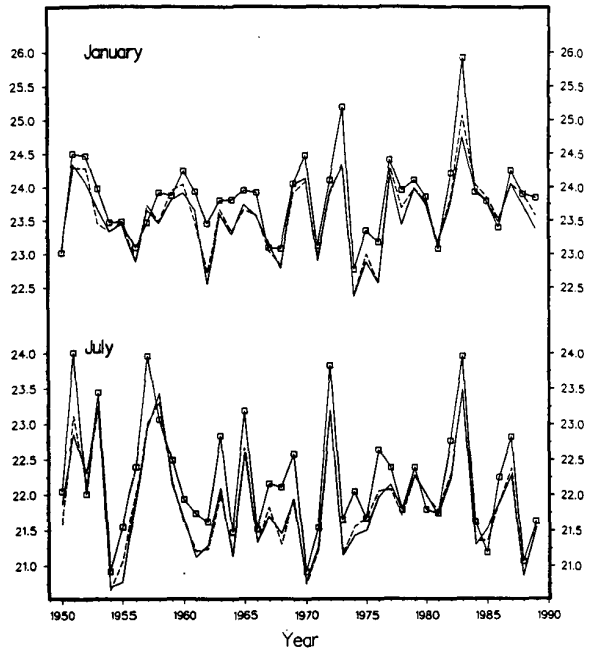


Figure 8. January and July area weighted time series of SST data for OI (solid line), kriging (dash line), and original gridded values (solid line with open boxes).

Acknowledgments

This work is part of an effort by several NOAA units, with funding from NOAA's Climate and Global Change Program, to develop a global climate perspectives system for climate diagnostics research and climate change detection efforts. We would like to thank Patricia Miller of NOAA/ERL/FSL for valuable discussions regarding optimal interpolation, and Naser Heravi of the Environmental Statistics and Modeling Division, Harry Reid Center, University of Nevada, Las Vegas, for helpful kriging discussions.

References Cited

Cressie, N.A.C., 1991: *Statistics for Spatial Data.* John Wiley & Sons, New York, 900 pp.

Eishceid, J.K., H.F. Diaz, R.S. Bradley, and P.D. Jones, 1991: *A comprehensive precipitation data set for global land areas.* U.S. Department of Energy report DOE/ER-69017T-H1, 82pp.

Englund, E. and A. Sparks, 1991: *Geostatistical Environmental Assessment Software User's Guide.* EPA 600/8-91/008, U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Las Vegas, Nevada, 89119.

Gandin, L.S., 1963: *Objective Analysis of Meteorological Fields.* Gidrometeorologicheskoe Izdatel'stvo (GIMIZ), Leningrad (translated by Israel Program for Scientific Translations, Jerusalem, 1965).

Matheron, G., 1963: Principles of geostatistics. *Economic Geology*, **58,** 1246-1266.

Woodruff, S.D., R.J. Slutz, R.L. Jenne, and P.M. Steurer, 1987: A comprehensive ocean-atmosphere data set. *Bull Amer. Meteor. Soc.*, **68,** 1239-1250.

COMBINING FORECASTS: AN OVERVIEW

Robert T. Clemen

University of Oregon
Eugene, Oregon

1. INTRODUCTION

When two or more forecasts differ, how can one reconcile the differences? One approach is to try to choose the model that is "best" in some sense (most accurate forecasting record, most appropriate model of the underlying process, most cost-effective to apply). Another possibility is to combine the individual forecasts into a single composite forecast.

Forecast combinations have been researched and applied in a wide variety of settings. Much of this research is summarized and reviewed by Clemen (1989), who covers contributions to the theory from economic and business forecasting, statistics, management science, and psychology. Many applications are described, including a number from meteorology, and an extensive annotated bibliography provides easy access to this diverse literature. For more recent additions to the literature, excellent resources are the *Journal of Forecasting* and the *International Journal of Forecasting*.

This paper provides a brief overview of forecast combination, especially as it relates to the field of meteorology. The next section proposes a Bayesian motivation for combining forecasts. Section 3 gives a brief history, and Section 4 describes meteorological and other applications. Section 5 concludes with a discussion of possible research topics.

2. A BAYESIAN PERSPECTIVE

From a conventional forecasting point of view, combining forecasts heads in the wrong direction. In contrast to finding the "true" underlying process, combining forecasts amounts to an admission that the forecaster is unable to build a properly specified model. If several different models can be combined to obtain a better forecast, it should (theoretically, at least) be possible to construct a single model that makes optimal use of the different kinds of information used by the individual forecasts. Forecast combination can be viewed as a compromise approach while awaiting the development of the more complete single model (Diebold, 1989).

A more useful approach is to view the problem as one in which a decision maker needs to make the best possible use of available information. Given a set of diverse forecasts (information), what is an appropriate posterior probability distribution for the event in question (Morris, 1974, 1977)? From this perspective, it is desirable to construct Bayesian models to help decision makers use multiple forecasts effectively. Virtually all of the combination models in use can be viewed from this Bayesian perspective.

3. HISTORY OF FORECAST COMBINATION

The seminal article on forecast combination is often thought to be Bates and Granger (1969), in which the authors develop and test several techniques for combining point forecasts. Although other work had appeared earlier, Bates and Granger provided the initial impetus for the development of theory in forecast combination. During the 1970s, a stream of articles followed in the *Operational Research Quarterly* developing various statistical models for combining. For example, in contrast to Bates and Granger's minimum-variance model, Bunn (1975) presents a combination based on how likely it is that one forecast will be more accurate than another.

At roughly the same time, Nelson (1972) devised methods for studying the efficiency of economic forecasts. These methods essentially use a composite forecast as a benchmark. He shows that certain econometric forecasts are inefficient in the sense that combining the econometric forecast with a naive time-series model significantly reduces the forecasting error. Granger and Newbold (1973) address similar issues in the context of forecast evaluation.

Many studies have examined real-world performance of forecast combinations. For example, Newbold and Granger's (1974) results demonstrate that in practice weighted averages that depend on estimated correlations perform poorly. Makridakis, et al. (1983) report on the "M-Competition," a forecasting competition in which a large variety of forecasting methods are applied to 1001 different economic time series. Two different combining schemes are studied along with a number of individual forecasting methods. Both of the combinations perform well relative to the individual techniques and are robust, having a relatively high degree of accuracy for many kinds of time series.

Work on forecast combination in recent years has formulated the combining problem in terms of a statistical regression model (Granger and Ramanathan, 1984). In such a model, the actual value is the response variable, and the forecasts are the explanatory or right-hand-side variables. With adequate data, one can estimate combining weights with appropriate statistical techniques. Much recent work explores the statistical properties of component forecasts and uses advanced econometric techniques to estimate the weights. For a more complete discussion, see Clemen (1989).

Although typically not thought of as falling within the mainstream of forecast-combination research, contributions to the field have been made by psychologists and management scientists. On the psychological side, a number of studies examine the mechanical combination of expert judgments. For example, Goldberg (1965) presents evidence that in a clinical situation, averages of judgments are more accurate than the individual judgments. Einhorn and Hogarth (1975) give a theoretical explanation of the strong performance of equally weighted combinations, and Einhorn, Hogarth, and Klempner (1977) use the average of the individual judgments as a baseline for the performance of group judgment.

A major contribution from statistics and management science addresses the combination of probability distributions. Genest and Zidek (1986) and French (1985) review this rich literature, which includes Morris's (1974, 1977) Bayesian paradigm mentioned above. Forecast combination methods also have been inspired by operations research models for dealing with multiple objectives. For example, Lawrence and Reeves (1981), show how multiple-objective linear programming can be used to combine forecasts. Gupta and Wilton (1987) combine Bunn's method with the mathematics of Saaty's (1980) Analytic Hierarchy Process, a multiple-objective decision-making tool, to create a novel subjective approach to forecast combination.

4. APPLICATIONS

Meteorologists have used forecast combinations, but not widely. Sanders (1963) discusses probability forecasting in meteorology and reports an experiment in which the average of two forecasters' subjective probabilities perform better than either individual probability. Staël Von Holstein (1971) reports a similar but more extensive experiment. Winkler, Murphy, and Katz (1977) present several probability-consensus models and measure the performance of these models in terms of a quadratic score for a set of probability forecasts. The simple average of the forecasts performs well compared to models that use past performance data, and performance improves as more forecasts are included in the consensus.

More recently, Thompson (1977) and Fraedrich and Leslie (1987) develop formal statistical techniques for combining forecasts. In particular, Fraedrich and Leslie use their model to combine both probability and binary forecasts of precipitation in Australia; the combined forecast substantially improves on the component forecasts as measured by the half-Brier score. Independently, Clemen and Murphy (1986a, b) and Murphy, Chen, and Clemen (1988) apply forecast combination techniques to measure informational contributions of component forecasts.

Other areas have used forecast combinations. The most extensive use has occurred in economic forecasting, including forecasting of inflation, money supply, exchange rates, stock prices, livestock prices, corporate earnings, revenues, electrical demand, and tourism. Some studies have even attempted to quantify potential economic benefits of combining forecasts. Still other fields include the prediction of social and technological events, population levels, psychiatric diagnosis, football game scores, wilderness area use, political risk, and sunspot cycles. Clemen (1989) provides references.

5. CONCLUSION AND FUTURE RESEARCH

Combining forecasts has been shown to be practical, economical, and useful. Underlying theory has been developed, and many empirical tests have demonstrated the value of the technique. The robustness of simple combination techniques in particular makes the practice of combining forecasts very straightforward. Perhaps the main issue is simply implementation; practitioners should be encouraged to consult multiple information sources and combine the information using simple techniques like averaging.

Research on forecast combination continues, but the focus has shifted away from a search for the optimal weighting methods. Instead, interest has turned to explanation of the robustness of simple combination techniques. For example, Winkler and Clemen (1992) and Gunter (1991) study the sensitivity of combination weights to statistical variation, and Miller, Clemen, and Winkler (1992) examine the effect of nonstationarity on the performance of various combining methods. Whether the robustness of these simple techniques would extend to the meteorological field is an intriguing question; with an abundance of data and a more or less stationary system, it may be possible for meteorologists to exploit interdependencies among forecasts to a much greater extent than can economists, who must deal with substantial nonstationarity in the economic system.

Taking the research another direction altogether, Batchelor and Dua (1991) study the effects of combining economic forecasts that arise from different methodologies, and they find substantial improvements in forecasting performance by doing so. Combining meteorological forecasts arising from different models could prove worthwhile because alternative models are often complementary. For example, Fraedrich and Leslie (1987) combine forecasts from very different models, while Clemen and Murphy (1986a, b) combine model-based forecasts and subjective forecasts. Again, the abundance of meteorological data may lead to improved forecasting performance by combining forecasts from very diverse models.

ACKNOWLEDGEMENT

This research was supported by the National Science Foundation under Grant SES-9022616.

REFERENCES

Batchelor, R., and P. Dua, 1991: Forecaster diversity and the benefits of combining forecasts. Unpublished ms, City University Business School, London.

Bates, J.M., and C.W.J. Granger, 1969: The combination of forecasts. *Op. Res. Q.*, **20**, 451-468.

Bunn, D.W., 1975: A Bayesian approach to the linear combination of forecasts. *Op. Res. Q.*, **26**, 325-329.

Clemen, R.T., 1989: Combining forecasts: A review and annotated bibliography. *Int. J. Forecast.*, **5**, 559-583.

Clemen, R.T., and A.H. Murphy, 1986a: Objective and subjective precipitation probability forecasts: Statistical analysis of some interrelationships. *Wea. Forecast.*, **1**, 56-65.

Clemen, R.T., and A.H. Murphy, 1986b: Objective and subjective precipitation probability forecasts: some methods for improving forecast quality. *Wea. Forecast.*, **1**, 213-218.

Diebold, F., 1989: Forecast combination and encompassing: Reconciling two divergent literatures. *Int. J. Forecast.*, **5**, 589-592.

Einhorn, H.J., and R.M. Hogarth, 1975: Unit weighting schemes for decision making. *Org. Behav. Hum. Performance.*, **13**, 171-192.

Einhorn, H.J., R.M. Hogarth, and E. Klempner, 1977: Quality of group judgment. *Psych. Bull.*, **84**, 158-172.

Fraedrich, K., and L. M. Leslie, 1987: Combining predictive schemes in short-term forecasting. *Mon. Wea. Rev.*, **115**, 1640-1644.

French, S., 1985: Group consensus probability distributions: A critical survey. In J. Bernardo, M. DeGroot, D. Lindley, & A. Smith (Eds.), *Bayesian statistics 2* (pp. 183-197). Amsterdam: North-Holland.

Genest, C., & Zidek, J.V. (1986). Combining probability distributions: A critique and annotated bibliography. *Stat. Sci.*, **1**, 114-135.

Goldberg, L.R., 1965: *Diagnosticians versus diagnostic signs: The diagnosis of psychosis vs. neurosis from MMPI.* Psychological Monographs, No. 79.

Granger, C.W.J., and P. Newbold, 1973: Some comments on the evaluation of forecasts. *App. Econ.*, **5**, 35-47.

Granger, C.W.J., and R. Ramanathan, 1984: Improved methods of forecasting. *J. Forecast.*, **3**, 197-204.

Gunter, S., 1991: Theoretical justification of the efficiency of simple average combinations. Unpublished ms, Temple University.

Gupta, S., and P.C. Wilton, 1987: Combination of forecasts: An extension. *Mgmt. Sci.*, **33**, 356-372.

Lawrence, K.D., and G.R. Reeves, 1981: Consensus time series forecasting. In J. Morse (Ed.) *Organizations: Multiple Agents with Multiple Criteria*, (pp. 199-204). New York: Springer-Verlag.

Makridakis, S., A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, and R. Winkler, 1983: *The forecasting accuracy of major time series methods.* London: Wiley.

Miller, C., R. Clemen, and R. Winkler, 1992: The effect of nonstationarity on combined forecasts. *Int. J. Forecast.*, **7**, in press.

Morris, P.A., 1974: Decision analysis expert use. *Mgmt Sci.*, **20**, 1233-1241.

Morris, P.A., 1977: Combining expert judgments: A Bayesian approach. *Mgmt Sci.*, **23**, 679-693.

Murphy, A.H., Y.-S. Chen, and R.T. Clemen, 1988: Statistical analysis of interrelationships between objective and subjective temperature forecasts. *Mon. Wea. Rev.*, **116**, 2121-2131.

Nelson, C.R., 1972: The prediction performance of the F.R.B.-M.I.T.-PENN model of the U.S. economy. *Am. Econ. Rev.*, **62**, 902-917.

Newbold, P., and C.W.J. Granger, 1974: Experience with forecasting univariate time series and the combination of forecasts (with discussion). *J. Roy. Stat. Soc.*, A, **137**, 131-149.

Saaty, T., 1980: *The Analytic Hierarchy Process.* New York: McGraw-Hill.

Sanders, F., 1963: On subjective probability forecasting. *J. Appl. Met.*, **2**, 191-201.

Staël Von Holstein, C.-A.S., 1971: An experiment in probabilistic weather forecasting. *J. Appl. Met.*, **10**, 635-645.

Thompson, P.D., 1977: How to improve accuracy by combining independent forecasts. *Mon. Wea. Rev., 105*, 228-229.

Winkler, R., and R. Clemen, 1992: Sensitivity of weights in combining forecasts. *Op. Res.*, **40**, in press.

Winkler, R., A.H. Murphy, and R. Katz, 1977: The consensus of subjective probability forecasts: Are two, three, ..., heads better than one? Preprint Volume: Fifth Conference on Probability and Statistics, Nov., 1977, Las Vegas, Nevada (pp 57-62). Boston, MA: AMS.

J3.2

COMBINATION OF WEATHER FORECASTS: SOME APPLICATIONS

Klaus Fraedrich

Institut für Meteorologie
Freie Universität Berlin
D-1000 Berlin 41, Germany

1. INTRODUCTION

Optimal weighting of forecasts has received much attention in economics, management and statistics literature. In meteorology it is also known that consensus forecasts and the (linear) combination of predictions provide more accurate results than the individual forecasts which comprise the consensus. Although 'this is the incontrovertible fact' it still 'does not appear to be widely recognized or accepted' (Thompson 1977). Furthermore, 'the problem of forecast complexing has not jet received the attention that it merits' (Zhunkovskii and Brunova 1978). In particular short-term weather and short term climate (or long-range forecasting (SECTION 2, 3) are areas where a combination of two or more independent predictions is a promising technique to directly improve the forecast accuracy, in particular, if one of the forecast models is based on statistical or analogue methods. Both very short-term and long-range forecasts cover basically the 'nowcasting' period of the evolution in the weather or the climate state space where persistence forecasts yield relatively high accuracy. In the medium to extended range (SECTION 4) predictions cover a larger part of the state space and the combination of an ensemble of forecasts may serve another, possibly weaker, purpose. Rather than to improve the skill, quantitative measures of the forecast confidence, or the predictability, can be obtained as an additional information. In the following discussion about forecast combinations the emphasis lies on the range of applications and the challenges of mixing forecasts. It is not intended to completely cover the whole field nor to present all details.

2. LONG-RANGE WEATHER FORECASTING

Long range weather forecasting is still considered as 'an empirical art....with a mix of techniques that are statistical to some extent' (Gilman 1985). A famous mix of statistical techniques is the first order autoregressive process, because it can be interpreted as an error minimizing combination of two anomaly predictions, persistence and climate. This will be shown in the introductory example where regression type point processes are considered (Fraedrich and Smith 1989, applying Thompson 1977). This is followed by multifield predictions based on past weather analogues (Livezey et al 1990, following Barnett and Preisendorfer 1978).

2.1 Point processes

Consider two normalized anomaly forecasts, $F(t) = A(t)$ or $B(t)$, predicting the observed anomaly $X=(x-\langle x\rangle)/s$, where x is the observable, X is the anomaly, $\langle\rangle$ is the ensemble mean and s the standard deviation. These anomaly forecasts can be combined in an error minimizing fashion (as shown, for example, by Thompson 1977):

$$F^*(t) = aA(t) + bB(t)$$

Minimizing the ensemble mean square error, $E^*=\langle(X-F^*)^2\rangle$, of the combination forecast with respect to the weights a,b leads to the weights $a=(R_A-rR_B)/(1-r^2)$, $b=(R_B-rR_A)/(1-r^2)$, and the hindcast skill of the combination $F^*(t)$:

$$S^* = 1-E^* = (R_A^2+R_B^2-2rR_AR_B)/(1-r^2) .$$

R_A, R_B are the simultaneous correlations between the forecasts (A,B) and the observable, X (or x); r is the correlation between the forecasts A and B. Individual anomaly predictions by a climate mean forecast, F=0, yield a skill $S_C=0$; a persistence forecast model, X(t-n), leads to a skill $S_P=2R_{XX}(t-n,t)$; an auto-regression, $F(t)=X(t-n)R_{XX}(t-n;t)$; a regression forecast with the predictor y for the anomaly X, $F(t)=y(t-n)R_{XY}(t-n,t)$, leads to a hindcast skill $S_{A,R}=R^2(t-n,t)$.

CLIMATE AND PERSISTENCE: The climate-persistence combination yields a first order autoregressive process with a skill, $S_a=R_{XX}^2(t-n,t)$, which is better than both forecasts alone: $S_A-S_P>0$, $S_A-S_C>0$ (see also FIGURE 1). It may be used to define the level of zero-skill for combined predictions which other combinations have to pass.

COMBINATION FORECAST (FIGURE 1; regression and auto-regression): An application of the combination technique to a variety of simple El Nino/Southern Oscillation point forecasts of the monthly mean SST for July in the equatorial Pacific is shown in FIGURE 1. In the average the linear error minimizing combination provides more accurate predictions than the individual forecasts alone.

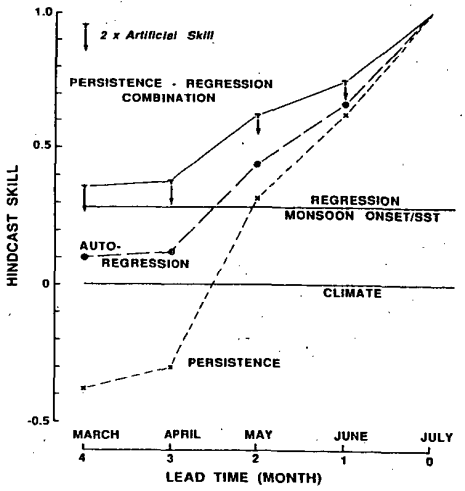


Fig. 1 (Fraedrich and Smith 1989, Fig. 1). Hindcast skill for the prediction of the monthly mean sea surface temperature (SST) of the eastern equatorial Pacific (for July) using the individual forecast schemes (climate, persistence, auto-regression and a regression between the data of the onset of the Australian monsoon in Darwin and the east equatorial Pacific SST) and the related combinations. A zero-order measure of the artificial skill is included.

2.2 Multifield predictions (past weather analogues)

PAST WEATHER ANALOGUES (PREDICTORS AND PREDICTANDS): Past weather analogs for long range forecasting are determined in a climate state space spanned, for example, by the principle components of sea-land surface and upper air data, circulation indices, position of troughs and ridges, etc.. These state vectors serve as predictors, which are a lead time (say a month or season) ahead of the forecast period. The predictands at forecast time are long-range weather variables, say, temperature or precipitation at single stations. The predictand may be measured in probabilistic terms of equally probable classes: above, near, below normal or, in categorical terms as normalized anomalies. Both analogs and anti-analogs (antilogs) of historic weather, which serve as the predictors, are selected from the past weather files by evaluating the Euclidean distance between a present, A, and all historic, B, states:

$$D^2(i) = \langle A(t)-B(t_i))^2\rangle \quad \text{or} \quad D^2(i) = \langle A(t)+B(t_i))^2\rangle.$$

The L best analogs or antilogs are only a small distance apart from the present state and ranked, r=1,...,L, by their increasing distance to obtain composites. Forecasts based on the best analog only, suffer from the large variability between the individual cases. Therefore, the analog time evolution needs to be combined before actual predictions can be made:

COMPOSITE (CLUSTER) ANALOG FORECASTS: A set of L analogs as predictors asociated with the predictand leads to a forecast in probabilistic or in categorical terms: forecast probabilities in quantile classes, P, are conditional on the present state at a station; the forecast anomaly, T, is estimated by weighted averages of the L best analogs:

$$P = \Sigma w(i)P(i) \quad \text{or} \quad T = \Sigma w(i)T'(i) \text{ for } i=1,..,L$$

P(i)=1 or 0, if the quantile class has been observed or not after the forecast lead time Δt, or $T'(i) = +/-T(t_i+\Delta t)$, if the past weather date 'i' was an analog/antilog situation. The weights, $w(i)=g/D^2(i)$, are large for good analogs; they depend on the normalized analog/antilog distance, $g=1/\Sigma D^{-2}$. Alternatively, a cluster analysis of L good analogs (after having evolved to the forecast time; Toth 1989) reduces the L best analogs to a subset L*<L, which, of course, does not necessarily comprise the best.

COMBINATION FORECASTS (ANALOG AND PERSISTENCE): The forecasts based on analogs and antilogs can now be made either by using the analogs alone or, in association with other predictands, by specifying the class with the largest transition probability, P, or using the weighted anomalies, T. The latter will contain an excess of near normal forecasts compared to those inferred from the most probable class, because opposite signs of T' will tend to cancel each other. These anomaly forecasts can be developed further, if combining them linearly with persistence of the anomaly. Persistence is a useful and independent reference scheme, both for short term climate prediction or long-range forecasting and short term weather prediction or nowcasting. Two combinations have been suggested:

PERSISTENCE/ANALOG: $T^*(t) = aT_P(t-n) + bT_A(t)$

The combination forecast, T*(t), is the linear combination of the persistence and the anomaly forecast, $T_P(t-n)$ and $T_A(t)$, where n is the lead time for persistence and t is the running forecast time. The weights a,b are determined by bivariate linear regression to minimize the mean square error of T* of the dependent data set (for more details, see Livezey et al. 1990).

PERSISTENCE/ANALOG TREND: $T^*(t) = T_P(t-n) + cdT_A$

The first term is the persistent anomaly, the second is the weighted trend. The weight may be fixed, c=1, or minimizing the square error on the dependent data, $c=\langle dT'dT_A\rangle/\langle dT_A^2\rangle$; the d's indicate the differences between future and past months or seasons, $dT'=T(t+n)-T(t)$, $dT_A'=T_A(t+n)-T_A(t)$; and <> denotes the average over the dependent sample.

RESULTS (FIGURE 2): The two basic combinations, that is the bivariate linear regression and the persistence plus analog trend, are overall more skillful than either of its components; the bivariate regression is slightly better.

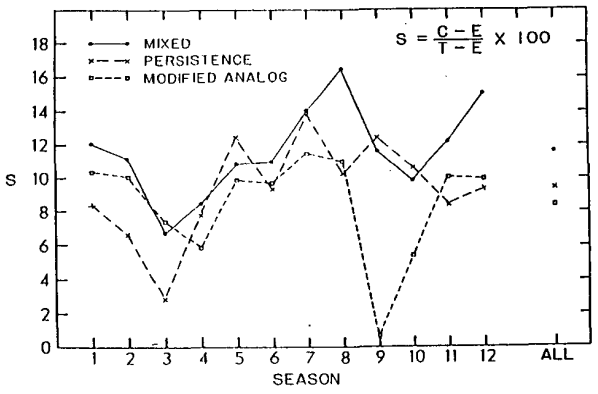


Fig. 2 (Livezey et al. 1990, figure 5). Skill versus 3 months forecast periods (number denotes the first month of the period) of the prediction of three equally distributed temperature categories made by persistence, the analog scheme and a combination; the skill, S=(C-E)/(T-E)is defined by the total number of forecasts, T, the number of correct, C, and the number expected to be correct by chance, E, respectively.

3. SHORT-TERM FORECASTING

In short term forecasting, the prediction schemes can also be combined to improve the forecasts. It has been shown theoretically and in real time trials that the combination of both stochastic and deterministic models leads, in the average, to superior forecasts. Here we will focus on probability forecasts. Again, point processes and multivariate combinations are described.

3.1 A point process: single station probability forecasts

Binary and/or probabilistic forecast schemes are used in this application which, for example, may be provided by forecasts generated by a numerical weather prediction (NWP) model with or without the use of model output statistics, MOS, and a stochastic forecast model, for example, a Markov chain. The observable or verification is interpreted as a binary variable X (=1 or 0; for example, the rainfall occurrence). Then two probabilistic forecasts can be combined:

$$PoP^*(t) = aP_A(t) + (1-a)P_B(t)$$

Minimizing the ensemble mean squared error (or half-Brier score), $B^*=\langle(PoP^*-X)^2\rangle$, yields $a=\{\langle XP_A\rangle-\langle XP_B\rangle+\langle P_B^2\rangle-\langle P_AP_B\rangle\}/\{\langle P_A^2\rangle+\langle P_B^2\rangle-2\langle P_AP_B\rangle\}$, so that the half-Brier score of the combination is obtained (Fraedrich and Leslie 1987):

$$B^* = \langle(X-P_B)^2\rangle - \{\langle XP_A\rangle-\langle XP_B\rangle+\langle P_B^2\rangle-\langle P_AP_B\rangle\}^2/\{\langle P_A^2\rangle + \langle P_B^2\rangle-2\langle P_AP_B\rangle\}$$

$B_B = \langle (X-P_B)^2 \rangle$ is the half-Brier of the model P_B; the other terms in brackets, <>, are variances (or co-variances) of predictions and observations, which are not taken as deviations from their respective means. The forecast schemes may now be combined subject to further testing.

CLIMATE AND PERSISTENCE: The climate-persistence mix may be used as a zero-skill level for combined predictions. The climate forecast, $P_C=\langle X\rangle=\langle P_C\rangle$, leads to a half-Brier score, $B_C=\langle (X-P_C)^2\rangle=P_C-P_C^2$, which corresponds to the natural variance. Persistence is interpreted as a binary model with a half-Brier score, B_P. Now the weight of the climate-persistence combination can be deduced, $a=1-B_P/2B_C$. The combined half-Brier score gives $B^*=B_P-(B_P/2B_C)^2$. If the relative frequency of incorrect predictions coincides with the natural variability, that is $B_P=B_C$ and $a=0.5$, so that one obtains a 25% improvement of the combined forecast, $B^*=3/4B_C$. In this sense categorical forecasts can be reliably recalibrated so that they are reliable probability forecasts (Murphy 1986).

COMBINATION FORECASTS (MARKOV CHAIN AND NWP LIMITED AREA MODEL): Markov chains of single station weather can be used to predict the probability of precipitation, $0<P_M<1$, say three or more hours ahead. Here it is based on four mutually exclusive states: three cloud cover and a rain state using 3-hourly international weather code information (see Fraedrich and Leslie 1987, 1988). The NWP model for regional short term forecasting is the Australian Bureau of Meteorology operational limited area model (LAM, 12 levels, 150 km resolution, primitive equations). Rainfall predictions are interpolated from the four nearest gridpoints of the weather station. The results of two real time short-term single-station forecast trials are briefly described: one for an Australian mid-latitude station (Melbourne), the other for the tropical wet season (Darwin).

* MID-LATITUDES (Markov and NWP): Probability of precipitation forecasts for the 12 hours local time was predicted for Melbourne each day for 3 months in winter (June-August 1986) using six differnt techniques: a Markov chain, the limited area NWP model, a weighted combination, a model output statistics, an analogue statistics and the manual official Bureau forecast. The results indicate that the skill of the combined Markow-NWP forecasts considerably exceeds the other techniques showing about 90% correct 12h-forecasts of rainfall occurrence. The Markov model was next followed by the other methods which were close. Care was taken as there were different lead times associated with NWP forecasts and the MOS and analogue schemes depending on it owing to the

* TROPICS (short-term Markov and synoptic persistence): A similar trial was performed for the tropical station Darwin during the wet season (December-February 1986/87). The results lead to a minimal model, which is proposed for the short term prediction of precipitation occurrence in the tropics. This model is purely statistical and consists of an optimally weighted linear combination of two prediction schemes: Markov chain and persistence based on surface observations only. It appears that persistence on a day to day basis describes the dry and wet spells of the Monsoon regime representing the synoptic and larger scale fluctuations. The Markov chain, which uses the observed data at the time the forecast is made, provides the embedded meso-scale information. This model is minimal in the sense that only surface observations are needed and computation requirements are almost nil. This scheme has been found to be the most skillful of all forecast methods and combinations tested.

MODEL	MELBOURNE	DARWIN
Markov	0.164	0.147(0.173)
NWP	0.258	0.461(0.539)
MOS	0.184	0.221(0.223)
Climate	0.259	0.153(0.216)
Markov and NWP	0.142	
Climate and NWP	0.196	
Markov and Persistence		0.127(0.167)
Persistence and Climate		0.142(0.209)
Persistence		0.222(0.356)

TABLE 1 (Fraedrich and Leslie 1987,1988,tables 1): Half-Brier score for various schemes for the real time short term trials in Melbourne (July 1986) and Darwin (December-February 1986/87). For Darwin half-Brier scores are given for both WW-rain (including rain in the neighbourhood) and RR-rain (at the station; in parentheses). The forecast period was the twelve hour interval 0600 to 1800 local time. Note that the Markov chain involves no lead time whereas NWP and MOS have 21 h lead time.

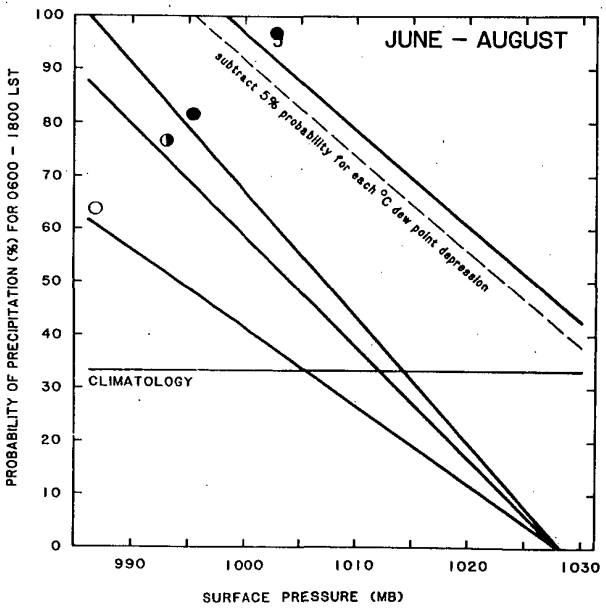


Fig. 3 (Fraedrich and Leslie 1987, Fig. 5). Melbourne single station probability of precipitation (0600-1200 LST) changing with surface pressure, cloud cover and rainfall states observed during the preceeding hour. From top to bottom are the states of rainfall and cloud cover (6-8, 3-5, 0-3 octas). Note that for the rainfall state the PoPs decrease with decreasing dewpoint depression.

3.2 Multivariate combination forecasts of tropical cyclone tracks (analog/statistical and NWP)

HURRAN/CLIPER: The hurricane analog (HURRAN) technique selects analogs for an existing tropical storm or hurricane. It examines all tropical cyclones in a particular region and identifies and selects those that have designated characteristics similar to an existing storm. Positions of storms selected as analogs are determined at successive time steps (of 6 or 12 h length) after the initial time. When the analog technique fails because of insufficent samples, a climate-persistence (CLIPER) nonlinear multiple regression scheme is applied fitting climatology

and persistence to the analog predictors (Hope and Neumann 1970, Neumann and Pelissier 1981, and developed at BMRC, Melbourne). Furthermore, we use a *regression* (REGR) technique developed by Keenan (1986) for the Australian region tropical cyclones with 6 h time steps. The BMRC tropical NWP model system consists of a univariate optimal interpolation analysis scheme and the semi-implicit primitive equations with a horizontal resolution of 100 km and ten levels.

COMBINATION FORECASTS (CLIPER AND TROPICAL NWP MODEL): Denoting the zonal and meridional storm displacements by X and Y, the linear combination forecast is

$$X = a_1X_A + a_2Y_A + a_3X_B + a_4Y_B + a_5$$
$$Y = a_6X_A + a_7Y_A + a_8X_B + a_9Y_B + a_{10}.$$

The values of the coefficients a_1-a_{10} are obtained by applying standard multiple regression techniques to more than 50 tropical cyclone best tracks observed during the seasons 1979/80-1983/84; the seasons 1984/85-87/88 were used as an independent data set to verify the performance of the combination. Each cyclone track was divided into 12-h sections and a different set of regression coefficients was obtained for each 12-h period. The CLIPER forecasts were weighted more heavily in the first 12 hours, but the NWP model gained more weight later in the forecast period.

METHOD	12-h	24-h	48-h
CLIPER	086;065;239	185;116;221	379;236;185
BMRC NWP	102;073;225	217;122;205	362;215;171
REGR	091;059;221	189;103;208	355;179;173
ECMW		227;117;196	369;207;165
COMB	069;041;225	157;089;205	312;147;171

TABLE 2 (Leslie and Fraedrich 1990, table 1): Mean tropical cyclone position errors based on indepent 'best track' data of five schemes evaluated in a comparison of performance over four tropical seasons 1984/85-1987/88: mean error (in km); standard deviation (in km); number of forecasts.

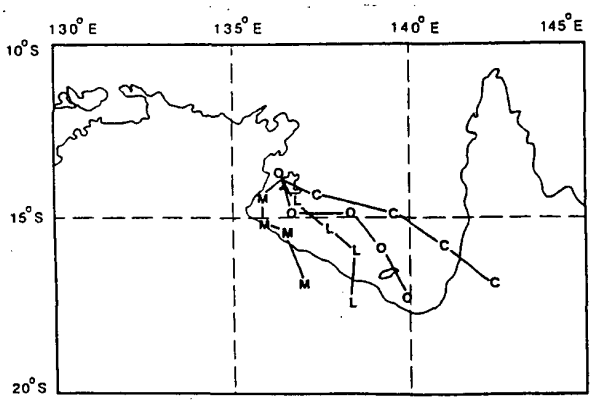


Fig. 4 (Leslie and Fraedrich 1990, Fig. 1). Comparison of forecast tracks of tropical cyclone JASON for the period 0000 UTC, 10 February 1987 to 0000 UTC 12 February 1987. The *observed 'best track'* is denoted by O; the CLIPER scheme by C; the BMRC NWP model by M; and the optimal linear combination of C and M by L, for 12-, 24-, 36- and 48-h forecasts.

RESULTS: The 24 and 48 h position error has been reduced by the combination (COMB). The errors were 15% less than the next best scheme (CLIPPER) and 17% less than the next best scheme (based on regressions) at 48 h. A standard t-test (with the number of degrees of freedom adjusted to allow for serial correlation between forecasts) rejected the null hypothesis at the 1% level or lower that the combination had an identical mean error with any of the other techniques. The error of the combination was lower than all the other schemes in every tropical cyclone season to which it was applied.

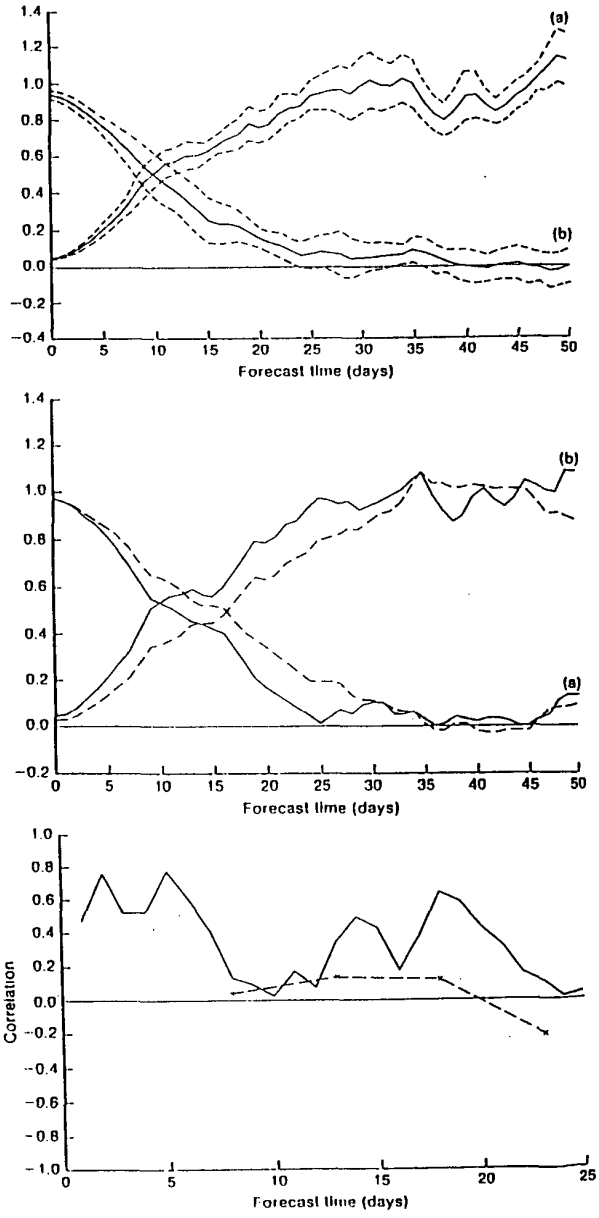


Fig. 5 (J.M. Murphy 1988, Fig. 3, 4 and 7). (1) SPREAD: (a) Ratio of variance of forecast probability density (fpd) to model climate variance; (b) squared anomaly correlation between centroid of fpd and all individual forecasts within fpd. Values are taken from the 'perfect model' random perturbation ensemble forecast statistics with associated standard deviations. (2) MEAN: (a) Daily perfect model forecast anomaly correlation and (b) normalized error variance: individual forecasts from unperturbed initial state (full) and ensemble mean forecast (dashed). (3) SPREAD AND SKILL: Correlation between 'perfect model' error variance of ensemble mean forecast and ensemble variance: daily fields (full) and 15-day-mean fields (dashed).

4. MEDIUM AND EXTENDED RANGE PREDICTION

ENSEMBLE AVERAGE FORECASTS AND PREDICTABILITY: The preceeding subsections on short-term weather and short-term climate predictions are mainly concerned with combinations of two independent forecast schemes leading to an improvement of the skill. In medium and extended range forecasting a different combination approach is being applied: ensemble forecasts. The aim is twofold: to improve the forecast and to predict the skill. These ensemble forecasts are basically Monte Carlo type predictions of a set of randomly perturbed initial conditions or using lagged forecasts. The lagged ensemble is generated by perturbations due to the forecast model predicting from (successively older) initial states forward to the time of the latest analysis. In the weather state space these predictions represent an ensemble of initial states evolving along diverging time trajectories (Leith 1974, see also Hayashi 1986, Seidman 1989). The associated measures used in analysing ensemble forecasts are the first and second moments, the mean and the spread or variance, which are sufficient to describe a Gaussian distribution (see for example J.M. Murphy 1988, 1990). Furthermore, ensemble probabilities are deduced to extract more information and to provide probability forecasts (A.H. Murphy 1977).

* ENSEMBLE MEAN AND SPREAD: While the ensemble mean forecasts lead to skill improvement compared with the mean skill of the individual forecasts (extending, hopefully, into the long-range), the spread of an ensemble distribution can, in principle, give an a priori indication of the forecast skill which is one approach towards skill prediction. For a 'perfect model' situation some results are shown in FIGURE 5.

5. CONCLUSION

This review has been prepared to classify the combination forecasts applied in meteorology. In both short-term weather and climate predictions statistical methods appear to be valuable. In short-term weather prediction it is the combination of deterministic and stochastic schemes covering the non-linear developments and the local characteristics of weather and climate.

In long-range weather forecasts the combination of empirical-statistical techniques is successfully used. The combination with deterministic forecasts may be a future development once extended range predictions are sufficiently informative. Finally, the medium and extended range forecast combinations are basically Monte Carlo type predictions to gain information about the forecast confidence and to improve the forecast skill by appropriate ensemble averaging or generating probabilistic forecasts.

Acknowledgement: Thanks are due to Ms. M. Scholz for preparing the manuscript.

REFERENCES (LONG-RANGE FORECASTING)

Barnett, T.P., and R.W. Preisendorfer, 1978: Multifield analog prediction of short term climate fluctuations using a climate state vector. J. Atmos. Sci., **35**, 1771-1787.

Fraedrich, K. and N. Smith, 1989: Combining predictive schemes in long-range forecasting. J. Climate, **2**, 291-294.

Gilman, D.L., 1985: Long range forecasting: The present and the future. Bull. Amer. Meteor. Soc., **114**, 159-164.

Livezey, R.E. and A.G. Barnston, 1988: An operational multifield analog/antilog prediction system for United States seasonal temperatures. J. Geophys. Res., **93**, 10953-10974.

Livezey, R.E., A.G. Barnston and B.K. Neumeister, 1990: Mixed analog/persistence prediction of seasonal mean temperatures for the USA. Intern. J. Climatol., **10**, 329-340.

Thompson, P.D., 1977: How to improve accuracy by combining independent forecasts. Mon. Wea. Rev., **114**, 228-229.

Toth, Z.,1989: Long range weather forecasting using an analog approach. J. Climate, **2**, 594-607.

REFERENCES (SHORT-TERM FORECASTING)

Fraedrich, K. and L.M. Leslie, 1987: Combining predictive schemes in short-term forecasting. Mon. Wea. Rev., **115**, 1640-1644.

Fraedrich, K. and L.M. Leslie, 1988: A minimal model for the short-term prediction of rainfall in the tropics. Weather and Forecasting, **3**, 243-246.

Hope, J.R. and C.J. Neumann, 1970: An operational technique for relating the movement of existing tropical cyclones to past tracks. Mon. Wea. Rev., **98**, 925-933.

Keenan, T.D., 1986: Forecasting tropical cyclone motion using a discriminant analysis procedure. Mon. Wea. Rev., **114**, 434-441.

Leslie, L.M. and K. Fraedrich, 1990: Reduction of tropical cyclone position errors using an optimal combination of independent forecasts. Weather and Forecasting, **5**, 158-161.

Murphy, A.H., 1986: Comparative evaluation of catagorical and probabilistic forecasts: two alternatives to the traditional approach. Mon. Wea. Rev., **114**, 245-249.

Neumann, C.J. and J.M. Plessier, 1981: Models for the prediction of tropical cyclone motion over the North Atlantic: an operational evaluation. Mon. Wea. Rev., **109**, 522-538.

REFERENCES (MEDIUM-EXTENDED RANGE FORECASTING)

Hayashi, Y., 1986: Statistical interpretation of ensemble-time mean predictability. J. Meteor. Soc. Jap., **67**, 164-181.

Leith, C.C., 1974: Theoretical skill of Monte Carlo forecasts. Mon. Wea. Rev., **102**, 409-418.

Murphy, A.H., 1977: The value of climatological, categorical and probability forecasts in the cost-loss ratio situation. Mon. Wea. Rev., **105**, 803-816.

Murphy, J.M., 1988: The impact of ensemble forecasts on predictability. Quart. J. R. Meteorol. Soc., **114**, 463-493.

Murphy, J.M., 1990: Assessment of the practical utility of extended range ensemble forecasts. Quart. J.R. Meteorol. Soc., **116**, 89-125.

Seidman, A.N., 1981: Averaging techniques in long-range weather forecasting. Mon. Wea. Rev., **109**, 1367-1379.

Zhunkovskii, E.E. and T.M. Brunova, 1978: Effective complexing of alternative forecasts (engl. translation). Soviet Meteorology and Hydrology, **5**, 11-18.

COMBINING VERSUS CHOOSING: IMPLICATIONS FOR FORECAST EVALUATION

Robert T. Clemen[1], Allan H. Murphy[2], and Robert L. Winkler[3]

[1] College of Business Administration, University of Oregon, Eugene, Oregon 97403, U.S.A.
[2] Department of Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331, U.S.A.
[3] Fuqua School of Business, Duke University, Durham, North Carolina 27706, U.S.A.

1. Introduction

In many weather forecasting situations, it is possible to produce the forecasts of interest by two or more different methods. The set of available methods may consist of alternative models (numerical and/or statistical), different forecasters, or a mixture of models and forecasters. In such situations, the basic problem is traditionally viewed as a problem of identifying - and subsequently choosing - the method that produces the best forecasts.

When two or more forecasting methods are available it is also possible to combine the individual forecasts, and methodological and practical issues related to combining have been studied extensively in various contexts during the last decade (see Clemen, 1989). Moreover, the possibility of combining forecasts produced by different methods has been investigated in various weather forecasting situations, with some notable successes in recent years (Fraedrich and Leslie, 1987, 1988; Fraedrich and Smith, 1989). In this context, the basic problem can be viewed as a problem of determining whether or not combined forecasts can be produced that are better than those provided by the individual methods.

This paper is concerned with evaluation methodology in the contexts of the choosing and combining scenarios. The phrase "evaluation methodology" is meant to include descriptions of relationships among forecasts and observations based on their joint distributions as well as measures that characterize forecasting performance. Specifically, we focus here on two basic evaluation tools: (a) the sufficiency relation (DeGroot and Eriksson, 1985; DeGroot and Fienberg, 1982, 1983, 1986) and (b) the concept of extraneous experts (Clemen, 1985; Clemen and Guerard, 1989). The sufficiency relation embodies the conditions under which one set of forecasts can be unambiguously judged to be better than another set of forecasts, whereas the concept of extraneous experts describes the conditions under which one set of forecasts contains all of the information (regarding the events of interest) possessed by another set of forecasts.

The objective of this paper is to investigate - and compare - the use of these evaluation tools in the contexts of the choosing and combining scenarios. Basic evaluation concepts and methods, including the sufficiency relation and the concept of extraneous experts, are briefly discussed in Section 2. Hypothetical examples are considered in Section 3 to illustrate the differences between the use of these tools in the choosing and combining scenarios. Section 4 summarizes some results of representative choosing and combining studies in which either the sufficiency relation or the concept of extraneous experts was used to evaluate operational weather forecasts. Section 5 contains a discussion of the implications of these results and some concluding remarks.

2. Forecast Evaluation: Basic Concepts and Methods

a. Forecast evaluation in the choosing problem

When the basic problem is viewed as a problem of choosing among available forecasting methods (hereafter referred to as the *choosing scenario),* the objective is to choose the best forecasting method. In this context, the primary role of forecast evaluation is to identify the method that produces the best forecasts. To realize this objective, it is necessary to define what is meant by the phrase "best forecasts."

For an individual forecasting method, forecast quality is fully described by the joint distribution of forecasts and observations (Murphy and Winkler, 1987). Moreover, conditional and marginal distributions associated with factorizations of this joint distribution characterize specific aspects of forecast quality. Thus, forecast quality is multidimensional in nature, and one or two measures of overall performance (e.g., a correlation coefficient or a skill score) generally cannot completely describe the quality of a set of forecasts.

Analogously, when the relative quality of two (or more) forecasting methods is of interest, comparison of the values of one or two overall performance measures is inadequate. It is necessary to consider the basic characteristics of forecast quality described by the two joint distributions and the conditional and marginal distributions associated with their respective factorizations. The sufficiency relation provides a means of comparing alternative forecasting methods that is consistent with the multidimensional nature of forecast quality, and this relation is defined in Section 2c.

b. Forecast evaluation in the combining problem

When the basic problem is viewed as a problem of combining the forecasts produced by two or more forecasting methods (hereafter referred to as the *combining scenario), the objective is to obtain combined forecasts whose quality* exceeds that of the forecasts produced by any individual method. Forecast evaluation plays two roles in this context. First, it is necessary to ascertain whether or not the methods in question contain independent information regarding future values of the variable of interest, in which case the quality of the combined forecasts necessarily equals or exceeds that of any individual method. Second, the quality of the combined forecasts must be assessed.

Combining forecasts produced by two forecasting methods can only lead to better forecasts when each method contains information regarding the variable of interest that is not included in the other method. Therefore, forecast evaluation in the context of the combining scenario is necessarily concerned with the independent information content of the alternative forecasting methods. To determine whether or not these methods contain independent information, the full multivariate distribution involving both types of forecasts and the observations must be examined. Thus, the fundamental difference between forecast evaluation in the choosing and combining scenarios is that it involves bivariate distributions in the choosing scenario and multivariate distributions in the combining scenario.

c. The sufficiency relation

The significance of the sufficiency relation in the context of forecast evaluation resides in its ability to order forecasting methods in terms of their relative quality and value. If one forecasting method can be shown to be sufficient for another forecasting method, then the former's forecasts are necessarily of higher quality in all respects - and of greater value to all users - than the latter's forecasts. The use of the sufficiency relation to evaluate weather forecasts has been investigated recently by Ehrendorfer and Murphy (1988, 1992), Krzysztofowicz and Long (1991a,b), and Murphy and Ye (1990) among others.

Consider two forecasting methods (or forecasters) A and B and suppose that these methods produce probabilistic forecasts f_A and f_B for a dichotomous variable X (i.e., X = 0 or 1). Thus, $f_A = P_A(X=1)$ and $f_B = P_B(X=1)$, where X = 1 if the event occurs. We will assume that these forecasts are well-calibrated (or reliable) in the sense that $P(X=1|f_A) = f_A$ and $P(X=1|f_B) = f_B$. Let $v_A(f)$ and $v_B(g)$ denote functions characterizing the refinement (or sharpness) of A's and B's forecasts, where $v_A(f) = P(f_A=f)$ and $v_B(g) = P(f_B=g)$, where f and g denote particular values from a finite set of permissible probabilities F (i.e., f,g e F, where $0 \le f,g \le 1$).

Given these definitions, method A is sufficient for method B if a stochastic transformation h(f|g) exists such that

$$\Sigma_f h(g|f)v_A(f) = v_B(g) \quad \text{for all } g \text{ e } F \qquad (1)$$

and

$$\Sigma_f h(g|f)fv_A(f) = gv_B(g) \quad \text{for all } g \text{ e } F. \qquad (2)$$

The function h(g|f) satisfies the conditions for a stochastic transformation if $0 \le h(g|f) \le 1$ and $\Sigma_g h(g|f) = 1$ for all f e F. Equation (1) embodies the main result; equation (2) ensures that if A's forecasts are well-calibrated, then the transformed forecasts will also be well-calibrated.

Equation (1) can be interpreted by noting that the stochastic transformation represents an auxiliary randomization that introduces noise into A's forecasts. The new forecasts produced by this process are distributionally equivalent to B's forecasts. Thus, B's forecasts contain greater uncertainty than A's forecasts. It should also be noted that the sufficiency relation induces only a quasi-order on the two sets of forecasts, in the sense that A's forecasts may be sufficient for B's forecasts, B's forecasts may be sufficient for A's forecasts, or A's and B's forecasts may be insufficient for each other.

d. The concept of extraneous experts

In the combining scenario, the trivariate distribution $P(f_A,f_B,x)$ represents the appropriate framework within which to evaluate the forecasters. This distribution contains information about the forecasts produced (individually) by methods A and B, as well as information about the relationships among A's forecasts, B's forecasts, and the observations. This distribution can be factored into univariate conditional distributions and a joint marginal distribution as follows:

$$P(f_A,f_B,x) = P(x|f_A,f_B)P(f_A,f_B), \qquad (3)$$

where $P(x|f_A,f_B)$ represents conditional distributions of the observations given A's and B's forecasts and $P(f_A,f_B)$ represents the joint (marginal) distribution of A's and B's forecasts. Since the conditional distributions $P(x|f_A,f_B)$ describe the relationship between the forecasts, as it relates to the observations, it seems natural to focus attention on these distributions in this context.

Recall that we are concerned here with the question of whether or not forecasting methods (or forecasters) provide independent information regarding future values of the variable of interest (denoted here by x). If a forecaster does not provide any such information, then he/she is referred to as an *extraneous expert* (Clemen, 1985). With regard to the conditional distributions $p(x|f_A,f_B)$, forecaster A is extraneous if

$$P(x|f_A,f_B) = P(x|f_B). \qquad (4)$$

Likewise, forecaster B is extraneous if

$$P(x|f_A,f_B) = P(x|f_A). \quad (5)$$

Expressions (4) and (5) describe conditions of conditional independence in this context; namely, x is conditionally independent of f_A in (4) and x is conditionally independent of f_B in (5). Thus, A is extraneous when x is independent of f_A (given f_B) and B is extraneous when x is independent of f_B (given f_A). If such conditions hold, then forecasters A and B, respectively, contain no independent information regarding future values of x. Forecasts based on a combination of f_A and f_B will be of higher quality than f_A or f_B alone only if neither A's nor B's forecasts are extraneous.

Evaluation methods based on these concepts have been used to study relationships between objective and subjective weather forecasts (Clemen and Murphy, 1986a; Murphy et al., 1988). Specifically, these studies have investigated whether either the objective or subjective forecasts are extraneous, in the sense that they provide no independent information regarding the weather events of interest. An example of the results obtained in these studies is presented in Section 4b of this paper.

3. Evaluation in Choosing and Combining Scenarios: Some Hypothetical Examples

In this section we present three simple hypothetical examples to illustrate the differences between the use of the sufficiency relation and the concept of extraneous experts in the choosing and combining scenarios. The basic difference between forecast evaluation in the choosing and combining scenarios can now be described by considering the use of the sufficiency relation. Suppose that the forecasts produced by methods A and B have been evaluated and it has been shown that A's forecasts are sufficient for B's forecasts. If the problem is viewed as a problem of choosing between these methods, then A is clearly the method of choice. A's forecasts are of higher quality - and greater value - than B's forecasts.

However, it is important to recognize that this result does not directly address the combining problem. Although B's forecasts are inferior in quality to A's forecasts, the former may still contain information regarding future values of the variable of interest that is not included in the latter. In this case, it may be possible to find a forecasting method C, whose forecasts are (a) a combination of A's and B's forecasts and (b) sufficient for A's (and B's) forecasts. If A's (or B's) forecasts are sufficient for the combined forecasts C, then clearly B's (or A's) forecasts can be ignored. Clemen (1985) uses the term "extraneous expert" to describe such a forecaster (or method) and also defines extraneousness equivalently in terms of conditional independence. In summary, the basic issue in the combining scenario relates to independent information content, whereas the basic issue in the choosing scenario relates to relative forecast quality.

a. Example 1

Consider two forecasters A and B, each of whom is well-calibrated and uses only probabilities of 0.1 and 0.9. Suppose that their refinement functions are $v_A(0.1) = P(f_A = 0.1) = 0.5$, $v_A(0.9) = P(f_A = 0.9) = 0.5$, $v_B(0.1) = P(f_B = 0.1) = 0.5$, and $v_B(0.9) = P(f_B = 0.9) = 0.5$. Thus, A and B are exchangeable forecasters. Moreover, A and B are sufficient for each other. The stochastic transformation is $h(g|f) = 1$ if $g = f$ and $h(g|f) = 0$ otherwise. That is, $h(f_B = 0.1|f_A = 0.1) = 1$, $h(f_B = 0.9|f_A = 0.1) = 0$, $h(f_B = 0.1|f_A = 0.9) = 0$, and $h(f_B = 0.9|f_A = 0.9) = 1$. This transformation simply translates one forecaster into the other forecaster. A decision maker faced with choosing between these two forecasters would be indifferent.

However, suppose that these forecasts possess the joint distribution $p(f_A,f_B)$ described in Table 1. That is, $p(f_A,f_B)$ is

Table 1. Joint distribution $p(f_A,f_B)$ for Example 1.

		Forecaster B	
		$f_B = 0.1$	$f_B = 0.9$
Forecaster A	$f_A = 0.1$	0.4	0.1
	$f_A = 0.9$	0.1	0.4

such that forecasters A and B make the same forecast 80% of the time (40% of the time they both say "0.4" and 40% of the time they both say "0.9"). Their forecasts differ only 20% of the time (10% of the time A says "0.1" and B says "0.9" and 10% of the time A says "0.9" and B says "0.1").

Moreover, suppose that the conditional distribution of the event X = 1 (e.g., the occurrence of precipitation) given the two forecasts, $p(X = 1|f_A,f_B)$, is as described in Table 2. This

Table 2. Conditional distribution $P(X = 1|f_A,f_B)$ for Example 1.

		Forecaster B	
		$f_B = 0.1$	$f_B = 0.9$
Forecaster A	$f_A = 0.1$	0	0.5
	$f_A = 0.9$	0.5	1

conditional distribution indicates that if both A and B say "0.9," then the decision maker can be sure that precipitation will occur. That is, $P(X = 1|f_A = 0.9, f_B = 0.9) = 1$. Analogously, if both forecasters say "0.1," the decision maker can be sure that precipitation will not occur [$P(X = 1|f_A = 0.1, f_B = 0.1) = 0$]. When A and B disagree, the decision maker's posterior probability of precipitation is 0.5 [$P(X =$

$1|f_A = 0.1, f_B = 0.9) = P(X = 1|f_A = 0.9, f_B = 0.1) = 0.5]$.

If A's and B's forecasts are combined using this information (the combined forecast is denoted by "C"), the combined forecasts would be well-calibrated and would possess the following refinement function: $v_C(0) = P(f_C = 0) = 0.40$, $v_C(0.5) = P(f_C = 0.5) = 0.20$, and $v_C(1) = P(f_C = 1) = 0.40$. That is, the combined forecast would indicate that the probability of precipitation is zero 40% of the time, one 40% of the time, and 0.50 20% of the time.

It is quite easy to show that the combined forecast C is sufficient for either A or B. The stochastic transformation has the following form: $h(f_A = 0.1|f_C = 0) = 1$, $h(f_A = 0.9|f_C = 0) = 0$, $h(f_A = 0.1|f_C = 0.5) = 0.5$, $h(f_A = 0.9|f_C = 0.5) = 0.5$, $h(f_A = 0.1|f_C = 1) = 0$, and $h(f_A = 0.9|f_C = 1) = 1$. That is, if $f_C = 0$, set $f_A = 0.1$, and if $f_C = 1$, set $f_A = 0.9$. If $f_C = 0.5$, toss a fair coin. If heads occurs, set $f_A = 0.1$, and if tails occurs, set $f_A = 0.9$. Identical expressions and instructions can be written for the stochastic transformation relating f_B and f_C. However, neither B nor A is sufficient for C, indicating that *neither A nor B is extraneous.*

This example illustrates a situation in which the trivariate distribution *contains information that is useful in combining* forecasts but doesn't help in choosing between the two forecasters. Moreover, *even though forecasters A and B are* exchangeable in the sense that it doesn't matter to the decision maker which forecaster he/she consults, neither forecaster is extraneous.

b. Example 2

Consider the same forecasters as in Example 1. However, now suppose that they possess the joint distribution $P(f_A,f_B)$ described in Table 3. In order for the combined forecast (C) to remain well-calibrated, if both forecasters say "0.1," C must also say "0.1." Analogously, C must say "0.9" when A and B say "0.9." As expected, these two forecasters are not only exchangeable, they are fully equivalent since they just parrot each other. In fact, it is easy to see that A (or B) is sufficient for C, indicating that B (or A) is extraneous.

Table 3. The joint distribution $P(f_A,f_B)$ in Example 2.

		Forecaster B	
		$f_B = 0.1$	$f_B = 0.9$
Forecaster A	$f_A = 0.1$	0.5	0
	$f_A = 0.9$	0	0.5

In contrast to Example 1, this example demonstrates extraneous forecasts. Once the decision maker has consulted either A or B, then the other is extraneous. Moreover, it is easy to see that the combined forecast C must be equivalent to both A and B. The differences between Examples 1 and 2 are particularly intriguing because the two joint distributions are quite similar (cf. Tables 1 and 3).

c. Example 3

Now consider a situation in which the two forecasters are not exchangeable. In this example forecaster A uses probabilities 0.4 and 0.6, whereas forecaster B uses probabilities of 0.1 and 0.9 (the same as in the two previous examples). Once again, it is assumed that both forecasters are well-calibrated.

Suppose that their respective refinement functions are as follows: $v_A(0.4) = P(f_A = 0.4) = 0.5$, $v_A(0.6) = P(f_A = 0.6) = 0.5$, $v_B(0.1) = P(f_B = 0.1) = 0.5$, and $v_B(0.9) = P(f_B = 0.9) = 0.5$. It is relatively easy to show that forecaster B is sufficient for forecaster A. The stochastic transformation is as follows: $h(f_A = 0.4|f_B = 0.1) = 5/8$, $h(f_A = 0.6|f_B = 0.1) = 3/8$, $h(f_A = 0.4|f_B = 0.9) = 3/8$, and $h(f_A = 0.6|f_B = 0.9) = 5/8$. That is, when B uses "0.1," then with probability 5/8 the transformed forecast (A) is "0.4" and with probability 3/8 the transformed forecast (A) is "0.6." (A fair 8-sided die, with five "red" sides and three "green" sides, could be used to determine which probability should be used by forecaster A.) The probabilities 5/8 and 3/8 are reversed when B uses "0.9."

Now suppose that the forecasts possess the joint distribution $P(f_A,f_B)$ described in Table 4. These forecasters are independent, at least in terms of this joint distribution. That is, $P(f_A = 0.4,f_B = 0.1) = P(f_A = 0.4)P(f_B = 0.1)$, etc.

Table 4. Joint distribution $P(f_A,f_B)$ for Example 3.

		Forecaster B	
		$f_B = 0.1$	$f_B = 0.9$
Forecaster A	$f_A = 0.4$	0.25	0.25
	$f_A = 0.6$	0.25	0.25

Further, suppose that the conditional probability of precipitation given the two forecasts, $P(X = 1|f_A,f_B)$, is as described in Table 5. Thus, if the forecasters "agree" (i.e,

Table 5. Conditional distribution $P(X = 1|f_A,f_B)$ for Example 3.

		Forecaster B	
		$f_B = 0.1$	$f_B = 0.9$
Forecaster A	$f_A = 0.4$	0	0.8
	$f_A = 0.6$	0.2	1

f_A = 0.4 and f_B = 0.1 or f_A = 0.6 and f_B = 0.9), the decision maker is sure about the outcome (as in Example 1). However, if the forecasters "disagree" (f_A = 0.4 and f_B = 0.9 or f_A = 0.6 and f_B = 0.1), the decision maker adopts a slightly modified version of B's forecast.

As a result, the combined forecast (C), which is well-calibrated, has the following refinement function: $v_C(0) = P(f_C = 0) = 0.25$, $v_C(0.2) = P(f_C = 0.2) = 0.25$, $v_C(0.8) = P(f_C = 0.8) = 0.25$, and $v_C(1) = P(f_C = 1) = 0.25$. It is easy to show that C is sufficient for B (and hence is sufficient for A). The stochastic transformation has the following form: $h(f_B = 0.1|f_C = 0) = 1$, $h(f_B = 0.9|f_C = 0) = 0$. $h(f_B = 0.1|f_C = 0.2) = 1$, $h(f_B = 0.9|f_C = 0.2) = 0$, $h(f_B = 0.1|f_C = 0.8) = 0$, $h(f_B = 0.9|f_C = 0.8) = 1$, $h(f_B = 0.1|f_C = 1) = 0$, and $h(f_B = 0.9|f_C = 1) = 1$. That is, if C uses "0" or "0.2," then set the transformed forecast (B) to "0.1." Otherwise (i.e., C = "0.8" or "1"), set the transformed forecast (B) to "0.9." However, it can be shown that B is not sufficient for C, indicating that A is not extraneous.

This example is interesting because method A is dominated by method B in the choosing scenario. However, from the perspective of the combining scenario, A contains independent information (i.e., information not contained in B's forecasts). Hence, the combined forecast C is sufficient for (and thus better than) B. In summary, B is sufficient for A but A is not extraneous

4. Evaluation in the Choosing and Combining Scenarios: Some Real-World Results

a. Choosing scenario

It is evident from Section 2c that the sufficiency relation is a potentially powerful tool in the context of the choosing scenario, since it provides a means of determining unambiguously whether forecasting method A is better - in all important respects - than forecasting method B (or vice versa). However, it is only recently that this relation has been applied to the problem of comparative evaluation of weather or climate forecasts. Moreover, in view of the fact that the sufficiency relation is a quasi-order (see Section 2c), its applicability in such real-world contexts warrants careful investigation.

Ehrendorfer and Murphy (1988) introduced the sufficiency relation into the meteorological literature. They studied its applicability in two-event situations involving primitive probabilistic (i.e., calibrated nonprobabilistic) forecasts and identified the conditions - on various characteristics of forecast quality - under which method A is sufficient for method B (or vice versa). More recently, Ehrendorfer and Murphy (1992) undertook a similar analysis involving prototypical climate forecasting systems. These studies have identified potential tradeoffs between various characteristics of quality and explored the likelihood of encountering conditions in which the two forecasting methods are insufficient for each other.

Krzysztofowicz and Long (1991a) formulated a measure of sufficiency, the so-called forecast sufficiency characteristic (FSC), based on a theorem advanced by Blackwell and Girshick (1954). The FSC provides a means of assessing the sufficiency of method A for method B (or vice versa) without the need to determine the existence of a stochastic transformation. They used this measure to investigate the sufficiency of objective and subjective probability of precipitation (PoP) forecasts (see also Murphy and Ye, 1990). In a companion study (Krzysztofowicz and Long, 1991b), the authors modeled the PoP forecasts with beta distributions and then used the FSC to compare the two types of forecasts. Not surprisingly, the use of beta models greatly reduced the likelihood of encountering conditions of insufficiency in this context.

Recently, Krzysztofowicz (1992) derived a measure of skill for nonprobabilistic forecasts of continuous predictands - the so-called Bayesian correlation score (BCS) - based on the theory of sufficient comparisons of experiments. The BCS is specified in terms of the parameters of a normal-linear model, which combines information regarding the climatology of the predictand with characteristics of forecast quality. Under these modeling assumptions, the BCS can be used to compare alternative forecasts of the same predictand or (in a limited sense) different predictands, and it orders alternative forecasting systems in terms of the value of their forecasts to users (in a manner similar to the sufficiency relation).

Regardless of the range of applicability of the sufficiency relation, it is desirable and useful in the context of comparative evaluation of forecasting methods to determine - and compare - the various characteristics of forecast quality. The appropriate framework within which to undertake this evaluation is the joint distribution of forecasts and observations, since this distribution contains all of the non-time-dependent information relevant to forecast quality (Murphy and Winkler, 1987). Moreover, the information contained in the joint distribution is more accessible when this distribution is factored into conditional and marginal distributions. The joint, conditional, and marginal distributions, together with summary measures of these distributions and various performance measures and their decompositions, provide a methodological framework for *diagnostic verification* - an approach to absolute and comparative verification that identifies the basic strengths and weaknesses in forecasts and provides insights into ways of improving forecasting performance. This approach recognizes the multidimensional nature of verification problems (Murphy, 1991) and reveals that traditional approaches to forecast verification - which generally involve the computation of one or two overall measures of forecasting performance (e.g., a skill score or a correlation coefficient) - are seriously deficient. Diagnostic verification methods have recently been applied to the comparative evaluation of objective and subjective weather forecasts (Murphy et al., 1989; Murphy and Winkler, 1992).

b. Combining scenario

Clemen and Murphy (1986a) (hereafter CM86) investigated the interrelationships between objective and subjective probability of precipitation (PoP) forecasts using the methodology of extraneous experts. In particular, they addressed two questions: (a) Do the subjective forecasts contain information not included in the objective PoP forecasts? (b) Do the subjective PoP forecasts make full use of the information contained in the objective PoP forecasts? To answer these questions they analyzed the joint calibration function $p(x|f_o,f_s)$ for various matched samples of objective forecasts (f_o), subjective forecasts (f_s), and observations (x). The joint calibration function indicates the relative frequency of precipitation as a function of both f_o and f_s.

To obtain qualitative insight into the relationships of interest, CM86 first examined the conditional calibration functions, $p(x|f_o)$ and $p(x|f_s)$, conditioned on various values of f_s and f_o, respectively. Typical results are presented in Figures 1 and 2. Figure 1 indicates that the conditional

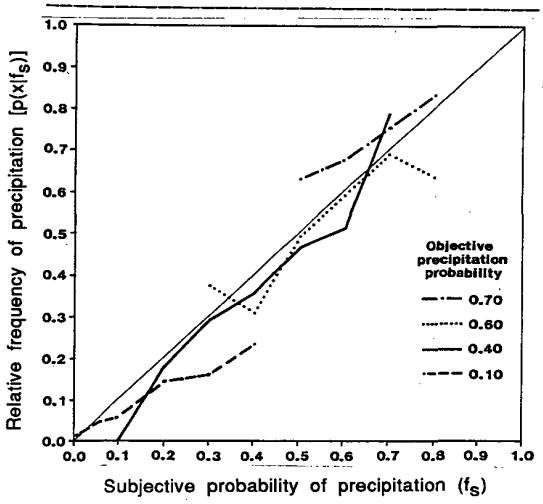


Figure 1. The conditional calibration functions for the subjective forecasts, $p(x|f_s)$, for selected values of the objective forecast (f_o).

calibration function for f_s increases as f_s increases. Thus, $p(x|f_o,f_s) \neq p(x|f_o)$, and f_s contains information not included in f_o (f_s is not extraneous).

The results in Figure 2 differ at least qualitatively from the results in Figure 1. The conditional calibration functions for f_o are quite flat, indicating that f_o contains relatively little information not included in f_s (recall that, in general, f_s is produced after f_o becomes available to the forecaster). This result suggests that the subjective forecasts contain most of the information included in the objective forecasts.

To assess these interrelationships in a quantitative manner, CM86 performed Chi-square tests on the basic joint calibration function $p(x|f_o,f_s)$ using a series of 2x2 contingency tables. They found strong evidence to support an affirmative answer to the first question - the subjective

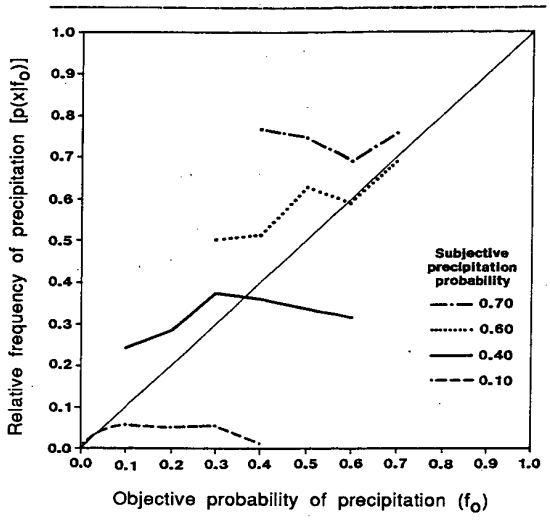


Figure 2. The conditional calibration functions for the objective forecasts, $p(x|f_o)$, for selected values of the subjective forecast (f_s).

forecasts do indeed contain information not included in the objective forecasts for most combinations of season, geographical area, and lead time. The quantitative results related to the second question are mixed. In the aggregate these results suggest that the subjective forecasts are not making complete use of the objective forecasts, especially for longer lead times. The subjective forecasts make the most complete use of the objective forecasts in the warm season and for the shortest lead times.

Overall, the results presented by CM86 indicate that neither f_s nor f_o is extraneous, in the sense that a decision maker must in general consult both forecasts in order to obtain the maximum amount of information (regarding the occurrence of precipitation). This fact suggests that a forecast produced by combining f_o and f_s might be better than either forecast alone. Clemen and Murphy (1986b), in a companion paper, investigated this possibility and found that modest improvements in forecasting performance could be achieved through averaging the objective and subjective forecasts. Moreover, Clemen and Winkler (1987) studied a variety of combining methods for a similar set of PoP forecasts.

In an analogous study involving nonprobabilistic forecasts of a continuous variable, Murphy et al. (1988) investigated the interrelationships between objective and subjective temperature forecasts using the methodology of extraneous experts. They obtained similar results; namely, the subjective forecasts contain information not included in the objective forecasts and the latter generally contain information not included in the former. In particular, they found that the incremental information content in the subjective forecasts generally substantially exceeds the incremental information content in the objective forecasts.

Once again, these results indicate that neither forecast is extraneous and that forecasting performance might be improved by combining the objective and subjective forecasts.

5. Discussion and Conclusion

Careful evaluation of forecasts is required whether the objective is to choose among alternative forecasting methods (or forecasters) or to use the available methods in the most effective manner by combining their forecasts. At a general level, the questions asked in any evaluation task are similar. How accurate are the forecasts? Do the forecasts possess a systematic bias? What are the implications of using the forecasts as a basis for decision making? However, as we have shown, the two tasks - choosing and combining - require different evaluation procedures. We have focused here on the use of the related concepts of sufficiency and extraneousness; sufficiency is an appropriate evaluation tool to use in the choosing scenario, but if the objective is to combine forecasts, the stronger concept of extraneousness is needed to determine whether any of the available forecasting methods can be eliminated from consideration.

The recent research that we have undertaken related to forecast evaluation has both methodological and practical implications. From a methodological point of view, more coherent and complete approaches to forecast evaluation are needed. As we have argued here and elsewhere, use of one or two overall evaluation measures (e.g., a measure of accuracy such as the mean square error, a skill score) leads to an impoverished understanding of the characteristics of a set of forecasts. In both the choosing and combining scenarios, a full examination of the joint distribution of forecasts and observations can yield a much more complete and insightful evaluation of the forecasts of interest.

The examples and results presented in this paper, together with the accompanying discussion, suggest that approaches to evaluation that employ the concepts of sufficiency and extraneousness can lead to a better understanding of the quality of forecasts, more informed choices among alternative forecasting methods, and a framework for studying the potential benefits (in terms of forecast quality) that can be achieved by combining forecasts. In particular, the concept of extraneousness permits the framing of questions regarding the complementarity of forecasts, questions that have not been widely addressed in the meteorological community.

While sufficiency and extraneousness are useful tools in the choosing and combining scenarios, further research is required to improve our understanding of these tools (and their relationships) and to develop new tools as required. For example, sufficiency has been related to other concepts (e.g., second-degree stochastic dominance; see DeGroot and Eriksson, 1985). The implications of such relationships, and possible extensions thereof, may lead to greater insight into the use of these tools as well as to the possible development of new approaches to evaluation. Because the sufficiency relation imposes only a quasi-order on forecasting methods (or forecasters), an important objective of future research is to develop evaluation tools beyond sufficiency and extraneousness that are still consistent with the principle of examining the entire joint distribution. Some tools may be more appropriate for analyzing the bivariate distributions in the choosing scenario, whereas other tools may be required for examining the multivariate distributions in the combining scenario.

Acknowledgments

This work was supported in part by the National Science Foundation under Grants SES-9022573, SES-9022616, and SES-9106440.

References

Blackwell, D., and M.A. Girshick, 1954: *Theory of Games and Statistical Decisions.* New York, Wiley, 355 pp.

Clemen, R.T., 1985: Extraneous expert information. *Journal of Forecasting,* **4**, 329-348.

Clemen, R.T., 1989: Combining forecasts: a review and annotated bibliography. *International Journal of Forecasting,* **5**, 559-583.

Clemen, R.T., and J. Guerard, 1989: Econometric GNP forecasts: incremental information relative to naive extrapolation. *International Journal of Forecasting,* **5**, 417-426.

Clemen, R.T., and A.H. Murphy, 1986a: Objective and subjective precipitation probability forecasts: statistical analysis of some interrelationships. *Weather and Forecasting,* **1**, 56-65.

Clemen, R.T., and A.H. Murphy, 1986b: Objective and subjective precipitation probability forecasts: some methods for improving forecast quality. *Weather and Forecasting,* **1**, 213-218.

Clemen, R.T., and R.L. Winkler, 1987: Calibrating and combining precipitation probability forecasts. *Probability and Bayesian Statistics* (R. Viertl, Editor). London, Plenum, pp. 97-110.

DeGroot, M.H., and E.A. Eriksson, 1985: Probability forecasting, stochastic dominance, and the Lorenz curve. *Bayesian Statistics 2* (J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith, Editors). Amsterdam, North-Holland, pp. 99-118.

DeGroot, M.H., and S.E. Fienberg, 1982: Assessing probability assessors: calibration and refinement. *Statistical Decision Theory and Related Topics III,* Volume 1 (S.S. Gupta and J.O. Berger, Editors). New York, Academic Press, pp. 291-314.

DeGroot, M.H., and S.E. Fienberg, 1983: The comparison and evaluation of forecasters. *The Statistician,* **32**, 12-22.

DeGroot, M.H., and S.E. Fienberg, 1986: Comparing probability forecasters: basic binary concepts and multivariate extensions. *Bayesian Inference and Decision Techniques* (P. Goel and A. Zellner, Editors). Amsterdam, North-Holland, pp. 247-264.

Ehrendorfer, M., and A.H. Murphy, 1988: Comparative evaluation of weather forecasting systems: sufficiency, quality, and accuracy. *Monthly Weather Review,* **116**, 1757-1770.

Ehrendorfer, M., and A.H. Murphy, 1992: Evaluation of prototypical climate forecasts: the sufficiency relation. *Journal of Climate,* **5**, in press.

Fraedrich, K., and L.M. Leslie, 1987: Combining predictive schemes in short-term forecasting. *Monthly Weather Review,* **115**, 1640-1644.

Fraedrich, K., and L.M. Leslie, 1988: Real-time short-term forecasting of precipitation at an Australian tropical station. *Weather and Forecasting,* **3**, 104-114.

Fraedrich, K., and N.R. Smith, 1989: Combining predictive schemes in long-range forecasting. *Journal of Climate,* **2**, 291-294.

Krzysztofowicz, R., 1992: Bayesian correlation score: a utilitarian measure of forecast skill. *Monthly Weather Review,* **120**, 208-219.

Krzysztofowicz, R., and D. Long, 1991a: Forecast sufficiency characteristic: construction and application. *International Journal of Forecasting,* **7**, 39-45.

Krzysztofowicz, R., and D. Long, 1991b: Beta likelihood models of probabilistic forecasts. *International Journal of Forecasting,* **7**, 47-55.

Murphy, A.H., 1991: Forecast verification: its complexity and dimensionality. *Monthly Weather Review,* **119**, 1590-1601.

Murphy, A.H., B.G. Brown, and Y.-S. Chen, 1989: Diagnostic verification of temperature forecasts. *Weather and Forecasting,* **4**, 485-501.

Murphy, A.H., Y.-S. Chen, and R.T. Clemen, 1988: Statistical analysis of interrelationships between objective and subjective temperature forecasts. *Monthly Weather Review,* **116**, 2121-2131.

Murphy, A.H., and R.L. Winkler, 1987: A general framework for forecast verification. *Monthly Weather Review,* **115**, 730-738.

Murphy, A.H., and R.L. Winkler, 1992: Diagnostic verification of probability forecasts. *International Journal of Forecasting,* **8**, in press.

Murphy, A.H., and Q. Ye, 1990: Comparison of objective and subjective precipitation probability forecasts: the sufficiency relation. *Monthly Weather Review,* **118**, 1783-1792.

TACTICAL DECISION MAKING USING SHORT-RANGE FORECASTS: TIMING ALFALFA HARVESTS

D. S. Wilks

Dept. Soil, Crop and Atmospheric Sciences
Cornell University

1. INTRODUCTION

Alfalfa (lucerne) is an important fodder crop especially for dairy production. Typically the harvested plant material is stored before feeding either as hay, where it is preserved by drying; or silage where it is preserved by an anaerobic, acid fermentation. For all haymaking, and for most silage making, the plant material must be left on the ground to dry before storage.

A central issue in harvest management is the detrimental effect of weather on the forage while it is drying. At least potentially, weather information could be used by farm operators to minimize weather damage and increase forage value. The optimal intervals of crop growth between successive cuttings in a given season, and the total number of cuttings taken in a season, are also important considerations in the operation of the forage system of a farm. In practice the decision to cut is often made with only crude guidelines for approximate date, forage maturity at harvest, and weather for the immediate future.

Described here is a model which calculates the economically optimum cutting strategy for alfalfa over the entire growing season for central New York conditions. A decision-analytic framework including explicit consideration of routinely available weather forecasts (Wilks 1991) is employed. The modeled cutting decision is faced daily, and it must balance increasing yield but decreasing forage quality if cutting is deferred, with potential weather-related losses if the cutting decision is made. These potential losses, as well as the probable length of time the forage must be dried before it can be stored, are evaluated using probability forecast information for the upcoming 24-hour period.

2. MODEL COMPONENTS

2.1 *Agronomic considerations*

Agronomic aspects of the model are described more fully in Wilks et al. (1992). Alfalfa growth is modeled following Selerio and Brown (1979), with parameter values for central New York state taken from Fick (1984). Forage composition is specified as a function of plant maturity using regression equations from Fick and Onstad (1988). Monetary valuation, depending on both yield and nutrient composition, is assigned through a modification of the approach of Wilkens and Fick (1988).

Three modes of forage preservation are considered: direct-cut silage, wilted silage, and dry hay. Hay must be dried to a much lower moisture content than wilted silage before it can be removed from the field (and the risk of weather damage) to storage. Direct-cut silage is not field-cured at all, but rather put directly into the silo, which eliminates weather damage risk. However, this option comes at the cost of a required preserving agent (formic acid) and subsequent loss of nutrients (and thus value) by seepage from the silo.

Respiration and rain losses are the primary causes of deterioration during drying. Both dry matter (DM) losses and nutrient composition (quality) changes occur during drying. If rainfall occurs, losses will be sustained both through physical detachment of leaves and through leaching of soluble constituents. These are modeled following Rotz et al. (1989), but combined into a single scalar measure of losses ("damage points"), to improve computational tractability.

2.2 *Forecast Information*

Of primary interest for cutting management are short-range (one-day ahead) rainfall probabilities. It is also necessary to consider temperature forecasts, for two reasons. First, weather events are represented in the decision-analytic framework employed (Wilks 1991) by proxy, using the statistical properties of the weather forecasts and their relationship to the predicted weather events. Thus the degree-day based growth and development of the crop in the model is controlled through the temperature forecasts. Second, pan evaporation, used to determine forage drying rates (Pitt 1985), depends in part on the temperature. Probability information concerning tomorrow's evaporation rate is contained in the current temperature forecast.

Four types of forecast information are considered. These are, in ascending order of information content, 1) climatological information, 2) "persistence" forecasts, 3) standard (i.e., operationally available) weather forecasts, and 4) hypothetical, perfect forecasts. The first two forecast types can be considered to be "baseline" information, perhaps corresponding to decision makers who are unaware of the relevance of weather forecasts to forage preservation. Results concerning the third type of information will be of primary interest, since this represents the presently available forecasts.

Fuller information regarding these four forecast types, in the current context, is provided in Wilks (1991).

The basic precipitation forecasts used here are the familiar Probability of Precipitation (PoP) forecasts issued twice daily by the United States National Weather Service (NWS). These are transformed into 24-hour PoP forecasts using results from Wilks (1990a). Specification of probabilities of rainfall amounts is through the product of the PoP and the conditional probabilities of precipitation amounts given occurrence (Wilks 1990b). Frequency of use of the PoP forecasts are represented using beta distributions, fit separately for the 3-month seasons March-April-May, June-July-August, and September-October-November, using forecast data for Syracuse, New York (Carter and Polger 1986). The day-to-day correlation structure of the forecasts, necessary for representation of event autocorrelation in the decision making model, is modeled as a first-order autoregression of a transformation of these beta distributions to the standard normal distribution (Wilks 1991).

Forecasts for mean daily temperature are obtained by averaging 12-hourly maximum and minimum temperature forecasts. In the model the day-to-day statistical dependence of the temperature forecasts are represented as following a Gaussian, first-order autoregressive process, with parameters fit using data for Syracuse, New York (Carter and Polger 1986). Probability information for the temperatures is obtained through the root-mean-squared error (RMSE) of the temperature forecasts, again using Syracuse data.

2.3 *Weather Events*

The algorithm used to solve the decision-analytic model requires evaluation of statistical expectations with respect to joint temperature and precipitation events. Specifically, while the growth of the alfalfa depends only on temperature, the damage accumulated by drying alfalfa depends both on temperature and rainfall. Similarly, the progress toward drying depends both on rainfall (if any) and daily pan evaporation. Here evaporation probabilities are modeled using Weibull distributions (fit to data from experimental farms at Ithaca and Aurora, New York), conditional on day of the year, average daily temperature (with separate distributions for T <10°C, 10°<T≤20°C, and T>20°C), and precipitation occurrence (separate distributions for wet and dry days).

3. THE DECISION MAKING MODEL

3.1 *Mathematical Framework*

The decision-analytic model is solved recursively, using

$$V_n[\lambda,\zeta] = E_{z|\zeta}\left[\max_i\left\{E_{\theta|z}V_{n-1}[H(A_i,\theta,[\lambda,z])]\right\}\right]. \quad (1)$$

$V_n[\bullet]$ denotes the expected monetary return ($/ha) to be realized between day n and the end of the season. The calculations proceed backward in time and the days are numbered anti-chronologically, so that $V_{n-1}[\bullet]$ pertains to the day following day n, but is calculated before $V_n[\bullet]$. The recursion is initialized at the end of the season, with a monetary penalty assessed if the final cutting has occurred too recently (reflecting detrimental effects of late cutting on winter survival, and on regrowth in the following spring). At the end of the recursion, when the computations have proceeded to the first day of the season, $V_n[\bullet]$ contains the expected return to be realized over an entire growing season. Here 244 time steps are used, starting at 30 November (day n=1) and ending on 1 April (day n=244).

Monetary return depends on the decision to cut, on weather, and on crop quality and yield. The arguments of $V[\bullet]$ consist of the elements of a state vector which for clarity is separated into a vector of nonmeteorological variables, λ, containing 4 elements; and a 2-element meteorological vector of the forecast variables (PoP and temperature) for the previous day, ζ. The state variables and their numerical ranges are summarized in Table 1. The four nonmeteorological state variables comprising λ describe the growth and harvest of the alfalfa stand.

The meteorological events bearing on the alfalfa and cutting management are contained in the three-element vector θ, which is comprised of 1) average daily temperature (relevant to growth, drying, and respiration damage), 2) precipitation amount (relevant to drying, rewetting, and rain damage), and 3) pan evaporation (relevant to drying rate). It is not known, at the time a cutting decision must be made, what the subsequent weather events will be, but probabilities for these three weather elements are related to the temperature and PoP forecasts.

Table 1. Nonmeteorological (λ) and meteorological (ζ) state variables used to represent the harvest system, their numerical ranges, and the number of discrete values each may take on in equation (1).

State Variable	Numerical Range	Discrete Values
λ (agronomic variables):		
Degree-days (5°C base) since previous cutting	0 – 750	76
Number of cuttings since 1 April	0 – 5 *	6*
Moisture content:		
Silage	4–2.3 (g water/g DM)	6
Hay	4–0.25 (g water/g DM)	8
Damage Points	0 – 25	26
ζ (previous day forecasts):		
PoP	0.0 – 1.0	6
Temperature	0° – 34°C	5

*More than 5 cuttings are allowed in the model by specifying that the sixth and subsequent cuttings have the same properties as the fifth, but in practice the model does not call for more than five.

The decision to cut or defer cutting for at least another day is represented by the binary variable, i, and is evaluated within the curly brackets of (1). For each of these two alternatives, the expected value with respect to the distribution of the meteorological events θ given each possible pair of forecasts **z** (denoted by the expectation operator $E_{\theta|z}[\cdot]$) of expected returns for day n-1 is computed. The action resulting in the larger expected return is selected as the optimal action $A^*(\lambda, z)$, which will depend on the day n, as well as the present state of the system λ, and the current forecasts, **z**.

The linkage of state λ for day n, eventual weather θ for day n, the choice A_i of whether or not to cut, and the value of the state vector for day n-1 is made by the transfer function H. This function specifies which state the harvest system will be in on day n–1, given that it starts in a particular state on day n, a particular combination of the three weather elements occurs, and one or the other of the two actions A_i is taken. That is,

$$[\lambda, \zeta]_{n-1} = H(A_i, \theta, [\lambda, z]_n) , \qquad (2)$$

which constitutes the bridge from one day to the next in the recursion (1). Thus the details of forage growth, drying, and weather damage models are contained within the structure of the transfer function H.

A second role played in equation (1) by the transfer function H is in the representation of the serial (day-to-day) correlation of the relevant weather events. In the approach used here (Wilks 1991) this is done by proxy through inclusion of forecasts for day n+1 as part of the state vector, which is also indicated in (2). Once the optimal actions $A^*(\lambda,z)$ have been computed for each of the possible temperature and precipitation forecast pairs **z**, the expected returns for the current day, $V_n[\cdot]$, are computed as the expectation with respect to (the frequency of issuance of) forecasts **z** for day n, conditional on forecasts ζ for day n+1.

3.2 *Monte-Carlo Simulation Based on Optimum Strategies*

Once an optimum cutting management (a complete set of $A^*[\lambda,z]$ for all n) is calculated from the recursion equation, the implementation of this management is stochastically simulated for 10,000 years of daily forecasts and weather. These simulations are effected by proceeding (conventionally) forward in time. Day-by-day weather events are then stochastically simulated, consistent with these forecasts and their real-world accuracy. Values of the state variables are updated through the course of each simulated year, and cutting decisions are specified according to A*. The distributions of dates of first and subsequent cuts, as well as other statistics pertaining to the harvest process, were tabulated from these runs.

4. RESULTS

4.1 *Optimal Actions and Cutting Dates*

The model typically chooses a four-cut management system for the central New York conditions investigated here. Optimal actions for each of the four cuttings are summarized in Figures 1, 2, 3, and 4, for the case of wilted silage. The vertical axes in Figures 1–4 show stand maturity, in terms of degree-days on the left (linear scale) and corresponding DM accumulations on the right. The shaded regions in these figures portray optimal actions A* as a function of stand maturity, and of the PoP forecast, for each cutting. In unshaded regions cutting is not the optimal action, regardless of the weather forecast. The lightest stippling indicates that stand should be cut only if the PoP forecast is < 0.1 (nominally 5%). Progressively heavier stippling indicates that cutting is optimal given higher probabilities of rain, but that the stand should still be cut for the drier rainfall probabilities as well.

The distributions of cutting dates produced through Monte-Carlo simulation using A* are shown at the bottom of each panel, in the form of modified box plots. Here the dots show earliest and latest dates for each cutting, and the other features of the box plots indicate the dates corresponding to particular percentiles of the distribution of cutting dates: the ends of the whiskers show the 5th and 95th percentiles, the ends of the boxes show the 25th and 75th percentiles, and the vertical line in the middle of the boxes indicates the median cutting date.

Figure 1 exhibits relatively little structure in comparison to the subsequent three figures. Simulated cutting dates using the optimal strategy are centered on late May and early June, in agreement with conventional practice. The stippled regions broaden slightly to the right of the figure, indicating that the optimal cutting strategy corresponds to a progressively smaller yield as time goes on, balancing effective use of what remains of the growing season for subsequent cuts. If the stand has passed the point where value per ha actually declines with increasing maturity (about 450 degree-days for the first cutting), the model calls for cutting even if the PoP is in the range 0.1 to 0.3 (nominally 20%), both to avoid future declines in value per ha as the stand matures further, and to be able to devote more of the remainder of the growing season to subsequent growth. With very minor exception, however, the model does not recommend cutting if the forecast is less favorable than this (PoP > 0.3), regardless of how overmature the stand may be.

Figure 2 portrays optimal actions for the second cutting. The end-of-season penalty is still sufficiently far in the future that the decision regions are relatively smooth and resemble those for the first cutting in Figure 1. Through the middle of July, at least, the decision region broadens with time. Cutting is only called for if the PoP is < 0.1, until the forage value per hectare reaches its maximum, at around 550 degree-days. As before, cutting is never called for if the PoP is greater than 0.3.

Figure 3 shows the optimal decision regions for the third cutting, where the mid-August bulge in the decision region reflects some urgency for the third cutting to be completed, so that sufficient regrowth for a fourth cutting can occur before late September. At the maximum of this bulge,

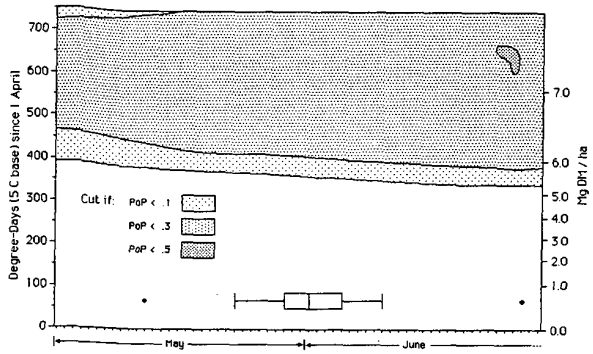


Figure 1. Modeled optimal actions using weather (PoP) forecasts for first-cut wilted silage, as a function of the date and of stand maturity (or, equivalently, yield). Boxplots portray the distribution of cutting dates over 10,000 simulated years managed according to the optimal actions.

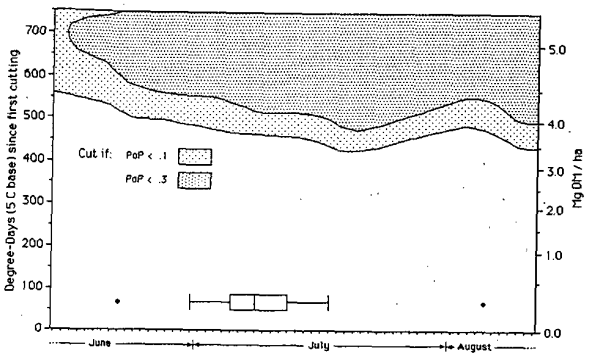


Figure 2. As Figure 1, for the second cutting.

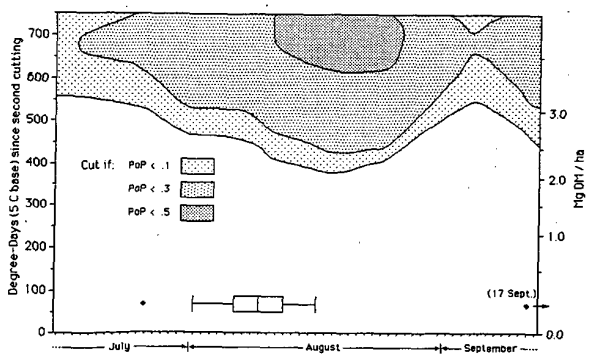


Figure 3. As Figure 1, for the third cutting.

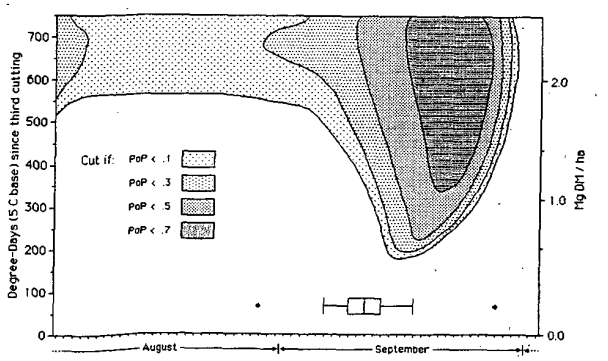


Figure 4. As Figure 1, for the fourth cutting.

cutting is called for with only about 400 degree-days accumulated, provided the rainfall forecast is favorable. The boxplot representing the distribution of simulated cutting dates indicates that this bulge is very effective in spurring a third cutting before mid-August in most simulated years. Again, the upward constriction of the decision region in early September effectively indicates an abandonment of the idea of four cuts, so that the motivation to cut is relaxed unless the stand is very mature, with the decision region for the third (now final) cut extending again through 500 degree-days in mid-September.

Figure 4 shows optimal decisions for the fourth (and usually final) cutting. Here the decision process is approaching the end-of-season penalty, and to avoid this the model will accept low yields and tolerate high probabilities of rain damage and extended drying periods.

Prescribed actions based on climatological or "persistence" information are qualitatively similar to those presented above for optimal use of weather forecasts, although of course inferior monetary returns are expected as a consequence of their lesser information content. Figure 5 illustrates the nature of this relationship for the second cutting. The shapes of the decision regions and the distributions of cutting dates resemble those for the more sophisticated weather forecasts shown in Figure 2. This is not surprising, since these two reference information sources are, in effect, relatively unskillful probability forecasts.

4.2 *Value of Information and Risk*

A comparison of the three forage preservation approaches and the four weather information sources, in terms of annual monetary return, is presented in Figure 6. The horizontal axis is expected (average) annual income per hectare, and the vertical axis is the coefficient of variation of annual income (i.e., the standard deviation of annual income divided by the expected value). Both of these quantities were computed from the outcomes of 10,000 stochastically simulated cropping years each, generated as described above. The expected income is also a result of (1), but its variability is not easily recovered from the recursion. Circles, squares, and

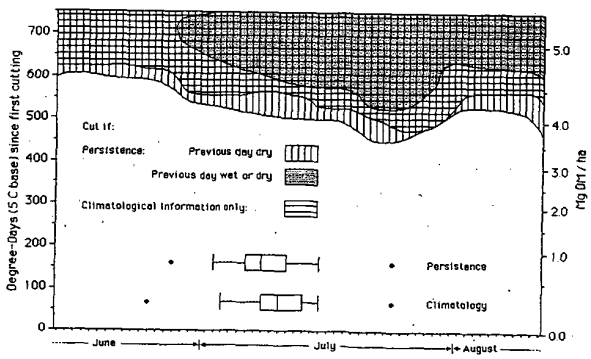


Figure 5. As Figure 1, but using climatological or persistence information, for the second cutting (compare Figure 2).

rotated squares represent hay, wilted silage, and direct-cut silage, respectively. The density of shading increases with increasing forecast information content, ranging from climatological to perfect information.

Clearly higher annual returns are more desirable so that points further to the right on the diagram represent superior outcomes, other things being equal. In terms of average income, wilted silage is simulated to be preferable to direct-cut silage, which in turn appears to be preferable to haymaking, for the central New York conditions simulated here. It does not appear that the weather damage avoided in the making of direct-cut silage is worth the cost of formic acid and nutrients lost in silo effluent, at least on an expected value basis. Note that while this comparison between wilted and direct-cut silage is probably reasonable, the relationship between expected income for hay and either silage system should be made cautiously. In particular, post-harvest losses will likely be lower for hay as compared to silage (where deterioration unrelated to weather conditions may occur), although average silo losses of about 25% (or 35%) in monetary value would be required before expected returns for hay and direct-cut (or wilted) silage would be comparable.

The information in Figure 6 also allows evaluation of the expected value of using weather forecasts to improve the day-to-day decision making. Conventionally, the value of information in this context is computed as the increase in expected return over that associated with a baseline, or reference information source (e.g., Winkler and Murphy 1985). Here it is probably most reasonable to represent this reference level by the persistence information (light stippling). The average annual increase in value to be derived from using short-range weather forecasts can thus be estimated as the difference in expected annual returns between the "persistence" information and the conventional weather forecasts (heavy stippling), amounting to about $84/ha (+5.4%) for wilted silage, and $106/ha (+11.1%) for hay. Both increases come with decreases in production risk. For wilted silage the increase in expected value constitutes about 54% of that for perfect 24-hour temperature and PoP forecasts, representing maximum information value. For the case of hay the expected value of forecast and perfect information are almost indistinguishable, which is less an indication of high forecast quality than of the fact that the haymaking process will typically extend well beyond the 1-day forecast horizon imposed in the model. Expected returns for direct-cut silage are virtually independent of the source of meteorological (in this case, only temperature) information, consistent with the aim of this preservation mode to minimize sensitivity to harvest weather.

The vertical dimension in Figure 6 is a portrayal of variability of income or, equivalently, of risk. One would anticipate that, for a given level of expected return, a real-world forage manager would prefer to minimize year-to-year variations in income. Thus, other things being equal, points lower on the diagram in Figure 6 would represent more

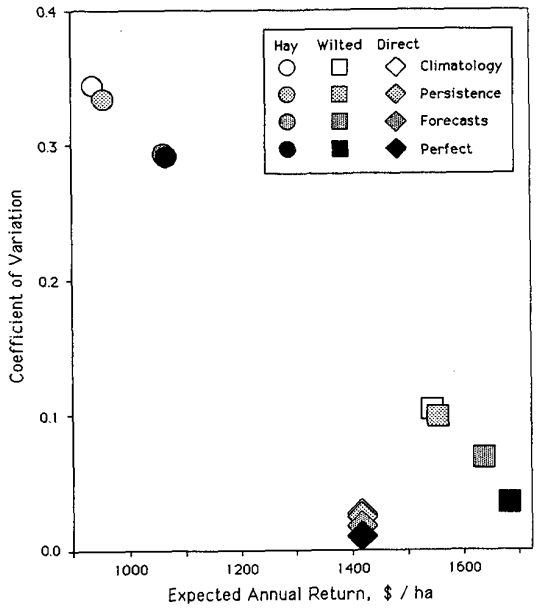


Figure 6. Relationship of expected annual return per hectare and variability of annual returns, for haymaking (circles), wilted silage (squares), and direct-cut silage (rotated squares). Quality of the weather information increases with density of shading from climatological probabilities only (open symbols) to hypothetical perfect forecasts (solid symbols).

desirable outcomes. Regardless of the type of meteorological information, haymaking is tremendously riskier than either of the modes of silage making considered. While expected returns for direct-cut silage are consistently lower than for wilted silage, the risk associated with this form of forage preservation is substantially lower. If a decision maker were sufficiently risk-averse it is possible that the decrease in risk associated with direct-cut would be more than worth the lower expected returns. This would be the case particularly if only climatological or persistence information were available. In all cases, increasing meteorological information decreases production risk.

5. CONCLUSIONS

The decision-analytic approach in this study has provided a set of prescriptions for optimal timing of alfalfa harvests using routinely available weather forecasts, for central New York conditions. Computed optimal actions result in a four-cut harvest system, which is increasingly coming into use in practice. A decision to cut is recommended only for very low rainfall probabilities unless the stand is advancing sufficiently beyond optimal maturity.

Results indicate that wilted silage is more reliable and produces higher forage values than dry hay production in central New York state. It should be recognized, however, that longer-range forecasts are also available to real-world decision makers, and a desirable extension to this work would be inclusion of these sources of information into the modeled

decision process. While direct-cut silage provides more consistent (less variable) returns, the cost of formic acid together with loss of nutrients leached from the silo combine to make this option less preferable, except perhaps to the most risk-averse decision makers.

Finally, the calculations performed here indicate that substantial monetary value should be expected if weather forecasts are used intelligently to schedule alfalfa harvest. Average annual increases of roughly 5 to 10% over baseline decision making are estimated to accrue from the use of routinely available weather forecasts, together with decreased risk. It thus seems well worth the alfalfa producer's time and effort to account for the forecasts in harvest decision making.

Acknowledgements: This work was supported by the USDA under regional project NE-132, and project NYC-125410.

REFERENCES

Carter, G.M., and P.D. Polger, 1986. A 20-Year Summary of National Weather Service Verification Results for Temperature and Precipitation. NOAA Tech. Memo. NWS FCST 31. NWS, Silver Spring, MD, 50 pp.

Fick, G.W., 1984. Simple simulation models for yield prediction applied to alfalfa in the northeast. *Agronomy Journal*, 76, 235-239.

Fick, G.W., and D.W. Onstad, 1988. Statistical models for predicting alfalfa herbage quality from morphological or weather data. *Journal of Production Agriculture*, 1, 160-166.

Pitt, R.E., 1985. Evaporation-based drying rate of forage: effects of dessicants and crop density. *Journal of Agricultural Science, Cambridge*, 105, 223-229.

Rotz, C.A., J.R. Black, D.R. Mertens, and D.R. Buckmaster, 1989. DAFOSYM: A model of the dairy forage system. *Journal of Production Agriculture*, 2, 83-91.

Selirio, I.S., and D.M. Brown, 1979. Soil moisture-based simulation of forage yield. *Agricultural Meteorology*, 20, 99-114.

Wilkens, P.W., and G.W. Fick, 1988. FORVAL: A computer program using chemical analysis and market data to price hay. *Journal of Agronomic Education*, 17, 122-127.

Wilks, D.S., 1990a. On the combination of forecast probabilities for consecutive precipitation periods. *Weather and Forecasting*, 5, 640-650.

Wilks, D.S., 1990b. Probabilistic quantitative precipitation forecasts derived from PoPs and conditional precipitation amount probabilities. *Monthly Weather Review*, 118, 874-882.

Wilks, D.S., 1991. Representing serial correlation of meteorological events and forecasts in dynamic decision-analytic models. *Monthly Weather Review*, 119, 1640-1662.

Wilks, D.S., R.E. Pitt, and G.W. Fick, 1992. Modeling optimal alfalfa harvest scheduling using short-range weather forecasts. *Agricultural Systems*, submitted.

Winkler, R.L., and A.H. Murphy, 1985. Decision Analysis. In: A.H. Murphy and R.W. Katz, eds., *Probability, Statistics, and Decision Making in the Atmospheric Sciences*. Westview, 493-524.

J4.3

ECONOMICALLY OPTIMAL DECISIONS AND THE VALUE OF METEOROLOGICAL INFORMATION

Lev S. Gandin[1], Allan H. Murphy[2], and Eugene E. Zhukovsky[3]

[1]National Meteorological Center, National Weather Service, Camp Springs, Maryland 20746, U.S.A.
[2]Department of Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331, U.S.A.
[3]Institute of Agronomical Physics, St. Petersburg 195220, Russia.

1. Introduction

Information related to future weather and/or climate conditions can take many different forms. Two classes of such information are distinguished here: (a) climatological information and (b) meteorological information. The former is information derived from the analysis of historical data and, under the assumption of statistical stationarity, this information can be "translated" into forecasts of future conditions. The latter is information derived from numerical, statistical, or conceptual models that explicitly forecast future conditions. This information can take the form of nowcasts, short-range or medium-range weather forecasts, or long-range weather - or short-range climate - forecasts.

The purpose of producing such information - an activity to which the meteorological community devotes a substantial fraction of its human and financial resources - is presumably to enable individuals and organizations whose activities are sensitive to weather and/or climate conditions to make good decisions. In this context, "good decisions" are decisions that enhance economic and social welfare (e.g., by increasing benefits or decreasing disbenefits). As noted by Winkler *et al.* (1983), information acquires value only through its ability to influence decisions made by information-sensitive users. Clearly, then, it is essential that both producers and consumers of weather and climate forecasts become aware of (a) the conditions under which various strategies for using this information are optimal and (b) the benefits (disbenefits) of making optimal (suboptimal) decisions.

A variety of studies of the use and value of weather/climate forecasts have been undertaken in the last 25 years (see Katz and Murphy, 1992; Murphy and Brown, 1982; Winkler and Murphy, 1985). These studies have ranged from surveys of selected groups of users to analyses in which the decision-making problems of specific users were described using decision-analytic models. For example, both single-stage ("static") and multiple-stage ("dynamic") models have been used in prescriptive studies of the optimal use and economic value of weather/climate forecasts (e.g., Brown *et al.*, 1986; Katz *et al.*, 1982; Wilks and Murphy, 1986). In most studies of this type, attention has focused primarily on obtaining forecast-value estimates and/or exploring the relationship between forecast quality and forecast value.

The overall purpose of this paper is to investigate the optimal use and economic value of weather/climate forecasts in prototypical decision-making problems. In particular, we are concerned with identifying the conditions under which the use of different types of information is optimal or suboptimal. To accomplish these objectives, two simple weather/climate-information-sensitive decision-making problems are considered: (a) a situation in which only two actions are available to the user and the weather/climate conditions are described by a binary variable (the *discrete example)* and (b) a situation in which the weather/climate variable is continuous and the user's decision involves the choice of a specific value for this variable (the *continuous example).*

Specifically, we investigate the optimal use of forecasts from the perspective of potential users who want to behave in an economically optimal manner. This perspective implies that a prescriptive approach be adopted to users' decision-making problems, and the basic elements of this approach are briefly described in Section 2. The discrete and continuous examples are considered in some detail in Sections 3 and 4, respectively. Section 5 discusses the conclusions that can be drawn from these examples and their implications for the meteorological and user communities. This section also contains a brief summary and some concluding remarks.

2. Elements of Decision-Making Problems: Prescriptive Approach

This paper focuses on individuals whose activities are sensitive to weather or climate conditions. In particular, it is assumed that these individuals must make decisions in the face of uncertainty regarding future meteorological conditions and that, as a result, they are potential users of weather or climate forecasts. For convenience, we refer to these individuals as *users* or *decision makers.*

We are concerned here with prescriptive analyses and models of users' decision-making problems. That is, the analyses and models of interest focus on the ways that users should behave in order to make optimal decisions and thereby maximize their economic welfare (as opposed to the ways that they might actually behave in the real world). In this regard, it may be useful to identify - and briefly describe - the basic elements of prescriptive decision-making problems:

(a) Set of actions. All such problems necessarily involve a set of two or more actions that are admissible (i.e., optimal

under some conditions) from the user's point of view.

(b) Set of events. The events describe a mutually exclusive and collectively exhaustive set of conditions that characterize the possible "states of the world." In this paper, these events are defined in terms of the values of a single weather/climate variable.

(c) Set of consequences. Each unique combination of action and event leads to a distinct consequence. For example, in a two-action, two-event decision-making problem (see Section 3a), four distinct consequences can be identified.

(d) Utilities of consequences. The utilities of the consequences describe the worth of these consequences to the decision maker. In general, the worth of a consequence may be expressed in economic or noneconomic terms. Here, we assume that all utilities are expressed in monetary terms (i.e., in terms of economic gains or losses).

(e) Probabilities of events. The probabilities of the events indicate the likelihoods of occurrence of the various events. These probabilities can be derived from climatological information or from forecasts based on numerical, statistical, or conceptual models. In the case of nonprobabilistic (or categorical) forecasts, the probabilities are restricted to values of zero and one.

(f) Decision criterion. It is necessary to specify a criterion which the user will follow in selecting a course of action. We will assume here that the user chooses the action that minimizes his/her expected (or average)losses, where the expected loss (associated with an action) is the probability-weighted average of the relevant individual losses. In effect, we are assuming that the decision maker's utility function is linear in monetary loss (e.g., see Winkler and Murphy, 1985).

The discrete and continuous examples considered in Sections 3 and 4, respectively, possess all of these elements. However, the number of actions and events - as well as the nature of the forecasts - differs between the two decision-making problems.

3. Discrete Decision-Making Problem: Example 1

a. Basic situation, with climatological information

Consider a decision-making problem involving two actions (a_1 and a_2) and two weather/climate events (x_1 and x_2). Let d_{ij} denote the loss incurred by the decision maker when action a_i is taken and event x_j occurs (i,j=1,2). The loss matrix characterizing this general two-action, two-event (i.e., 2x2) problem is presented in Table 1. It is a generalization of the expense matrix associated with the familiar cost-loss ratio situation (e.g., Thompson, 1962; Murphy, 1977).

Table 1. General 2x2 decision-making problem.

		Events	
		x_1	x_2
Actions	a_1	d_{11}	d_{12}
	a_2	d_{21}	d_{22}

It is assumed initially that only climatological information is available to the decision maker. Let $p_1 = P(x = x_1)$ and $p_2 = P(x = x_2) = 1 - p_1$ denote these climatological probabilities. As indicated in Section 2, it is assumed here that the user chooses the action that minimizes expected (or average) loss, where this expectation is taken with respect to the probability distribution over the individual losses.

To determine the optimal action, the expected losses associated with a_1 and a_2 must be compared. These expected losses are $D_1 = p_1d_{11} + p_2d_{12}$ and $D_2 = p_1d_{21} + p_2d_{22}$, respectively. Thus, the user prefers a_1 if $D_1 < D_2$, or $(d_{21} - d_{11})/(d_{12} - d_{22}) > p_2/p_1$, and a_2 if $D_1 > D_2$, or $(d_{21} - d_{11})/(d_{12} - d_{22}) < p_2/p_1$. The former is said to be inclined toward a_1 and the latter is said to be inclined toward a_2. Only those users for whom $D_1 = D_2$, or $(d_{21} - d_{11})/(d_{12} - d_{22}) = p_2/p_1$, exhibit no such inclination. These so-called *ideal users* (Obukhov, 1955) are indifferent between a_1 and a_2.

Note that it is only the difference between d_{21} and d_{11} (and between d_{12} and d_{22}), rather than their actual values, that influences the choice of the optimal action. For decision-making purposes, $d_{21} - d_{11}$ and $d_{12} - d_{22}$ can be replaced by d_1 and d_2, respectively. These differences in loss are referred to here as *meteorological losses,* since they represent the incremental losses that are incurred due to incomplete knowledge as to which event will occur. The 2x2 problem is described in terms of these meteorological losses in Table 2.

Table 2. General 2x2 problem, with meteorological losses.

		Events	
		x_1	x_2
Actions	a_1	0	d_2
	a_2	d_1	0

The ratio d_1/d_2 characterizes users in the general 2x2 problem in the same way that the cost-loss ratio characterizes users in the cost-loss ratio situation. In this regard, users inclined toward a_1 would occupy the finite line segment $0 < d_1/d_2 < p_2/p_1$ and users inclined toward a_2 would occupy the infinite line segment $d_1/d_2 > p_2/p_1$. Moreover, it can be shown that the expected economic value realized through the use of forecasts (or any other type of meteorological information) is a linear function of this ratio over one of these two line segments and a nonlinear function of the ratio over the other line segment. To avoid these shortcomings, it is convenient to define a user parameter u as follows:

$$u = \begin{cases} 1 - (p_1d_1/p_2d_2) & \text{if } p_1d_1 \geq p_2d_2 \\ (p_2d_2/p_1d_1) - 1 & \text{if } p_1d_1 \leq p_2d_2. \end{cases} \quad (1)$$

The parameter u, the value of which characterizes each user, is negative for users inclined toward a_1, positive for users inclined toward a_2, and zero for ideal users.

b. Forecasts from one forecasting system

Now it is assumed that, in addition to climatological information, a nonprobabilistic forecast f - produced by a forecasting system F - is available to the decision maker. Let f_i denote a forecast of the event x_i (i=1,2). Since the quality of forecasts is completely described by the joint distribution of forecast and observed events (Murphy and Winkler, 1987), the quality of the forecasts produced by system F can be characterized by a 2x2 performance matrix. Such a matrix is depicted in Table 3, in which $p_{ij} = P(f=f_i, x=x_j)$. Table 3 also contains the marginal distributions of the forecasts and

Table 3. Performance matrix in two-event situation, with nonprobabilistic forecasts.

		Observations		
		x_1	x_2	
Forecasts	f_1	p_{11}	p_{12}	q_1
	f_2	p_{21}	p_{22}	q_2
		p_1	p_2	1

observations, $q_i = P(f = f_i)$ and $p_j = P(x = x_j)$, respectively (i,j=1,2). If $q_1 = p_1$ (in which case $q_2 = p_2$ as well), system F's forecasts are (unconditionally) unbiased.

The presence of the forecasts implies that four strategies are available to the user: (1) S_1 - always take action a_1; (2) S_2 - always take action a_2; (3) S_3 - take a_1 (a_2) when f = f_1 (f_2); and (4) S_4 - take a_2 (a_1) when f = f_1 (f_2). When the strategies S_1 and S_2 are followed, the user ignores the forecasts. In the case of S_3 the user takes the forecasts at face value, whereas in the case of S_4 the user (in effect) reverses the labels on the forecasts. The latter strategy is included to treat the situations in which x_2 occurs more frequently than x_1 when f = f_1 or x_1 occurs more frequently than x_2 when f = f_2. If we assume that the labels on such forecasts are always reversed before they are used, then strategies such as S_4 need not be of further interest.

A user who follows strategy S_1 (a_1 always; u < 0) incurs expected losses $L_1 = p_2 d_2$. Analogously, a user who follows strategy S_2 (a_2 always; u > 0) incurs expected losses $L_2 = p_1 d_1$. On the other hand, if the user follows strategy S_3, his/her expected losses are $L_3 = p_{12} d_2 + p_{21} d_1$. If the economic value of the forecasts (V) is defined as the relative decrease in the user's expected losses, when he/she makes use of the forecasts, then $V = (L_1 - L_3)/L_1$ when $u \leq 0$ and $V = (L_2 - L_3)/L_2$ when $u \geq 0$. Substituting the expressions for L_1, L_2, and L_3 into the expressions for V yields

$$V = \begin{cases} K + p_1^* u & u \leq 0 \\ K - p_2^* u & u \geq 0, \end{cases} \qquad (2)$$

where $K = (p_{11}p_{22} - p_{12}p_{21})/(p_1 p_2)$, $p_1^* = p_{21}/p_1$, and $p_2^* = p_{12}/p_2$. The numerator of K is the determinant (D) of the performance matrix and p_1^* (p_2^*) is the conditional probability of f = f_2 (f_1) given x = x_1 (x_2). It should also be noted that K is identical to the sample version of Kuiper's performance index (see Murphy and Daan, 1985), a familiar measure of overall forecasting performance in the situations of interest here. Moreover, K is closely related to the coefficient of qualitative correlation r, where $r = D/(p_1 p_2 q_1 q_2)^{1/2}$. It can now be seen that strategy S_3 (S_4) is always preferred to strategy S_4 (S_3) when r > (<) 0 [or, equivalently, K > (<) 0].

Examination of the expression for V in (2) leads to several interesting conclusions regarding the use and value of the forecasts. First, for all r > 0, at least some users can be identified who can benefit from the forecasts (i.e., for whom S_3 is the optimal strategy). For r < 0, an analogous statement could be made with respect to S_4. Only those forecasts for which r = 0 are of no benefit to any users.

Second, the value of the forecasts is a maximum for ideal users (u = 0). This maximum value never exceeds one. It approaches one when p_{12} and p_{21} approach zero.

Third, the value of the forecasts to users inclined toward a_1 or a_2 decreases linearly with u. This rate of decrease is less than the rate of increase in expected loss occasioned by the use of the inappropriate climatological strategy (i.e., using S_2 instead of S_1 when u < 0 or using S_1 instead of S_2 when u > 0).

Fourth, V = 0 for users with $u = u_{min} = -K/p_1^*$ and V < 0 for users with $u < u_{min}$. Analogously, V = 0 for users with $u = u_{max} = K/p_2^*$ and V < 0 for users with $u > u_{max}$. Thus, S_1 is optimal for users with $u < u_{min}$ and S_2 is optimal for users with $u > u_{max}$. In other words, some users always exist whose inclinations toward the climatological strategies are strong enough such that a strategy of ignoring the forecasts (S_1 or S_2) is preferred to a strategy of using the forecasts (S_3 or S_4). The only users able to benefit from all forecasts for which r > 0 (or r < 0) are ideal users. This result suggests that users might benefit by restructuring their decision-making problems to approximate ideal users (e.g., by making the ratio of their meteorological losses, d_1/d_2, more closely correspond to the relevant climatological odds, p_2/p_1).

These results can be illustrated in a schematic diagram (diagram omitted), in which V is plotted against the user parameter u for situations in which r ≥ 0. The "space" of users (the infinite line segment) is divided into three intervals: (1) I_1 ($u < u_{min}$) for which S_1 is optimal; (2) I_2 ($u_{min} < u < u_{max}$) for which S_3 is optimal; and (3) I_3 ($u > u_{max}$) for which S_2 is optimal. Note that V is positive in I_2, attaining its maximum value for u = 0 (ideal users). On the other hand, V is negative in both I_1 and I_3. Users whose values of u fall in these latter intervals should ignore the forecasts and base their decisions on climatological information.

c. Forecasts from two forecasting systems

Now consider a situation in which forecasts produced by two forecasting systems, F and G, are available to the decision maker. Since these forecasts relate to the same events, the climatological probabilities, p_1 and p_2, are the identical for both systems. The quality of both F's and G's forecasts can be characterized by a performance matrix similar (in format) to that depicted in Table 3. Although the performance matrix for

each system is available to the user, it is assumed here that the user does not have access to information (e.g., the trivariate distribution of F's forecasts, G's forecasts, and the observations) describing the joint performance of the two systems. Thus, the decision maker is able to use the forecasts produced by each system separately, but he/she is unable to combine the forecasts produced by the two systems in the process of identifying the economically optimal strategy.

Under these assumptions, six strategies are available to the user: (1) S_1 - always take a_1; (2) S_2 - always take a_2; (3) S_3 - take a_1 if $f = f_1$ and a_2 if $f = f_2$; (4) S_4 - take a_2 if $f = f_1$ and a_1 if $f = f_2$; (5) S_5 - take a_1 if $g = g_1$ and a_2 if $g = g_2$; and (6) S_6 - take a_2 if $g = g_1$ and a_1 if $g = g_2$. In brief, S_1 and S_2 are climatological strategies, S_3 and S_4 are strategies involving the use of F's forecasts (S_3 - take F's forecasts at face value, S_4 - relabel F's forecasts); and S_5 and S_6 are strategies involving the use of G's forecasts (S_5 - take G's forecasts at face value, S_6 - relabel G's forecasts).

Assuming that the climatological probability p_1 (= 1 - p_2) is known, each performance matrix possesses two degrees of freedom (i.e., two additional "parameters" must be specified to determine the p_{ij}; see Table 3). We first consider several special situations in which an additional restriction is placed on the values of the p_{ij} in both performance matrices. Such restrictions reduce the number of degrees of freedom from two to one in each case and substantially simplify the analysis. Then the general situation is briefly considered.

Suppose that systems F and G provide unbiased forecasts. Then $p_1 = q_1$ and $p_{12} = p_{21} = p$ for both systems. In addition, $D = p_1p_2 - p$, $p_1^* = p/p_1$, and $p_2^* = p/p_2$. Let p_F and p_G denote one-half of the probability that F and G, respectively, produce an incorrect forecast, and suppose that $p_F < p_G$. It follows that $D_F > D_G$, $p_1^*(F) < p_1^*(G)$, and $p_2^*(F) < p_2^*(G)$. Moreover, the fact that $u_{min} = -K/p_1^*$, $u_{max} = K/p_2^*$, and $K = D/(p_1p_2)$ reveals that F's forecasts are better than G's forecasts in all respects. Specifically, the former may be profitably applied by a wider range of users than the latter, and the former provide greater economic benefit than the latter.

This result can be illustrated by means of a schematic diagram (omitted) in which it is assumed that $D > 0$ for both F and G. In this case, the space of users is an infinite line segment divided into five intervals: (1) I_1 [$u < u_{min}(F)$] for which S_1 is optimal and the value of both F's and G's forecasts is negative; (2) I_2 [$u_{min}(F) < u < u_{min}(G)$] for which S_3 is optimal and the value of F's (G's) forecasts is positive (negative); (3) I_3 [$u_{min}(G) < u < u_{max}(G)$] for which S_3 is optimal and the value of both F's and G's forecasts is positive; (4) I_4 [$u_{max}(G) < u < u_{max}(F)$] for which S_3 is optimal and the value of F's (G's) forecasts is positive (negative); and (5) I_5 [$u > u_{max}(F)$] for which S_2 is optimal and the value of both F's and G's forecasts is negative. Thus, for all values of u between $u_{min}(F)$ and $u_{max}(F)$, the economic value of F's forecasts is positive and exceeds the economic value of G's forecasts. In this sense, F's forecasts "dominate" G's forecasts. Outside of this interval [i.e., $u < u_{min}(F)$ or $u > u_{max}(F)$], users should ignore both F's and G's forecasts and base their decisions on climatological information.

This same result is also obtained in the more general situation in which systems F and G are assumed to produce forecasts that are equally biased. In this case, the difference $p_1 - q_1$ possesses the same value for both systems, and $D = p_1p_2 - e + b(p_2 - p_1)$, in which $b = (q_1 - p_1)/2$ and $e = (p_{12} + p_{21})/2$. Note that b is the same for both F and G, whereas e is different for the two systems. Thus, this situation leads to the same qualitative conclusions as those arrived at in the situation in which the forecasts are assumed to be unbiased (see previous paragraph).

A different situation (with different results) arises when it is assumed that systems F and G produce forecasts with the same value of the determinant D. First, note that this assumption also implies that both systems possess the same value of K. It can easily be shown that $p_{12} = p_1p_2 - D + 2bp_1$ and $p_{21} = p_1p_2 - D - 2bp_2$. The first two terms on the right-hand side of these expressions are the same for both forecasting systems. However, the third term is different and possesses opposite signs in the two expressions. This difference has important consequences - it implies that, for every F and G, one system can be used more profitably than the other system by users inclined toward strategy S_1 and the opposite is true for those users inclined toward strategy S_2. Moreover, users will always exist who can profitably apply system F but not system G, and vice versa.

A schematic diagram illustrating this situation divides the infinite line segment into six intervals: (1) I_1 [$u < u_{min}(G)$] for which S_1 is optimal and the value of both F's and G's forecasts is negative; (2) I_2 [$u_{min}(G) < u < u_{min}(F)$] for which S_5 is optimal and the value of F's (G's) forecasts is negative (positive); (3) I_3 [$u_{min}(F) < u < 0$] for which S_5 is optimal and the value of both F's and G's forecasts is positive; (4) I_4 [$0 < u < u_{max}(G)$] for which S_3 is optimal and the value of both F's and G's forecasts is positive; (5) I_5 [$u_{max}(G) < u < u_{max}(F)$] for which S_3 is optimal and the value of F's (G's) forecasts is positive (negative); and (6) I_6 [$u > u_{max}(F)$] for which S_2 is optimal and the value of both F's and G's forecasts is negative. Thus, the value of F's forecasts is positive and exceeds the value of G's forecasts for all users for whom $0 < u < u_{max}(F)$ and the value of G's forecasts is positive and exceeds the value of F's forecasts for all users for whom $u_{min}(G) < u < 0$. Users with values of u outside these intervals should ignore both F's and G's forecasts and base their decisions on climatological information.

The significance of these results becomes apparent when traditional practices in forecast verification are considered. Generally, a performance measure such as the correlation coefficient (r), or Kuiper's performance index (K), is used to determine whether system F's forecasts are better (or worse) than systems G's forecasts. But the arguments set forth in the previous paragraph reveal that some users always exist who can realize greater economic gain by relying on the "inferior" forecasts than by relying on the "superior" forecasts. Moreover, some users also exist who can realize benefits from the inferior forecasts but whose "benefits" would be negative if they relied on - and believed in - the superior forecasts.

The same arguments apply in the general case in which the performance matrix possesses two degrees of freedom.

Suppose that F is found to produce better forecasts than G in such a situation according to some performance measure (e.g., the correlation coefficient r). Despite this ordinal ranking in performance, some users exist who can benefit more in an economic sense from G's forecasts than from F's forecasts, and some users exist who can benefit from G's forecasts yet can only realize disbenefits from F's forecasts. Similar results have been obtained in the context of the basic cost-loss ratio sitiuation by Murphy and Ehrendorfer (1987).

4. Continuous Decision-Making Problem: Example 2

a. Basic situation, with climatological information

In many weather/climate-information-sensitive decision-making problems, the variable of interest is a continuous variable rather than a discrete (or even dichotomous) variable. Let X denote such a variable. If X is also the decision variable, in the sense that the decision maker must choose a particular value of this variable, then an essentially infinite number of possible decisions (or actions) exists. We consider here a situation in which the meteorological losses incurred by the decision maker depend only on the difference between the actual value of the variable (x) and its assumed or chosen value (y). Specifically, the loss function L(x,y) is assumed to be of the following form:

$$L(x,y) = \begin{cases} A_+(y - x) & \text{if } x \le y \\ A_-(x - y) & \text{if } x \ge y, \end{cases} \qquad (3)$$

where A_+ and A_- are positive weights that transform errors of overestimation ("positive" errors) and errors of underestimation ("negative" errors), respectively, into losses. In general, it is assumed that $A_+ \neq A_-$, and this inequality indicates that the loss function is asymmetric.

It is convenient here to introduce a nondimensional index α based on A+ and A-, where $\alpha = (A_+ - A_-)/(A_+ + A_-)$. Then $-1 \le \alpha \le 1$, with $\alpha = 0$ representing a symmetric loss function ($A_+ = A_-$), $-1 \le \alpha < 0$ representing situations in which positive errors ($y - x > 0$) produce larger losses, and $0 < \alpha \le 1$ representing situations in which negative errors ($y - x < 0$) produce larger losses.

The mean (or average) loss as a function of the chosen value y, $\langle L(y)\rangle$, can be determined by integrating L(x,y) over all values of x. That is, $\langle L(y)\rangle = \int_x L(x,y)f(x)dx$, where f(x) denotes the probability density function of x (the values of x range from $-\infty$ to $+\infty$). If x possesses a Gaussian distribution with mean μ and variance σ^2, then it can be shown that

$$\langle L(y)\rangle = (2^{1/2}\sigma/\pi^{1/2})F(v), \qquad (4)$$

where $F(v) = [(A_+ + A_-)/2][e^{-v^2} + \pi^{1/2}v\,\mathrm{erf}(v) - \pi^{1/2}\alpha v]$, in which $v = (y - \mu)/2^{1/2}\sigma$ and erf(v) is the error function with argument v.

Suppose that y is chosen to be the mean of x (i.e., $y = \mu$). Then $v = 0$ and $\langle L(y)\rangle = L^*$, where

$$L^* = [\sigma/(2\pi)^{1/2}](A_+ + A_-). \qquad (5)$$

Combining (5) with the expression for $\langle L(y)\rangle$ in (4) yields

$$\langle L(y)\rangle = L^*[e^{-v^2} + \pi^{1/2}v\,\mathrm{erf}(v) - \pi^{1/2}\alpha v]. \qquad (6)$$

The ratio $\langle L(y)\rangle/L^*$ is plotted as a function of v in Figure 1, for selected values of the asymmetry parameter α. It can be seen from this figure that the ratio $\langle L(y)\rangle/L^*$ attains its minimum value when $v = 0$ ($y = \mu$) only if $\alpha = 0$. That is, the mean loss is minimized by choosing the mean value of the variable X only in the case in which the loss function is symmetric ($A_+ = A_-$ and $\alpha = 0$). For asymmetric loss functions ($A_+ \neq A_-$ and $\alpha \neq 0$), the minimum value of this ratio occurs for values of v - and y - larger (smaller) than μ when α is positive (negative). This result is illustrated in Figure 1 for users for whom $\alpha > 0$.

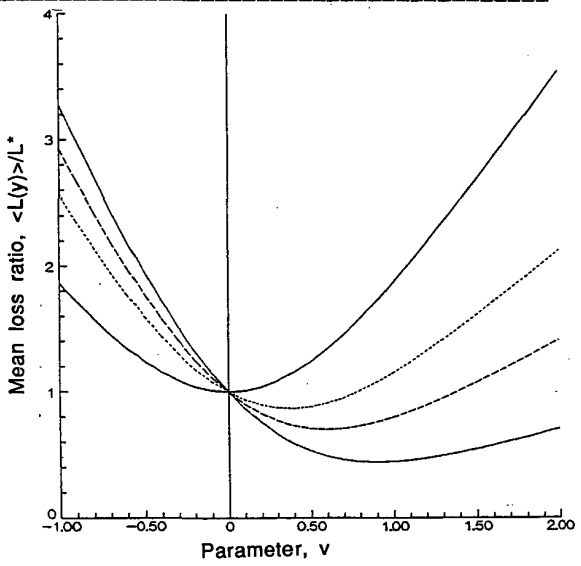


Figure 1. The ratio of mean losses, $\langle L(y)\rangle/L^*$, as a function of the parameter v, for selected values of the asymmetry parameter α. Key: $\alpha = 0$ (solid-heavy), $\alpha = 0.4$ (dotted), $\alpha = 0.6$ (dashed), $\alpha = 0.8$ (solid-light).

The climatologically optimal strategy can be found by differentiating $\langle L(y)\rangle$ in (6) with respect to v and setting the resulting partial derivative equal to zero. This process yields $\mathrm{erf}(v_C) = \alpha$ or $v_C = \mathrm{erf}^{-1}(\alpha)$, where v_C is the climatologically optimal value of v. In this case, $\langle L(y)\rangle = L_C$ (for "climatological"), where

$$L_C = L^*e^{-v_C^2} = L^*e^{-[\mathrm{erf}^{-1}(\alpha)]^2}. \qquad (7)$$

Note that $L_C \le L^*$. That is, the mean loss associated with the climatologically optimal strategy is always less than or equal to the mean loss incurred by choosing $y = \mu$, with equality only in the case of a symmetric loss function ($\alpha = 0$).

It can also be seen that v_C is negative (positive) when α is negative (positive), with $v_C = 0$ when $\alpha = 0$ (symmetric loss function). The ratio L_C/L^*, representing the mean loss associated with the climatologically optimal strategy divided

by the mean loss when $y = \mu$, is symmetric about $\alpha = 0$, with the largest value of L_c/L^* (= 1) occurring for $\alpha = 0$. Thus, the gains associated with following a climatologically optimal strategy increase as the asymmetry of the loss function increases (in a negative or positive direction).

b. Optimal decisions and economic value, with forecasts

Now suppose that a nonprobabilistic (or "point") forecast of the continuous variable X is available to the user, and let this forecast be denoted by z. We shall assume that x and z possess a bivariate Gaussian distribution, f(x,z), with means μ and η, variances σ^2 and τ^2, and correlation coefficient ρ. The assumption of a bivariate Gaussian distribution for x and z implies that z possesses a univariate Gaussian distribution f(z). Moreover, the conditional distribution of x given z, f(x|z), is also Gaussian with mean $\mu' = \mu + \rho(\sigma/\tau)(z - \eta)$ and variance $\sigma'^2 = (1 - \rho)^2\sigma^2$.

Under the assumption of a bivariate Gaussian distribution for f(x,z), it follows that the mean loss when the choice of y depends on the forecast z [i.e., $y = y(z)$] is $\langle L[y(z)]\rangle$, where

$$\langle L[y(z)]\rangle = L^*(1 - \rho^2)^{1/2}[e^{-w^2} + \pi^{1/2}we\,rf(w) - \pi^{1/2}\alpha w], \quad (8)$$

in which $w = (y - \mu')/[\pi^{1/2}(1 - \rho^2)^{1/2}\sigma]$. Note that this expression differs from the expression for the mean loss in the case in which the forecasts are not available (or are ignored), $\langle L(y)\rangle$, in two ways. First, the expression in (8) contains a factor $(1 - \rho^2)^{1/2}$ that characterizes the relationship between x and z. Second, v in (6) is replaced by w in (8); however, the functional dependence of (8) on w is the same as that of (6) on v.

In this context, several different strategies involving the use of the forecasts are available to the decision maker. Three such strategies are considered here: (1) a strategy involving "blind" belief in the forecast; (2) a strategy involving belief in a "regressionized" version of the forecast; and (3) a strategy involving optimal use of the forecast together with climatological information. The strategy involving blind belief in the forecast implies that $y = z$, and it leads to a mean loss $\langle L[y(z)]\rangle = L_b$ (for "blind"), where

$$L_b = L^*[1 - 2\rho(\tau/\sigma) + (\tau/\sigma)^2]^{1/2}. \quad (9)$$

[To obtain this expression, it is necessary to integrate (8) over all values of z. Recall that this variable possesses a Gaussian probability distribution.] The expression for L_b in (9) can be readily evaluated for various combinations of values of ρ and τ/σ. Here we consider only the case in which $\tau = \sigma$ (i.e., $\tau/\sigma = 1$). Such forecasts can be characterized as "detailed," in the sense that the variability of the forecasts is the same as that of the observations. In this case, $L_b = L_{bd}$ (for "blind" and "detailed"), where

$$L_{bd} = L^*[2(1 - \rho)]^{1/2}. \quad (10)$$

Note that $L_{bd} \geq L^*$ unless $\rho > 1/2$. That is, only when the correlation between the forecasts and observations exceeds 0.5 is the mean loss associated with a strategy of blind belief in detailed forecasts ($y = z$, with $\tau = \sigma$) less than the mean loss associated with the strategy of choosing the mean value of the underlying variable ($y = \mu$).

Following a strategy involving belief in a regressionized version of the forecast leads to a different result. In this case, $y = \mu + \rho(\sigma/\tau)(z - \eta)$, and the mean loss $\langle L[y(z)]\rangle = L_r$ (for "regressionized"), where

$$L_r = L^*(1 - \rho^2)^{1/2}. \quad (11)$$

Note that $L_r \leq L^*$, implying that the use of such forecasts generally leads to smaller mean losses than the use of the mean value of the observations. On the other hand, L_r can exceed the loss associated with the strategy involving optimal use of climatological information, L_c [cf. (7)].

Finally, following a strategy involving optimal use of the forecasts together with climatological information leads to a mean loss of $\langle L[y(z)]\rangle = L_o$ (for "optimal"), where

$$L_o = L^*(1 - \rho^2)^{1/2}e^{-w_o^2} = L^*(1 - \rho^2)^{1/2}e^{-[erf^{-1}(\alpha)]^2}. \quad (12)$$

This expression is obtained by differentiating $\langle L[y(z)]\rangle$ in (8) with respect to w and setting the resulting partial derivative equal to zero. The expression for L_o in (12) involves two multipliers of L^* - namely, $(1 - \rho^2)^{1/2}$ and $e^{-[erf^{-1}(\alpha)]^2}$ - whose values are less than one and thereby reduce the magnitude of L_o (from L^*, its upper limit).

The mean losses associated with five different strategies are depicted in Figure 2 as a function of the correlation coefficient ρ for a situation involving a moderately asymmetric loss function ($\alpha = 0.4$). These strategies include: (1) use of the mean of the underlying variable (mean loss L^*); (2) the climatologically optimal strategy (mean loss L_c); (3) blind belief in the forecast (mean loss L_{bd}); (4) the strategy of belief in a regressionized forecast (mean loss L_r); and (5) the strategy involving optimal use of the forecasts together with climatological information (mean loss L_o). Several conclusions can be drawn from the contents of Figure 2, and these conclusions are briefly summarized here:

(1) For situations in which the correlation coefficient ρ (a one-dimensional measure of the degree of association between the forecasts and observations) is less than 0.5, the mean loss associated with blind belief in the forecasts, L_{bd}, is greater than the mean loss associated with choosing the mean value of the underlying variable, L^*.

(2) Until the correlation coefficient ρ reaches a relatively large value, in this case about 0.6 (this value exceeds 0.5 in all cases), the mean loss associated with the climatologically optimal strategy, L_c, is less the mean loss associated with strategy involving blind belief in the forecasts, L_{bd}.

(3) Belief in a regressionized version of the forecast, with mean loss L_r, is always better than (or at least as good as) choosing the mean of the underlying variable, with mean

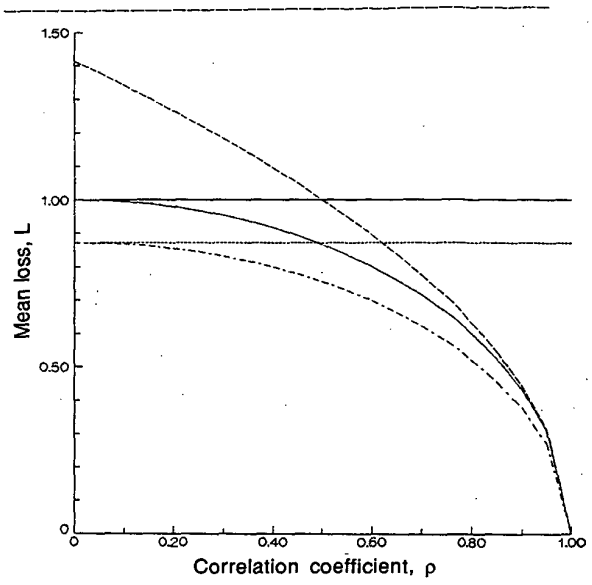


Figure 2. Mean losses L^* (line-solid), L_c (line-short dashes), L_{bd} (curve-long dashes), L_r (curve-solid), and L_o(curve-short/long dashes) as a function of the correlation coefficient ρ for $\alpha = 0.4$. See text for additional details.

loss L^*. However, a critical value of ρ exists, ρ_c say, such that belief in a regressionized forecast is worse than the climatologically optimal strategy when $\rho < \rho_c$. In this case, ρ_c is approximately 0.5.

(4) The strategy involving optimal use of the forecasts and climatological information, with mean loss L_o, is better than all of the other strategies. That is, $L_o \leq \min(L_r, L_{bd}, L_c, L^*)$ for all values of ρ, with equality only in the case of L_c when $\rho = 0$ (i.e., $L_o = L_c$ when $\rho = 0$).

These conclusions are based on results obtained under the assumptions of a bivariate Gaussian model for the forecasts and observations and a simple linear loss function. However, additional studies (by the authors) based on different assumptions have yielded qualitatively similar results.

5. Discussion and Conclusion

This paper has investigated the optimal use and economic value of weather/climate forecasts from the perspective of rational, information-sensitive users of such forecasts who want to minimize their expected (or mean) losses in the context of prototypical decision-making problems. Particular attention has been focused here on the relative merits, in economic terms, of alternative strategies involving the use and non-use of forecasts. Two simple decision-making problems have been considered: (a) a discrete problem involving two actions and two events and (b) a continuous problem involving an essentially infinite number of actions and events. The former is a generalization of the familiar cost-loss ratio situation. To obtain analytical results, the latter relies heavily on the use of a bivariate Gaussian probability model for the joint distribution of forecasts and observations and a simple parameterization of the loss function.

The results obtained in the discrete example indicate that some decision makers almost always exist who can benefit economically from using the forecasts. In particular, so-called ideal users whose loss matrices are such that the ratio of their meteorological losses corresponds closely to the odds ratio for the events realize the greatest benefit from the forecasts. On the other hand, the loss matrices of other users - especially users who show strong inclinations towards a strategy of always taking one of the two actions based on climatological information alone - may be such that they only realize disbenefits from using the forecasts. These users should ignore the forecasts and base their decisions on climatological information. However, if these users could modify their loss matrices in such a way as to approximate the loss matrices of ideal users, then they might be able to benefit from the forecasts. When forecasts from two forecasting systems are available (but cannot be combined), it can be seen that the use of traditional verification measures (e.g., a correlation coefficient or a skill score) as surrogates for measures of economic welfare is a practice fraught with danger. This example demonstrates that users exist for whom inferior forecasts (according to the verification measure) are superior in terms of economic benefit, and vice versa.

The implications of the results obtained in the continuous example (see Section 4a for a summary of these results) become quite vivid when what might be referred to as "equivalent correlation coefficients" are considered. For example, consider a situation in which $\rho = 0.6$ and α is given, and suppose that the optimal strategy of using the forecasts together with climatological information is followed. It might then be asked what value of ρ would be required for the strategy of blind belief in the forecasts to yield the same mean loss as the optimal strategy. The result is quite surprising. Even for loss functions with only moderate asymmetry (e.g., $\alpha = 0.4$), the value of this correlation coefficient, r_e say, must exceed 0.8 for the two mean losses to be equal. Thus, substantially higher correlations are required in order that users employing the blind-belief strategy attain the same mean loss as users employing the optimal strategy.

Two general conclusions can be drawn from this example. First, improvements in the strategies employed by users of weather and climate forecasts offer a way of enhancing the value of such information that is at least as promising as that of improving the quality of the forecasts themselves. Second, it is often profitable in an economic sense to smooth the forecasts. Moreover, it is usually profitable to increase the amount of smoothing as the correlation coefficient (or, more generally, the quality of the forecasts) decreases.

The analyses described in this paper -and the underlying decision-making problems - could be extended in many different ways. In the case of the discrete problem, we could consider situations involving additional actions and/or events. It would also be of interest to investigate the optimal use and economic value of other types of forecasts in this context, such as probabilistic forecasts. Although forecasts produced by two forecasting systems were considered in the context of this example, it was assumed that the forecasts could not be

combined. Relaxation of this assumption in future studies of this type would make it possible to assess the relative merits of a richer set of alternative sources of forecast information.

In the case of the continuous example, possible extensions include the consideration of other types of loss functions. For instance, we could employ a quadratic loss function (i.e., a loss function in which the loss is proportional to the square of the difference between the chosen value y and the observed value x) or a loss function that is bounded or truncated in some way for large (or small) losses. Alternatively, a loss function might be considered in which the loss depended on both the difference y - x and the value of x itself. In this way, situations in which (for example) losses are higher when severe weather/climate conditions prevail could be explored.

Other possible extensions of this example include the optimal use of forecasts of continuous variables from two or more forecasting systems. For instance, we could consider situations in which the dependence among such forecasts is taken into account. In addition, the effect of using decision criteria other than the minimization of mean (or expected) loss could be investigated. An example of an alternative criterion is the minimization of the likelihoood of large losses, a criterion that is more conservative than that employed here.

It should be evident by now that both producers and users of meteorological information can benefit in a variety of ways from studies of the optimal use and economic value of weather/climate forecasts. For example, such studies can identify the type of information that is most beneficial to specific users as well as the best available strategies (of employing this information) and the expected benefits of following these strategies. Unfortunately, disciplinary, bureaucratic, financial, and other obstacles exist that make it difficult to conduct such studies and to disseminate their results. The current unsatisfactory situation in this area can be illustrated by drawing an analogy with the field of medicine. Suppose that physicians only diagnosed diseases but did not attempt to cure them. Medical patients would then be in the same position that almost all recipients of weather/climate forecasts find themselves today - they would possess considerable information about the conditions of interest (although it might not be precisely the information that they need), but they would not be unable to make effective use of it!

Acknowledgments

L.S. Gandin is a University Corporation for Atmospheric Research scientist at the National Meteorological Center. A.H. Murphy was supported in part by the National Science Foundation under Grant SES-9106440.

References

Brown, B.G., R.W. Katz, and A.H. Murphy, 1986: On the economic value of seasonal-precipitation forecasts: the fallowing/planting problem. *Bulletin of the American Meteorological Society,* **67**, 833-841.

Katz, R.W., and A.H. Murphy, Editors, 1992: *Economic Value of Weather and Climate Forecasts.* Cambridge, United Kingdom, Cambridge University Press, in preparation.

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.

Murphy, A.H., 1977: The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation. *Monthly Weather Review,* **105**, 803-816.

Murphy, A.H., and B.G. Brown, 1982: User requirements for very-short-range weather forecasts. *Nowcasting* (K. Browning, Editor). London, United Kingdom, Academic Press, pp. 3-15.

Murphy, A.H., and H. Daan, 1985: Forecast evaluation. *Probability, Statistics, and Decision Making in the Atmospheric Sciences* (A.H. Murphy and R.W. Katz, Editors). Boulder, Colorado, Westview Press, pp. 379-437.

Murphy, A.H., and M. Ehrendorfer, 1987: On the relationship between the accuracy and value of forecasts in the cost-loss ratio situation. *Weather and Forecasting,* **2**, 243-251.

Obukhov, A.M., 1955: On the successfulness of alternative forecasts. *Izvestia of the U.S.S.R. Academy of Sciences* (Geophysical Series), No. 4, 339-350.

Thompson, J.C., 1962: Economic gains from scientific advances and operational improvements in meteorological prediction. *Journal of Applied Meteorology,* **1**, 13-17.

Wilks, D.S., and A.H. Murphy, 1986: A decision-analytic study of the joint value of seasonal precipitation and temperature forecasts in a choice-of-crop problem. *Atmosphere-Ocean,* **24**, 353-368.

Winkler, R.L., and A.H. Murphy, 1985: Decision analysis. *Probability, Statistics, and Decision Making in the Atmospheric Sciences* (A.H. Murphy and R.W. Katz, Editors). Boulder, Colorado, Westview Press, pp. 493-524.

Winkler, R.L., A.H. Murphy, and R.W. Katz, 1983: The value of climate information: a decision-analytic approach. *Journal of Climatology,* **3**, 187-197.

THE GREENHOUSE HYPOTHESIS: MODEL PROJECTIONS IN COMPARISON WITH OBSERVATIONAL STATISTICS

Christian - D. Schönwiese

J.W. Goethe University, Institute for Meteorology and Geophysics, Frankfurt a.M., FRG

1. INTRODUCTION

In recent years, the discussion of man-made global climate change due to the enhanced 'greenhouse' effect has enormously intensified (WMO, 1979; Bolin et al., 1986; Houghton et al., 1990). In preparation of the Second World Climate Conference (SWCC, 1990) and the U.N. Conference on Environment and Development (UNCED, 1992) the International Panel on Climate Change (IPCC) has prepared and recently updated a related state-of-the-art report (Houghton et al., 1990, 1992). It is evident from this background material that human activities lead to an atmospheric concentration increase of particular 'greenhouse' gases GHG (CO_2 etc.) and that this process enhances the natural 'greenhouse' effect. In 1991 the Mauna Loa CO_2 concentration was c. 355 ppm and the CO_2 equivalent where the effect of additional GHG is added to the CO_2 concentration was approximately 410 ppm. Following the IPCC scenario A (trend extrapolation, 'business-as-usual') a doubling of these CO_2 equivalents compared to the preindustrial level of approximately 300 ppm may occur roughly in the years 2025 - 2030.

Climate model projections, particularly coupled atmosphere-ocean general circulation models (GCM) predict a pronounced global climate change in response to this anthropogenic forcing, although due to the GCM shortcomings (which are discussed elsewhere) there are substantial uncertainties in respect to both the magnitude and the regional patterns of this response. In the following, a few selected aspects are addressed where GCM predictions are compared with observational climatic data statistics. Some aspects of natural climate fluctuations have also to be involved.

2. GLOBAL MEAN TEMPERATURE

Even in case of the GHG forced global mean temperature change, the GCM predicted magnitude varies from model to model. Fig. 1 (from IPCC, updated) illustrates not only the correlation of the predicted temperature and precipitation rise (global averages) but also their quantitative variety which covers in case of temperature 1.9 - 5.2 K for the equilibrium response to a CO_2 doubling. At the bottom of this Figure the related transient predictions are added specifying the temperature rise which may have realized until a CO_2 doubling: 1.3 - 2.3 K (Houghton et al., 1992).

It is a matter of fact that climate is not only forced by atmospheric GHG variations but also by an enormously complicated diversity of natural forcing mechanisms. In consequence, if one tries to detect anthropogenic change in observational climatic data, it is not justified to look simply on observational trends and their significance. It is possible, however, to reproduce a predominant part of the observed temperature variance by means of multiforced statistical regression models driven simultaneously by volcanic, solar, ENSO (El Niño, southern oscillation), and GHG parameter time series; details see elsewhere (Bayer and Schönwiese, 1992; Schönwiese, 1991; Schönwiese and Stähler, 1991). Furthermore, it is possible to evaluate hypothetically from such statistical models the enhanced 'greenhouse' effect and some other signals under consideration.

In Fig. 2 the reconstructed global mean temperature anomalies, land and marine data (IPCC; Jones, 1991) 1861 - 1985 are combined with the 1985-2085 prediction of the coupled MPI (Max Planck Institute for Meteorology, Hamburg, FRG; Cubasch et al., 1991) transient GCM experiment. Moreover, a reproduction and extrapolation of the enhanced GHG

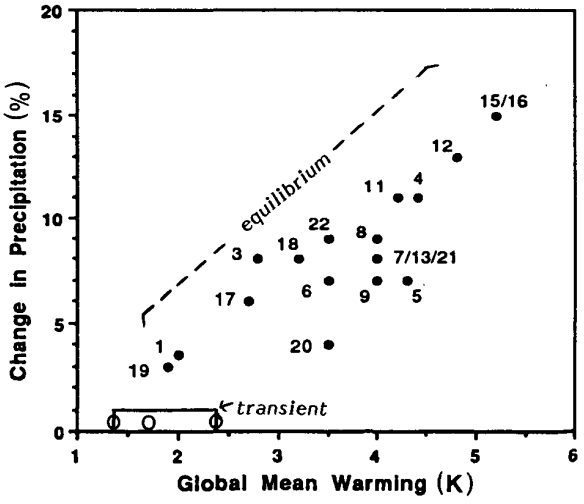


Fig. 1. Correlation of predicted global mean surface air temperature and precipitation rise, from 21 CO_2 doubling GCM equilibrium experiments, 20-22 high resolution, details see Houghton et al. (1990). At the bottom the temperature results of recent transient experiments are added (from Houghton et al., 1992; MPI 1.3 K, UKMO 1.7 K, GFDL 2.3 K).

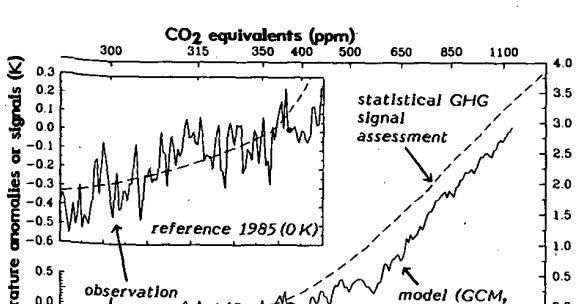


Fig. 2. Observed global mean annual surface air temperature variations 1861-1985 (from IPCC; Jones, 1988, 1991; land and marine data), transient IPCC scenario A GCM simulation 1985-2085 (MPI, Cubasch et al., 1991) and statistical assessment of the enhanced greenhouse effect, dashed line, explanation see text; from Bayer and Schönwiese (1992).

signal is added arising from multiple statistical assessments as mentioned before (GHG, volcanic, solar, and ENSO forcing, logarithmic GHG-temperature relation with a time lag of 20 yr; details see again Bayer and Schönwiese, 1992). The MPI model indicates a relatively low sensibility (1.3 K transient temperature response to a CO_2 doubling) so that this response is in between the low and best IPCC estimate (Houghton et al., 1990). The statistical assessment is very similar to the IPCC best estimate.

There is, however, another point of interest which should be taken into consideration. This concerns the background variability produced by the GCM and deviating from the long-term trend. If one computes the Pearson correlation coefficient (Spearman rank correlation very similar in this case) of the 1891 - 1991 observed and the 1985 - 2085 predicted time series, the result amounts to 0.71 but only 0.26 if detrended time series are compared. Fig. 3 disintegrates this result in terms of coherence spectra where again, for detrended data, the similarities are very poor. If one looks on the background variability which is produced by the GCM since 2025 (note outstanding change in both warming trend and variance) this correlation (observational data same sample size, 1925 - 1985) drops to an totally insignificant number of - 0.07 (rank correlation -0.16). It seems that substantial efforts are necessary to improve not only the prediction of the GHG signal but also the reproduction of the total variance.

Searching for more or less long-term trends in observations, which may be similar to GCM predictions, some additional statistical problems must not be neglected (details see Sneyers, this conference; Denhard and Schönwiese, 1992). Fig. 4 compares an usual linear trend analysis of the mean global temperature series (same as in Fig. 2) with the Mann-Kendall trend test. In this case the linear trend (0.5 K) and trend-to-noise ratio (2.6 where the 'noise' is represented by the annual data standard deviation) are confirmed by the nonparametric Mann-Kendall test which reveals an even higher level of confidence. There are, however, also some significant subinterval trends detectable, see again Fig. 4.

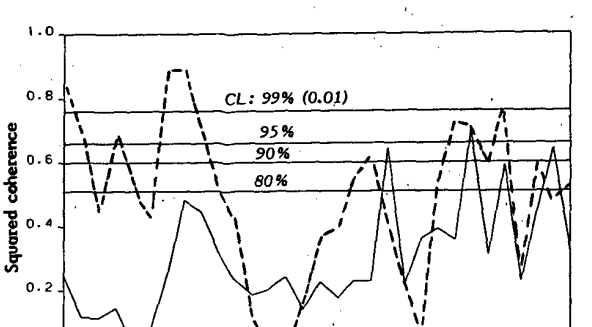


Fig. 3. Squared coherence spectra including confidence levels (CL) of the observed 1891-1991 and predicted 1985-2085 temperature variations, see Fig. 2, solid line detrended data, dashed line including trend (original data).

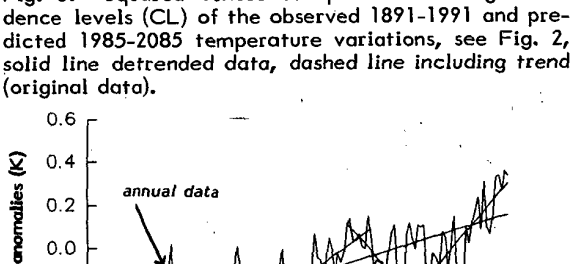


Fig. 4. Trend analysis of the observed global mean temperature variations, land and marine data as shown in Fig. 2, period 1861-1991. The linear trend amounts to 0.5 K, trend-to-noise ratio 2.6. The Kendall trend statistic Q values (details see Denhard and Schönwiese, 1992; Q can be interpreted as modified standard deviatiation factor) are as follows. 1861-1991: +10.1; 1865-1893: -3.14; 1903-1945: +6.15; 1937-1950: -3.02; 1964-1991: +4.69.

Generally, all observational time series statistics vary in time and space. As far as the spatial behaviour is concerned, some selected aspects are briefly addressed in the next section. In respect to all statistics, especially trends, one should not forget that these trends per se cannot tell us anything about forcing.

3. REGIONAL TRENDS

It was already mentioned that regional GCM predictions are much more uncertain than their global averages (Houghton et al., 1990; Schlesinger, 1991). Even in case of temperature which is a relatively 'certain' element in model projections and observations serious problems arise. One of these problems is illustrated in the Figures 5 and 6. Fig. 5 shows the observed linear summer temperature trends (again surface) 1890 - 1985 derived from the data set of Hansen and Lebedeff (1987) in a form where the globe is subdivided into c. 100 subareas. Note that this aspect shows Europe (except Scandinavia) as a cooling area whereas a more regionalized analysis based on c. 150 station records reproduces a quite different impression of European

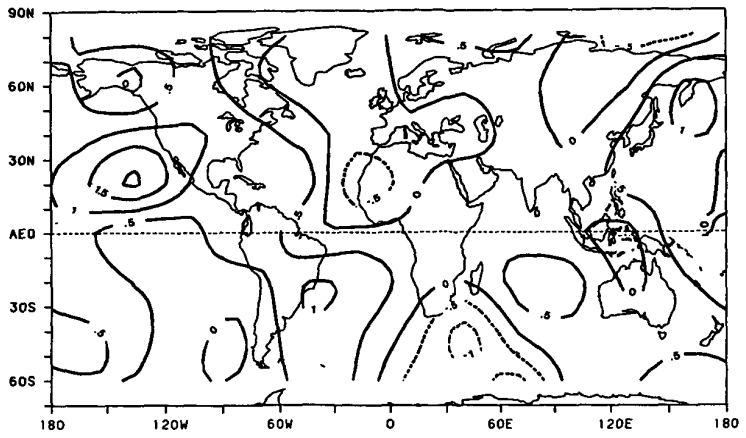


Fig. 5. Observed linear trends of surface air summer temperatures 1890-1985 (in K) based on the Hansen and Lebedeff (1987) global data set; from Schönwiese and Stähler (1991).

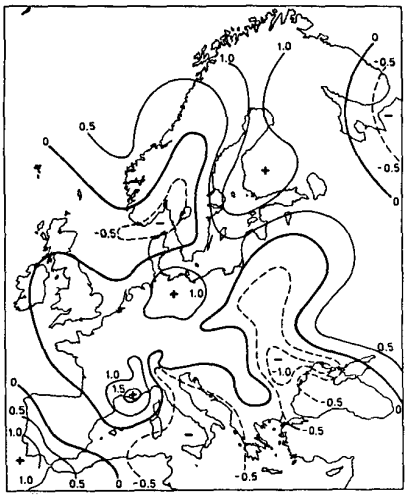


Fig. 6. Observed linear trends of surface air summer temperatures 1889-1988 (in K) based on c. 150 European station records.

temperature trends (Central Europe now warming), see Fig. 6. The trend-to-noise ratios are often poor but it seems that a corresonding Mann-Kendall trend test can reveal more confident results. In the moment, at Frankfurt University (FRG) an European trend atlas is in preparation which considers in addtion to temperature also precipitation, pressure and humidity (Schönwiese at al., 1992b), annual, seasonal and monthly trends. It may be possible or probable that a 'downscaling' of GCM experiments will lead also to different regional patterns of climate change.

If one considers zonal averages of temperature, again on a global observational scale and based on the Hansen and Lebedeff data set (1987) as before, the maximum trends appear in the arctic winter (> 5 K) exceeding a trend-to-noise ratio of 2; this is also the case in the c. 20° N zone spring and winter. What are, however, the reasons of these trends?

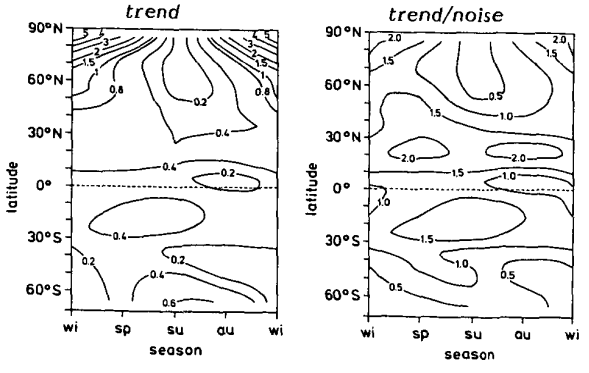


Fig. 7. Observed linear trends of surface air temperatures 1890-1985 (in K) based on the Hansen and Lebedeff (1987) data set as a function of season and latitude and related trend-to-noise ratios; from Schönwiese and Stähler (1991).

4. REGIONAL SIGNALS

Coming back to the multiple (multiforced) statistical regression technique addressed in Section 2 and using this technique on a regional and seasonal scale, the zonally averaged results are as follows. The temperature signals which may be hypothetically forced by an atmospheric CO_2 increase 280 (pre-industrial) to 350 ppm (roughly nowadays) reveal a pattern, see Fig. 8, which is very similar to the observed trends (Fig. 7) but the signals exceed a little bit the magnitude of the trends. At the same time this pattern is qualitatively very similar to the prediction of some GCM models (see e.g. Schlesinger, 1991; Houghton et al., 1990).

The corresponding maximum signals which may have been forced by explosive volcanic activity, see again Fig. 8, show once more a similar pattern but roughly half of the magnitude compared with the GHG signals (further details see Cress and Schönwiese, 1992). A corresponding analysis of solar signals (solar activity, not shown, see Schönwiese et al., 1992a) leads again to a similar pattern where the signals, however, are smaller than the volcanism-induced signals.

5. CONCLUSION

Only a few aspects of the greenhouse hypothesis have been discussed in this paper. It is evident, however, that statistics are urgently needed for validation of GCM control experiments, verification of their predictions and GHG signal detection in the observations. It is very important for the future that not only GHG but also natural forcing has to be taken into account which is in competition with GHG forcing (similar scale in time and space). Therefore, more efforts should be focussed on multiple (multiforced) simulations and assessments in both areas of research, GCM modelling and statistics. A detection of the GHG signal in observations which would be highly significant in a statistical sense is not or not yet possible but some related statistical hypotheses are on the way.

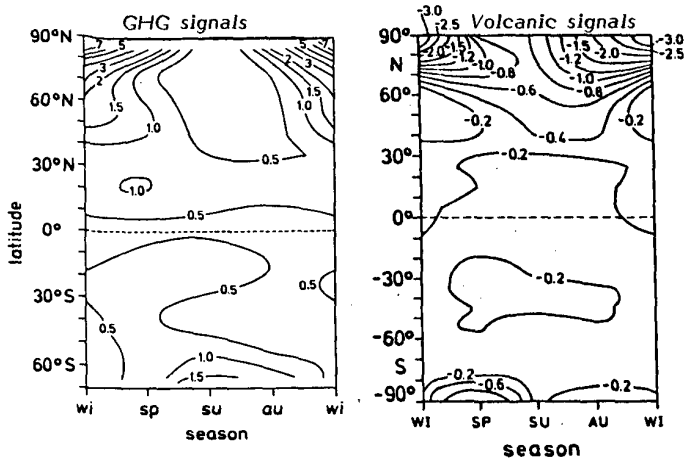


Fig. 8. Comparison of greenhouse-gas (GHG) and volcanism-induced surface air temperature signals (in K) derived from the data shown in Fig. 7 (multi-forced statistical assessments, method see text). The GHG signal refers to a CO_2 increase from 280 (preindustrial) to 350 ppm.

REFERENCES

Bayer, D. and C.D. Schönwiese, 1992: Hypothetical explanation of natural and anthropogenic temper ature variations by a nonlinear multiple regression model. Paper presented at the Internat. Meeting on Statist. Climatol., Toronto.

Bolin, B. et al., 1986: The Greenhouse Effect. SCOPE 29. Wiley, Chichester, 541 pp.

Cubasch, U. et al., 1991: Time-Dependent Green house Warming Computations With a Coupled Ocean-Atmosphere Model. Max-Planck-Institut für Meteorologie, Hamburg, Report No. 67, 18pp.

Cress, A. and C.D. Schönwiese, 1992: Statistical signal and signal-to-noise assessments of the seasonal and regional patterns of global volcanism temperature relationships. Atmósfera, **5**, 31-46.

Denhard, M. and C.D. Schönwiese, 1992: Non-parametric trend statistics and rank correlations of long European sea level pressure time series. Paper presented at the Internat. Meeting on Statist. Climatol., Toronto.

Hansen J. and S. Lebedeff, 1987: Global trends of measured surface temperature. J. Geophys. Res., **92**, 13345-13372.

Houghton, J.T. et al., 1980: Climate Change. The IPCC Scientific Assessment. Cambridge Univ. Press, Cambridge, 365 pp.

Jones, P.D., 1988: Hemispheric surface temperature variations: recent trends and an update to 1987. J. Climate, **1**, 654-660; priv. comm., 1991.

Schlesinger, M.E. (ed.), 1991: Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations. Elsevier, Amsterdam, 615 pp.

Schönwiese, C.D., 1991: A statistical hypothesis on global greenhouse-induced temperature change. Theor. Appl. Climatol., **44**, 243-245.

Schönwiese, C.D. and U. Stähler, 1991: Multiforced statistical assessments of greenhouse-gas-induced surface air temperature change 1890-1985. Clim. Dyn., **6**, 23-33.

Schönwiese, C.D. et al., 1992a: Inst. Meteorol. Geophys. Univ. Frankfurt (FRG), Report No. 92 (Solar climate relations, in German), in the press.

Schönwiese, C.D. et al., 1992b: An European Climate Trend Atlas (in German), in preparation.

World Meteorol. Org., 1979: Proceedings of the World Climate Conference. WMO Publ. No. 537, Geneva.

* * * *

Acknowledgement. *This study was supported by the German Government Climate Research Programme (BMFT, Project Number KF 2012 8).*

USE AND MISUSE OF STATISTICAL METHODS FOR THE DETECTION OF CLIMATE CHANGE

R. Sneyers

Institut Royal Météorologique de Belgique,
Brussels Belgium

1. INTRODUCTION

Though the evidence of the existence in the past of climate changes (Le Roy Ladurie, 1967; Lamb, 1972), until the 70es, it was believed among the majority of climatologists that our climate was stable. Even if, already in 1947 and 1948, Lewis, Lysgaard and Vandenplas put the attention on a possible warming of the climate during this century, the general belief was that the observed temperature increase resulted from urbanization around the observatories (heat island effect).

Witnesses of this opinion are i.e. Dettwiller-'s paper on the air temperature evolution in Paris (1970) and Prof. Landsberg last activities in observational experiences on the heat island effect (WMO-CCL Session 1982).

The discovery of the growth of the concentration of trace gases in the atmosphere having a greenhouse effect, growth confirmed especially in the case of the CO_2 concentration, and the first predictions produced by the GCMs of an enhanced greenhouse effect, changed completely this opinion with the result of a general jump of the climatologists into the problem of detection of the induced temperature increase in the climatological series of observations.

Actually, due to the random character of interannual variability, this search for temperature change is a statistical problem and the literature of the subject shows that, if some valuable work has been done on this topic, correct use of misleading methods or unproper use of appropriate ones presented in seminars, conferences and even in published papers, often led to dubious or to weakly significant results.

The error is not necessarily due to a lack of experience in the case of beginners in these procedures, but affects even experienced scientists.

The aim of this paper is to simultaneously report on and point out the weaknesses of the statistical procedures used for solving the problem of climate change detection during the last fourty years. It is completed with recent results given by what we think to be the rational way for solving the problem.

2. GENERALITIES

Synthesis of the statistical properties of large sets of observations is the fundamental aim of statistical analysis. To reach this aim, two quite different ways may be used: the empirical one and the theoretical one.

The empirical one produces general descriptions by means of what we shall call empirical methods. Examples are: averages, average square root deviations, moving or weighted averaging now commonly called filtering, empirical correlation or serial correlation, Fourier analysis or more generally spectrum analysis, least squares adjustment, principal component analysis or empirical orthogonal functions, and other still more sophisticated procedures.

The object of these methods is often to eliminate from the series the so-called "noise" in order to extract some special signal, forgetting that noise is an unseparable characteristic of a statistical phenomenon.

On the other hand, the theoretical method is based on the theory of mathematical statistics which gives the way for statistical inference, way involving testing of hypothesis for adequate modelling and for fixing the size of the estimation errors when adjusting the chosen model.

At this stage of statistical procedures, different misuses are already to be mentioned. They affect either the use of the empirical procedure or the interpretation of the result of statistical tests when applying the theoretical one.

The first mistake rises from confusing empirical and theoretical analysis. It consists in accepting the description given by the analysis without discussing the influence of the estimation errors in the properties apparently revealed by this analysis and thus without testing neither their validity nor their predictive power.

For the second case, undue conclusions may follow from incomplete analysis of the total sample or from unproper use of significance levels.

As an example of incomplete analysis, let us remember that if an underlying hypothesis is involved in a test of hypothesis, when this underlying hypothesis is not fulfilled, this may lead to rejection of the tested hypothesis though it is true.

For significance levels, considering α_0 as the adopted critical level and α the level corresponding to the observed value of the test statistic for a given sample, the null hypothesis is then rejected as soon as $\alpha < \alpha_0$.

It is obvious that consistence of statistical analysis implies that the chosen critical level remains constant once for all.

This means that when the analysis consists in multiple testing as i.e., when simultaneously testing independent subsamples or independent hypotheses on the same sample, proceeding to a particular test should take account of this pecularity by a convenient modification of the corresponding critical level.

In this special case, the reason for this caution is that when adopting for each subsample the level α_0 as critical level, the repeated application of the same test to each of the subsamples

increases the probability of finding under the null hypothesis a significant result, which means that the fixed critical level of significance for the total sample is no longer α_0.

3. THE SEARCH FOR CLIMATE CHANGE

The best means to review the different procedures to detect such changes is to remember the way covered since the first searches to solve this problem.

3.1 Moving averaging and filtering

Moving averaging or more generally filtering is the oldest procedure used for characterizing the climate evolution. In fact, the first assumptions born in the 1940es of a warming climate were based on the results provided by this method and even nowadays it is still used for this purpose (see i.e. Lewis, 1947, Lysgaard, 1948, Vandenplas, 1948 and a 1991 WMO Bulletin).

This method is mostly applied in order to damp the interannuel variability in view of getting a more clear look on the long term evolution of the non random component of the series.

It consists by replacing in the series each value x_i by the value x'_i defined by the relation

$$x'_i = \Sigma\, a_j x_{i+j}, \text{ where } j=-k, -k+1, \ldots, k \qquad (1)$$

the sum of the coefficients a_j being 1.

Moreover, when a_j is constant, the method reduces to the moving averaging one.

To point out the misleading character of this procedure, consider the case of a simple random series i.e. the elements of which are independently and identically distributed (i.i.d.).

If a is the mean of the distribution, the model defining the random process is then defined by the relation

$$x_i = a + e_i \qquad (2)$$

where e_i is i.i.d. with zero mean and zero autocorrelations.

Applying the procedure gives rise to the new relation

$$x'_i = a + e'_i \qquad (3)$$

e' being related to e_i in the same way as x' to x.

Computing variances and covariances of e'_i gives

$$\text{var } e' = (\text{var } e)/(2k+1) \qquad (4)$$

while from the autocovariances it appears that autocorrelations are

$$r_j = 1 - j/(2k+1) \text{ for } j < 2k+1$$

and

$$r_j = 0 \qquad \text{for } j \geq 2k+1. \qquad (5)$$

These relations remain the same for x'.

Moreover, if for a series of size n, we consider Bartels' equivalent number of repetitions Ω_n defined by the relation (see Sneyers, 1975 or 1990, p.168 and 174)

$$\text{var } \hat{a}'_n = \Omega_n\, (\text{var } x')/n, \qquad (6)$$

where $\hat{a}'_n$ is the average of the series of x'_i,

putting $l = 2k+1$, we find, after some reduction

$$\Omega_n = 1 + (l-1)[(1/n - (l+1)/nl] \qquad (7)$$

which with (6) and neglecting the higher order term in nl gives

$$\text{var } \hat{a}'_n = (l/n)\, \text{var } x'$$

With var x' = (var x)/l, we have finally

$$\text{var } \hat{a}'_n = (\text{var } x)/n. \qquad (8)$$

The conclusions are

(a) as expected, moving averaging leads to no improvement in estimating the non random component;

(b) on the contrary, the procedure introduces an artificial smoothing effect giving an apparent look of the climate evolution quite independent of the non random component and, at the same time, modifies the true properties of this component when non linear.

In particular, this explains the fact that the occurrence of exceptional high values at the beginning or at the end of the series may give to the series either a decreasing or an increasing effect, appearance to be assigned uniquely to the analysing procedure (see i.e. Parkinson, 1989).

It follows that conclusions based on the results of this procedure are completely misleading and thus, that this method has to be avoided both for representing the climate evolution as for the search for climate changes.

3.2 Testing the homogeneity of two independent samples. Student's t-test.

Climate change destroys the homogeneity inside independently distributed series of observation.

In the special case of a resulting division of the total series into two independent series, Student's t-test may be considered as an appropriate way for the purpose of climate change detection.

Note, however that, if the normality assumption for the distribution of the elements of the series is not a drawback for its use, owing to the robustness of the test, the non homogeneity of the sample variances may introduce a bias in its result.

Moreover, if for a given critical level of significance, its use gives valid conclusions in the case of a single series, for instance to the series of annual values, its simultaneous application to seasonal or monthly series without adjusting the corresponding critical level of significance may lead to invalid conclusions.

As an example, let us mention that comparing by means of this test the period 1939-1953 of available observations for Brussels with the reference period 1901-1930 (Sneyers, 1954), led correctly, at the 5% level, to the conclusion of a significant increase of the annual mean of the daily maximum of the air temperature.

On the other hand, the detection of three seasons and six months affected by a significant increase at the same level, if confirming by their number the existence of a climate change at the scale of the year, it did not allow a similar conclusion for a fixed season or month.

For the selection, the correct level α adjusted to the global level α_0 is then given by the

relation

$$(1-a)^m = 1 - a_0 \qquad (9)$$

m being equal to 4 or to 12, respectively in the case of seasons or months.

The correct selection leads then, at the global 5% level, to a significant increase for each of the three seasons (spring to autumn), but for only one among the six previously selected months.

Another interesting application of the t test, in its Fisher variance ratio form, for the detection of climate change was made for the period 1874-1945 by Vialar (1952) for the air temperature at Paris.

The method consisted in applying the test to the two partial series defined by a point running from the beginning to the end of the series. The first point corresponding to a significant value of the test statistic has then been considered as the beginning of a climate change, the 5% significance level remaining the same for each trial.

Obviously, this procedure involves two misuses:

(1) no account is taken for the repetition of the test;

(2) the fact that a dividing point leads to a significant value of the test statistic does not mean that it is an outlier for the first of the two series; thus, it may not be considered as the starting point for a climate change.

In the particular case of a change point dividing the series into two homogeneous random series with different means, the ML estimate of the change point is given by the point for which the significance level is least.

For this case, Vialar's procedure reduces to the parametric change point estimation developed by Hawkins (1977) and Worsley (1979).

3.3 Least squares adjustment of a regression model.

If in a process, the value y of an element depends exhaustively of a set of values of k elements x_1, x_2,... , x_k, this dependence may be characterized by a relation of the form

$$y = f(x_1, x_2, \ldots, x_k) \qquad (10)$$

Moreover, if the analytical form of the function (10) is well known, except the values of parameters involved in this function, estimates for these parameters may be achieved by applying the least squares procedure to a set of simultaneous observations of the elements y and x_j, j = 1, 2,... , k.

In this case, the true values differing generally from the observed ones, for the observed values, the correct relation is then

$$y = f(x_1, x_2, \ldots, x_k) + e \qquad (11)$$

were e involves both the errors of observations on y and on f.

When using this procedure, the covariance matrix of the estimation errors on the parameters may be well determined under the condition that the error e is a random one, identically and independently distributed following a normal distribution.

Two kinds of traps may lead to dubious conclusions: incorrect evaluation of the statistical properties of the error e, or non exhaustivity of the relation (10).

For the statistical properties of e, the normality assumption is of minor importance in the case of large samples, due to the asymptotic normality of the errors on the parameter estimates.

On the contrary, the independence of e is a major assumption, the absence of which may introduce strong biasses into the conclusions.

Moreover, when testing the independence of e, care has to be taken of the fact that the distribution of the serial correlation coefficient depends on the number of parameters under estimation (see Kendall, 1976, p. 164).

In the case of linear regression, the distribution of a transform of this coefficient, which is actually the inverse of the old Abbe criteria for homogeneity (Conrad, 1925, Sneyers, 1957), has been studied by Durbin and Watson (1950, 1951 and 1971).

For avoiding errors of inference due totally or partly to non exhaustivity or to non appropriate choice of relation (10), the best way is to divide the information into two independent parts, one being used for estimation and the other, for testing the goodness of fit of the adjusted model (Panofsky and Brier, 1958, p.179).

Note that if the condition for appropriate choice of model (10) is major in the case of non linearity, when the partial derivatives of f are finite, this condition weakens for small departures of the observations from corresponding central values of the elements. On the contrary, for large departures, the linear approximation is misleading.

In the case of climate change research, special applications generally used of this method are linear trend detection, search for enhanced effects on climate and Fourier or spectral analysis.

3.4 Linear trend detection

In this case, the function f reduces to a linear function of the time or, the considered values being generally annual ones, of year i corresponding to the observation y_i:

$$y_i = a_i + e_i \qquad (12)$$

Testing the stability of the series comes then out to testing the hypothesis a = 0, test which may be reduced to testing zero correlation between y_i and i.

Our first use of this method showed the existence of significant increasing trends for the series beginning in 1885 and ending at years varying from 1930 to 1950, for Brussels-Uccle, Greenwich-Kew, Paris-St-Maur, Frankfurt a/Main, Kleve and De Bilt (Sneyers, 1956). However, Anderson's serial correlation test (1942) applied to the deviations of the series from the linear model, did not allow to conclude to the linearity of the trends.

This method is still largely in use in the climatological literature without testing the randomness of the deviations, the most awfull manner being the one of applying it to the moving averages of the global air temperature data set prepared by Jones (1985, 1989).

3.5 Search for enhanced effects on climate.

In this case, y is the global average of the

air temperature and f is a linear function of variables representing different effects among which the CO_2 concentration.

The first caution to keep in mind is that the adjusted linear regression may be but an approximation of a true non linear relation. Extrapolation for large deviations from the observed values may then lead to dubious predictions.

Another point worth to be mentioned is that air temperature and CO_2 concentration series being simultaneously growing with the year, the fact that this last variable is not taken into account in the model modifies probably the detected correlations.

The same occurs, for instance, with the correlation between the monthly sunshine and the corresponding air temperature. It is well known, that in our country, sunshine is differently correlated with temperature for January and for July. However, correlating the observations of twelve consecutive months leads to the opposite result due to the simultaneous strong seasonal variation for the two variables.

This spurious result is thus a special case of non exhaustivity of the adjusted model.

3.6 Fourier analysis and spectrum analysis.

Already during the 19th century the search for periodicities in the climatological series has been a topic of great interest for climatologists. In particular, Köppen (1873) tried to detect in these series the influence of the variation of the solar activity, characterized by the Wolf sunspot number.

The method now mostly used for this purpose is spectrum analysis, which generalizes the Fourier analysis.

The fact that detection of the possible periodic components of the series is based on the spectral density which gives the relative contribution to the total variance for each argument α corresponding to any given periodicity, reduces the problem of significance of these components to the search for an adequate multiple test.

For the difficulties raised by this problem, the best reference remains still the comments on this subject made by Kendall and Stuart (Vol. 3, 1966).

A way for avoiding this difficulty is to note that the results given by the usual Fourier analysis for a given period (i.e. the seasonal variation) are, in fact, those procuced by the adjustment by least squares of the total Fourier development to the series (Sneyers, 1975 and 1990). It follows that the periodicites existing in a series being more clearly apparent in the empirical autocovariance function, at least in its first part, the graph of this last function may be used for detecting the possible existing periodic components, least squares being then used for estimating and testing significance.

This procedure was applied to the series of the Wolf sunspot numbers (Sneyers, 1976) and its results were confirmed by a sequential spectrum analysis (De Meyer, 1981).

However, testing these results with an independent series (Sneyers and Cugnon, 1986) reduced the number of estimated periodic components from 26 to three ones explaining only 48% of the total variance, the remaining been divided into an autoregressive process (37%) and its associated i.i.d. random variable (15%).

In fact, the serial correlation affecting the residual was responsible for overestimating the number of significant components.

In any case, this gives simultaneously a good example for both the possibilities of the method and of the usefulness of Panofsky and Brier's recommendation.

3.7 Cumulative sums of deviations (CUSUMS).

The method consists in computing the deviations

$$e_i = x_i - m(x) \qquad (13)$$

between the element x_i of a series with its general average $m(x)$.

The sums $S_k = \Sigma e_j$, where j is extended from 1 to k, are graphically represented in function of k, for k = 1, 2,... , n, and abrupt changes in the mean of the series are then characterized by slope changes in the progression of the graph.

Craddock (1979) used this procedure for the verification of the homogeneity of climatological series by applying it to the differences between the observations of stations under control and the corresponding ones of a reference station.

Abrupt climate changes were also detected by this procedure by Coops and Schuurmans (1986) for the air temperature at De Bilt (The Netherlands).

Though the method has been found especially useful for the detection of possible heterogeneities, the search for confirmation being then possible by archives examination, the absence of a selection test for separating random changes from significant ones makes it unsuitable for climate change detection.

Note that the rank version of this procedure due to Pettitt (1979) for the case of a single change point gives the right answer to the mentioned drawback, when each of the two subseries are i.i.d.

4. NON PARAMETRIC STATISTICAL DETECTION OF CLIMATE CHANGE.

A general method avoiding the different above described drawbacks involves the simultaneous use of a trend test (Mann, 1945), a change point test (Pettitt, 1979) and a test of randomness against serial correlation (Wald and Wolfowitz, 1943), all the tests being non parametric ones (Sneyers, 1992).

In particular, for analyzing the internal structure of the series, the trend test is applied in a sequential manner. Moreover, the Wald Wolfowitz test being, for large samples, practically identical to Anderson's test, replacing in the statistic, the original values by the corresponding ranks as suggested in Sneyers (1975 and 1990) makes the test equivalent to the rank version of von Neumann's ratio test developed by Bartels (1982).

The conclusion appearing from the application of this method to long climatological series is that, except for local influence of increasing urbanization, climate changes seem generally to proceed in an abrupt manner.

Using the sequential version of the trend test combined with the test of randomness, this conclusion has been established for the series of temperature:

(a) at the seasonal scale, for Brussels and Paris (Sneyers, 1958) with changes round 1910 for winter and spring and round 1930 for summer and

autumn, changes confirmed in 1971 and 1989 for Brussels (Sneyers, 1971 and Sneyers et al.,1989).

Note that the purpose of the paper of 1958 was to show that a climate change was responsible for an apparent correlation between the seasonal air temperatures detected for Brussels and Paris;
(b) at the annual level, about 1920 for Basle (Switzerland), De Bilt, Brussels and Paris and for the Jones series for the Northern Hemisphere (Goossens, 1986 and Goossens and Berger, 1986).

The complete analysis led, at the seasonal scale, to the above mentioned confirmation for Brussels and to similar results (Sneyers et al., 1991) for the Austrian series homogenized by Böhm (1990) and for the series of averages computed for Pic du Midi (France) by Bücher (1990).

At the annual scale, confirmation of the existence of the change point round 1920 was given by Demarée (1990) for both the Northern and the Southern Hemispheres.

Moreover, the analysis of the Austrian series reveals that actually, the last abrupt increase counterbalances an abrupt decrease which occurred round 1828.

For the annual amount of rainfall, results to be mentioned are: an abrupt increase round 1910 for Brussels (Sneyers et al. 1989), and for Africa: abrupt droughts about 1971, for Upper Volta and Burkina Faso, but without using the trend test (Snijders, 1983 and 1986), for Mauritania round 1967 (Demarée, 1990) and on the contrary, an abrupt increase during the period 1961 to 1972 in Burundi (Demarée and Sneyers, 1991, unpublished paper) which might explain the increase of the flow of the river Zaïre, observed during this period (Bultot and Dupriez, 1987).

Note that, for Brussels, concomitant abrupt changes occurred for the variance of both temperature and rainfall.

5. CONCLUSION

At the level of statistical modelling, the mentioned results show that the simple random models may no longer be considered as representing correctly the climate evolution, but that at least bi-modality has to be taken into account.

Note, for this point, that an adequate answer seems to be given by dynamical stochastic models as the one considered by Demarée and Nicolis (1990).

For the more general climatological problem of climate change, the importance of such results is emphazised by the fact that for quality control of theoretical climate models, no other alternative exists to this control than statistical climate modelling.

At the same time, it shows the degree of poorness of results given by climate change analysis limited to hemispheric or global series, due to the considerable loss of information involved in such a procedure.

If still necessary, this leads us back to the too often forgotten necessity for the availability of world wide extended long homogeneous series of observations.

6. REFERENCES.

Anderson, R.L.,1942. Distribution of the serial correlation coefficient. Ann. Math. Statis., 13, 1.

Bartels, R., 1982. The Rank Version of von Neumann's Ratio Test for Randomness. J. Am. Statist. Assoc., 77, 40-46.

Böhm, R. 1990. Temperature trends of high altitude stations in the Austrian Alps. 21. Intern. Tagung Alp. Meteor., Engelberg, Switserland, Tagungsbericht, 1. Teil, 322.

Bücher, A., 1990. Un Siècle d'Observation de la Température en Altitude dans les Hautes Pyrénées, France. 21. Intern. Tagung Alp. Meteor., Engelberg, Switserland, Tagungsbericht, 1.Teil, 342-345.

Bultot, F. and Dupriez, G.L., 1987. Niveaux et débits du fleuve Zaïre à Kinshasa (Régime-Variabilité-Prévision). Acad. Royale Sc. Outre-mer, Sc. Techn., Mém. in-4°, Fasc.2, Bruxelles, 49 pp.

Conrad, V., 1925. Homogenitätsbestimmung meteorologischer Beobachtungen. Met. Zeitschr., 42, 482-485.

Coops, A.J. and Schuurmans, C.J.E., 1986. Detection of the CO_2-effect upon the climate of Western Europe. Theor. Appl. Clim., 37, 111-125.

Craddock, J.M., 1979. Methods of comparing annual rainfall records for climatic purposes. Weather, 34, 332-346.

Demarée, G., 1990. Did an abrupt global climate warming occur in the 1920s?, in: Contributions à l'étude des changements de climat. Inst. R. Météor. de Belgique, Pub. A, n°124, 32-37.

id., 1990. An Indication of Climate Change as Seen from Rainfall Data of a Mauritanian Station, J. Theor. Appl. Climat. 42, 139-147.

Demarée, G. and Nicolis, C., 1990. Onset of Sahelian drought viewed as a fluctuation induced transition. Quart. J. R. Meteor. Soc., 116, 221-238.

De Meyer, F., 1981. Mathematical modelling of the sunspot cycle. Solar Physics, 70, 259-272.

Dettwiller, J., 1970. Evolution séculaire du climat de Paris. Influence de l'urbanisation. Mémorial de la météorologie, Paris, 83 pp.

Durbin, J. and Watson, G.S., 1950, 1951, 1971. Testing serial correlation in least squares regression. Biometrika, 37, 409; 38, 159; 58, 1.

Goossens, Chr., 1986. Etude spatio-temporelle des données climatiques européennes. Dissertation doctorale, Louvain-la-Neuve.

Goossens, Chr. and Berger, A.,1986. Annual and Seasonal Climatic Variations over the Northern Hemisphere and Europe during the Last Century. Annales Geoph., 4, 383-400.

Hawkins, D.M., 1977. Testing a Sequence of Observations for a Shift in Location, Journ. Amer. Statist. Assoc., 72, 180-182.

Jones, P.D., 1985. Northern Hemisphere Temperature 1851-1984. Clim. Mon. 14, 2, 42-50.

id., 1989. Global temperature variations since 1861: The influence of the southern oscillation and a look at recent trends. Proc. 28th Intern. Astrophysical Colloquium, Liège, 287-314.

Kendall, M.G., 1976. Time-Series. 2nd ed. Charles Griffin & Co.Ltd. 197 pp.

Kendall, M.G. and Stuart, A., 1966. The Advanced Theory of Statistics, Vol. 3, London, Griffin, 552 pp.

Köppen, W., 1873. Über mehrjährige Perioden der Witterung, insbesondere über die elfjährige Periode der Temperatur. Öster. Meteor. Zeitschr., 8, 241.

Lamb, H.H., 1972. Climate, Present, Past and Future, 2 vol., Methuen & Co, Ltd, London.

Le Roy Ladurie, E., 1967. Histoire du climat depuis l'an mil. Flammarion, 366 pp.

Lewis, L.F., 1947. Variations of temperature in London 1764-1939. Meteor. Mag., 135-138.

Lysgaard, L., 1948. Recent climatic fluctua-

tions. Nature, 161, 442-443.

Mann, H.B., 1945. Non parametric test against trend. Econometrika, 13, 245-259.

Panofsky, H.A. and Brier, G.W., 1958. Some Applications of Statistics to Meteorology. The Pennsylvania State University, University Park, Pennsylvania, 224 pp.

Parkinson, C.L., 1989. Dangers of multilinear averaging in analyses of long-term climate trends. Climate Dynamics, 4, 39-44.

Pettitt, A.N., 1979. A Non-parametric Approach to the Change-point Problem. Appl.Statis. 28, N°2, 126-135.

Sneyers, R., 1954. Les caractères climatiques de la période 1939-1953 et le temps en 1953. Inst. R. Météor. de Belgique, Pub. B, n°11, 52pp.

id., 1956. Sur quelques propiétés statistiques de la température de l'air en Belgique, Inst. R. Météor. de Belgique, Pub. A, n°4. 62 pp.

id., 1957. Sur la détermination de l'homogénéité des séries climatologiques. Journ. Scient. Météor., VII, 28, 359-372.

id., 1958. Connexions thermiques entre saisons consécutives à Bruxelles-Uccle. Inst. R. Météor. de Belgique, Pub. B, n°23, 24 pp.

id. 1971a. La détermination de la stabilité du climat par l'analyse statistique des séries d'observations; un exemple: la température de l'air et l'eau recueillie à Bruxelles-Uccle de 1833 à 1969. Ann. Meteor., Neue Folge, 5, 205-208.

id. 1971b. Statistical analysis of temperature and rainfall data at Brussels-Uccle from 1833 to 1969. WMO, Proc. Symp. on Physical and Dynamic Climatology, Leningrad, 390-392.

id. 1975. Sur l'Analyse Statistique des Séries d'Observations, Note Technique N°143, Organisation Météorologique Mondiale, Geneva, 192 pp.

id. 1976. Application of least squares to the search for periodicities. J. Appl. Meteor. 15, 387-393.

id. 1990, On the Statistical Analysis of Series of Observations, Technical Note Nr143, World Meteorological Organization, Geneva, 192 pp.

id. 1992. On the use of statistical analysis for objective determination of climate change, accepted for publication in Meteorologische Zeitschrift.

Sneyers, R. and Cugnon, P., 1986. On the predictability of the Wolf sunspot number, Annales Geoph., 4, A, 1, 81-86.

Sneyers, R., Vandiepenbeeck, M., Vanlierde, R. and Demarée, G., 1989. Climatic changes as appearing from the homogenized series of observations made in Brussels-Uccle, in: Brazdil, ed., 1990, Climate Change in the historical and industrial Periods, Masaryk University, Brno, 170-174.

Sneyers, R., Böhm, R. and Vannitsem, S., 1991. Climatic changes in the Austrian Alps as appearing from the homogenized series of temperature for the period 1775-1989. Contributed paper presented at the VIth Italian Glaciological Meeting, Gressoney, Italy, 26-28 September 1991.

Snijders, T.A.B., 1983. A study of the variability in space and time of Upper Volta rainfall. II. International Meeting on Statistical Climatology, Sept. 26-30, 1983, Lisboa, 3.5.1-3.5.7.

id., 1986. Interstation Correlation and Non-stationarity of Burkina Faso Rainfall, J. Theor. Appl. Climat., 25, 524-531.

Vandenplas, A., 1948. Variation séculaire de la température à Bruxelles-Uccle. Inst. R. Météor. de Belgique. Misc. n°35, 16p.

Vialar, J., 1952. Etude statistique des températures de Paris-Saint-Maur pour la période 1874-1945, 4, 21, 6.

Wald, A. and Wolfowitz, J., 1943. An exact test for randomness in the non-parametric case based on serial correlation. Ann. Math. Statist., 14, 378-388.

WMO-Bulletin, 1991. The world climate system in 1990. 40, 3, 215.

Worsley, K.J., 1979. On the Likelihood Ratio Test for a Shift in Location of Normal Populations. Journ. Amer. Statist. Assoc., 74, 365-367.

J5.3 EAST AUSTRALIAN RAINFALL EVENTS: INTERANNUAL VARIATIONS, TRENDS, AND RELATIONSHIPS WITH THE SOUTHERN OSCILLATION.

Neville Nicholls and Alex Kariko

Bureau of Meteorology Research Centre
Melbourne, Australia

1. INTRODUCTION

Many studies have considered the relationship between Australian rainfall and the Southern Oscillation. Such studies have used monthly, seasonal or annual rainfall *totals* as their dependent variable. The general relationship, that drier than normal conditions usually occur during periods with large negative Southern Oscillation Index (SOI) values (El Niño episodes) is well known. None of the studies have, however, discussed how rainfall *events* might be related to the Southern Oscillation, i.e., are negative values of the SOI associated with fewer than normal rain events, or shorter events, or less intense events? Information about SOI relationships with such variables might lead to increased understanding of the mechanisms by which the Southern Oscillation affects Australia's climate.

Numerous studies have also documented historical rainfall trends in Australia. Again, these studies have concentrated on trends in total rainfall, e.g. in annual rainfall totals. Little attention has been given to trends in the parameters describing rain events, apart from Yu and Neil (1991) who examined the relationship between east Australian rainfall totals, high intensity rain events, and global temperatures. Some examinations of rainfall events in numerical models with doubled CO_2 (e.g., Mearns et al., 1990) have suggested that increased intensity of rain events might be expected from an enhanced greenhouse effect. Information about recent historical trends in rain events might provide a basis for determining if a particular regional rainfall change was due to an enhanced greenhouse effect. Thus, if the models predict greater intensity of rain events, and no evidence of this is found in the recent record, we might conclude that the recent changes are *not* due to the enhanced greenhouse effect, or that model predictions of increased intensity may be incorrect.

In this paper the secular trends in rain events and relationships between rain events and the SOI are examined for five stations in east Australia. The stations have been chosen because another study (Lavery et al., 1992) determined that their daily rainfall data were of high quality and were unlikely to have been contaminated through changes in exposure, observational apparatus or techniques, or shifts in location. The observers operating these stations were also determined, through a series of statistical tests and careful searches of station documentation, to have been accurate and thorough. A sample of only five stations was chosen for this initial examination to facilitate examination of the numerous relationships possible *viz:* inter-station correlations, intra-station correlations, relationships with the SOI, trends, etc. With a larger number of stations an overwhelming number of relationships would have resulted.

2. DATA AND METHODOLOGY

Figure 1. Locations of the five stations used in this study.

The locations of the stations selected for analysis are shown in Figure 1. They are all located in an area where rainfall is strongly affected by the El Niño - Southern Oscillation. Each station has daily rainfall data from at least 1910. The data record is almost complete at each station, with very few missing days, or periods when the rainfall was accumulated over more than a single day. As noted above, the data at these stations were considered to be very reliable. These data are, therefore, suitable for examination of the behaviour of rain events, where events are defined in terms of daily data. The data cannot, of course, be used to examine rain events of shorter duration.

For this study a rain event was defined as a continuous period with recorded rainfall on each day. The length of an individual event was the number of contiguous days

with recorded rainfall. Three variables were used to describe interannual variations in rain events:
Number - number of events in a year
Length - mean length of events (days)
Intensity - mean rain per rainday (mm)
These three variables were calculated for each year from 1910 to 1988, for each of the five stations. Only annual totals and averages were calculated for this study. The total annual rainfall in a year is the product of these three variables (i.e., Number*Length*Intensity) for that year.

The interannual variations of the three variables, and the annual rainfall, at the five stations were correlated with each other, with the corresponding variable at the other stations, and with year (to examine linear trends) and the SOI. Values of the SOI were provided by the National Climate Centre of the Australian Bureau of Meteorology, as were the original daily rainfall data. The results are discussed below in the following order: mean values of annual rainfall, and mean number, length, and intensity of rain events; intra-station correlations between the four variables; inter-station correlations for each of the four variables; correlations of each variable at each station with the annual SOI; and trends in each of the variables at each station. Unless otherwise stated correlations denoted as "significant" were statistically significant at the 5% level.

3. RESULTS

3.1 Mean values

The mean number of events ranges between 22 and 60, with the largest number at Carrick, the station located at highest latitude. The events tend to last about two days, and their intensity decreases polewards.

3.2 Intra-station relationships

As was noted above, the annual total rainfall is the product of the number, length and average intensity of rain events for the particular year. All three variables do not, however, contribute equally to the interannual variations in annual rainfall. Their respective contributions depend on the correlations between them, and their variability. In Table 1 the correlations of annual rainfall with the other three variables are listed for each station. At each station, the average intensity (i.e. the average rainfall per rainday) is the major determinant of the interannual variations in annual rainfall totals. The relative contributions of the average length of events and the number of events vary, with no obvious pattern.

At each station, correlations between all four variables were calculated. In general, there were no significant relationships between the other three rain-event variables (number, length, and intensity) were found. Thus, the number, length, and intensity of rain events can be considered to be independent variables at each station.

3.3 Inter-station correlations

Inter-station correlations between the four variables are listed in Table 2. The annual rainfalls at all stations are positively and significantly correlated with annual rainfall at all the other stations (except between Gatton-Lawes and Carrick where the correlation of 0.21 is significant at 7%). These inter-station relationships of annual rainfall result mainly from inter-station relationships in the numbers of rain events. All these correlations are positive, and most are significant.

The other two variables (length and intensity) reveal little consistency between stations. Very few of the relationships are significant and some are even negative. So, we can conclude that the large-scale nature of Australian rainfall fluctuations is largely caused by an influence operating on the number of rainfall events. However, most of the inter-station relationships for intensity and length are positive, although generally weak. This indicates that they are contributing to some extent to the large-scale relationships between rainfall totals at the various stations. This is also indicated by the fact that the correlations between the annual rainfalls are generally larger than the correlations between the number of events at the different stations.

3.4 Relationships with the Southern Oscillation

Correlations between the annual mean SOI and the four variables at each station are presented in Table 3. All correlations are positive, indicating that the El Niño - Southern Oscillation affects the number, length and intensity of rain events at all stations, as well as influencing annual rainfalls. The correlations with annual totals are the largest, followed at three stations by the correlations with the number of rain events. The correlations with length are small and not significant, except at Carrick. Carrick's relationships with the SOI are rather different from the other four stations. At Carrick the number of events displays the weakest correlations with the SOI, whereas elsewhere this variable is significantly correlated with the SOI. The influence of the SOI on the length of events appears to increase with latitude, only at Carrick reaching significance.

The correlations between the SOI and annual rainfall reflect the combined influence of the SOI on all three event variables: number, length, and intensity. At every station each of these three variables is positively correlated with the SOI and therefore contributes to the SOI-annual rainfall correlation. So the SOI does *not* just affect annual rainfall by influencing just one index of rain events (e.g. number of events).

3.5 Trends

Correlations with year were calculated for each variable at each station to identify significant linear trends. These correlations are listed in Table 4. There is no evidence of a general increase in intensity of rain events. In fact four of the stations show negative correlations between intensity and year. Four of the five stations exhibit significant positive correlations between year and the number of events. The weak positive trends in rainfall at most of the stations are due, in general, to these increases in the number of events. The increases in annual rainfall are weaker than is the case for the number of events

because decreases in the intensity of events tend to offset the increases in rainfall caused by increased numbers of events. This is particularly the case at Winton where the strong decrease in intensity completely offsets the increased numbers of events, resulting in no trend in annual rainfall.

Canary Island is the only station without a significant positive trend in numbers of events. It is also the only station with a positive trend in intensity and this trend results in a weak positive trend in rainfall total.

At Peak Hill and Gatton-Lawes there are significant positive trends in the length of events. These positive trends enhance the increased annual rainfall caused by the increases in the numbers of events. As a result the positive trends in annual rainfall are strongest at these two stations. Nicholls and Lavery (1992) identified positive trends in rainfall, mainly summer rainfall, in the region including these two stations.

4. DISCUSSION

4.1 Rainfall trends and the "greenhouse" effect.

As was noted earlier, some numerical model studies of the climatic effects of increased atmospheric concentrations of carbon dioxide have suggested that intensity of rain events might increase. In east Australia the intensity of rain events has not, in general, increased through the twentieth century, despite rainfall totals increasing. Indeed, intensity has tended to decrease, offsetting some of the increased rainfall due to higher frequency of rain events. Thus, the pattern of rain event trends does not appear to coincide with that expected from the atmospheric model studies, implying that the twentieth century rainfall trends in east Australia (see Nicholls and Lavery, 1992) may *not* be the result of an enhanced "greenhouse" effect. Yu and Neil (1991) examined high intensity rainfall events in eastern Australia and concluded that they did not exhibit a simple and consistent relationship with global temperatures and rainfall totals. High rainfall intensities have occurred during periods of low or average rainfall and low temperature, as well as high temperature and rainfall. Their results also tend to cast doubts on the concept that global warming will necessarily lead to increased intensity of rain events.

4.2 The Southern Oscillation and rain events.

The results noted earlier indicate that the Southern Oscillation produces the large-scale nature of east Australian rainfall fluctuations by influencing, primarily, the *number* of rain events. That is, during El Niño episodes, the number of rain events at most stations is generally below normal. The Southern Oscillation also affects the length and intensity of rain events at all the stations examined here (Table 3). However, other factors must also affect the length and intensity of rain events in different ways at each station, leading to only weak inter-station correlations between these variables. By contrast, the inter-station correlations of the number of events are significant for most station pairs, suggesting that other factors do less to disrupt the Southern Oscillation influence on the number of events.

The Southern Oscillation does, however, affect annual rainfall at *individual* stations through its influence on all three characteristics of the rain events: number, length, and intensity. The correlations of SOI with annual rainfall are greater than the correlations with the individual rain event characteristics. The strength of the SOI-rainfall correlations in this area forms the basis of the seasonal outlook service provided by the Australian Bureau of Meteorology. This service uses the significant lag correlations between the SOI and seasonal rainfall (McBride and Nicholls, 1983). It might be possible to improve the SOI-rainfall lag relationships (and thus the outlooks) by predicting number, length, and intensity of rain events separately and then combining the predicted characteristics to produce a "predicted" total rainfall.

5. CONCLUSIONS

This study of east Australian rain events at five stations has found that:

- ★ Annual rainfall variations are primarily caused by variations in intensity.
- ★ The number, length, and intensity of rain events are essentially independent of each other, at each station.
- ★ The large-scale geographical nature of east Australian rainfall fluctuations is mainly due to strong inter-station correlations in the *number* of rain events.
- ★ The Southern Oscillation affects rainfall mainly by influencing the number and the intensity of rain events. Its influence on the length of events is, in general, weaker.
- ★ Twentieth century increases in rainfall have been due, primarily, to increased numbers of rain events. Intensity of rain events has generally declined, offsetting some of the increase in rainfall expected from the increased numbers of events.

REFERENCES

Lavery,B.M., Kariko,A.P. and Nicholls,N., 1992. 'A high-quality historical rainfall data set for Australia', *Aust. Meteorol. Mag.*, (accepted).

Mearns,L.O., Schneider,S.H., Thompson,S.L., and McDaniel,L.R., 1990. 'Analysis of climate variability on general circulation models: comparison with observations and changes in variability in $2xCO_2$', *J. Geophys. Res.*, 95, 20469-20490.

McBride,J.L., and Nicholls,N., 1983. 'Seasonal relationships between Australian rainfall and the Southern Oscillation', *Mon. Weath. Rev.*, **111**, 1998-2004.

Nicholls,N., and Lavery,B., 1992. 'Australian rainfall trends during the twentieth century', *Int. J. Climatology*, (accepted).

Yu,B., and Neil,D.T., 1991. 'Global warming and regional rainfall: The difference between average and high intensity rainfalls', *Int. J. Climatology*, **11**, 653-661.

Table 1. Correlations of annual total rainfall with the number of rain events, their average length in raindays, and average intensity per rainday, for each station. All correlations significant at 5%. Data from 1910-1988.

Station	Number	Length	Intensity
Winton	0.55	0.64	0.71
Gatton-Lawes	0.48	0.54	0.59
Peak Hill	0.62	0.58	0.66
Canary Is.	0.63	0.34	0.71
Carrick	0.29	0.57	0.76

Table 2. Inter-station correlations. Correlations in each box are, in order, between annual rainfall totals, number of rain events, length of rain events, and intensity of rain events. Correlations significant at 5% level are underlined. Data from 1910-1988.

Station	Winton	Gatton-Lawes	Peak Hill	Canary Is.
Gatton-Lawes	0.26 0.47 0.12 -.11			*Format*
Peak Hill	0.42 0.29 0.03 0.15	0.57 0.32 0.32 0.02		*Annual* *Number* *Length* *Intensity*
Canary Is.	0.42 0.34 0.11 -.03	0.34 0.05 -.10 0.04	0.60 0.41 0.12 0.26	
Carrick	0.33 0.18 0.05 0.18	0.21 0.26 0.07 0.24	0.37 0.26 0.13 0.23	0.58 0.31 0.22 0.14

Table 3. Correlations of annual SOI with number, length and intensity of rain events, and annual total rainfall, at five stations. Correlations significant at 5% are underlined.

	Number	Length	Intensity	Annual
Winton	0.35	0.13	0.27	0.44
Gatton-Lawes	0.33	0.12	0.24	0.47
Peak Hill	0.30	0.15	0.38	0.44
Canary Is.	0.37	0.22	0.19	0.45
Carrick	0.19	0.37	0.41	0.58

Table 4. Correlations between year and the number, length and intensity of rain events, and annual rainfall totals, at five stations. Correlations significant at 5% are underlined.

Station	Annual	Number	Length	Intensity
Winton	-.03	0.29	0.02	-.29
Gatton-Lawes	0.26	0.36	0.38	-.21
Peak Hill	0.21	0.25	0.31	-.07
Canary Is.	0.16	0.00	-.20	0.28
Carrick	0.05	0.33	0.03	-.18

J5.4

CROSS-VALIDATED PREDICTION MODELS FOR 1 DECEMBER AND 1 AUGUST FORECASTS OF SEASONAL TROPICAL CYCLONE ACTIVITY IN THE ATLANTIC BASIN

P.W. Mielke, K. J. Berry, W.M. Gray and C.W. Landsea

Colorado State University
Fort Collins, Colorado

Prediction models for 1 December and 1 August forecasts have been developed for seasonal tropical cyclone activity in the Atlantic Basin (Gray *et al.*, 1992a,b). Seven distinct dependent variables [number of named storms (NS), number of named storm days (NSD), number of hurricanes (H), number of hurricane days (HD), number of intense hurricanes (IH), number of intense hurricane days (IHD) and hurricane destruction potential (HDP)] are predicted from common sets of five and nine independent variables for 1 December and 1 August forecasts, respectively, using 41 years (1950-1990) of data. The 1 December and 1 August prediction equations possess a common structure given by

$$y=\beta_0+\beta_1[a_1U_{50}+a_2U_{30}+a_3\,|\,U_{50}-U_{30}\,|]+\beta_2[a_4R_s+a_5R_g]$$

$$+\beta_3[a_6(SLPA)+a_7(ZWA)+a_8(SOIA)+a_9(SSTA)]$$

where β_0,β_1,β_2 and β_3 are ordinary linear model regression parameters (note that $\beta_3=0$ in the 1 December prediction equation); $a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8$ and a_9 are selected constants associated with the two and three physically related groups associated with the 1 December and 1 August forecasts, respectively; U_{50} and U_{30} are extrapolated Quasi-Biennial Oscillation (QBO) data consisting of 50 mb and 30 mb upper-air zonal winds, respectively; R_s and R_g are African rainfall data involving August to September Western Sahel and August to November Gulf of Guinea rainfall, respectively; and SLPA, ZWA, SOIA and SSTA are Caribbean Basin and El Niño-Southern Oscillation (ENSO) data involving the Sea Level Pressure Anomaly in the lower Caribbean basin for June and July, the Zonal Wind Anomaly in the Caribbean basin for June and July, the Tahiti minus Darwin sea level pressure difference (Southern Oscillation Index Anomaly) for June and July, and the Sea Surface Temperature Anomaly in Niño 3 for June and July, respectively. The method for obtaining cross-validated (jackknife) predictions involves [1] using Least sum of Absolute Deviations (LAD) regression and [2] grouping the independent variables into two and three groups having related physical properties for 1 December and 1 August forecasts, respectively. For three of the seven 1 December prediction models and six of the seven 1 August prediction models, the jackknife and non-jackknife results are almost identical (this feature holds for both the subsequently-defined measure of agreement and the squared Pearson correlation coefficient). Comparisons of LAD and Least sum of Squared Deviations (LSD) regression raise some unanticipated concerns regarding the almost universal use of LSD regression in various applications at the present time.

Let y and $\hat{y}$ denote the observed and predicted dependent variables, respectively. The measure of agreement (ρ) is defined by

$$\rho=(\mu_\delta-\delta)/\mu_\delta$$

where $\delta=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i|$ and μ_δ is the average value of δ over all $n!$ equally likely permutations of $y_1,\ldots,y_n$ relative to $\hat{y}_1,\ldots,\hat{y}_n$ under the null hypothesis that the $n(n=41)$ pairs (y_i and $\hat{y}_i$ for $i=1,\ldots,n$) are merely the result of a random assignment (Mielke, 1984, 1991). While non-jackknifed predicted values are based on all n collections of dependent and independent observations, the jackknifed predicted values are based on the $n-1$ collections of dependent and independent observations with the removal of the collection associated with the value being predicted (in this sense each of the n predicted values is cross-validated by being independently predicted from the remaining data). While the squared Pearson cross-product correlation coefficient (r^2) is strictly a measure of linearity (i.e., $r^2=1$ implies all observed and predicted value pairs fall on a line which does not necessarily have unit slope or pass through the origin, ρ is a measure of agreement and $\rho=1$ implies all observed and predicted value pairs fall on a line with unit slope which passes through the origin.

The intent of this presentation is to examine the influence on both the measure of agreement and the squared Pearson correlation coefficient if [1] LSD regression is utilized instead of LAD regression and/or [2] non-grouped (NG) prediction models are utilized instead of the grouped (G) prediction models (i.e., non-grouping or grouping of the independent variables of the prediction models). Complete descriptions of the dependent and independent variables of these models are described in Gray *et al.* (1992a,b). The independent variable constants for each group comprising the grouped prediction models are taken from the relative proportions for the non-jackknifed LAD and LSD regression fits. The results of this study are presented in Tables 1 and 2. While the

highest squared Pearson correlation coefficients are achieved by the non-jackknifed LSD regression model (as expected), the highest squared Pearson correlation coefficients for the jackknifed results are obtained by the grouped LAD prediction model for 13 of the 14 cases in question (the grouped LSD prediction model barely yielded a higher value than the grouped LAD prediction model for one case). The findings stated in the last sentence are definitely unanticipated. The highest measure of agreement coefficients are achieved by the non-jackknifed LAD regression model (as expected) and the highest measure of agreement coefficients for the jackknifed results are obtained by the grouped LAD prediction model for all 14 cases (not surprising in light off the previous findings). As indicated by Gray *et al.* (1992a), a search routine yielded even larger jackknifed values of both the measure of agreement and the squared Pearson correlation for four of the seven 1 December prediction models (this is due to the fact that maximizing the measure of agreement is a non-linear criterion whereas the LAD procedure is a linear criterion). The parenthetical results in Tables 1 and 2 are the jackknifed results for 1 December given in Gray *et al.* (1992a); no parenthetical result implies the LAD procedure yielded an identical result. As a matter of speculation, it appears that the LAD procedure yields an optimized measure of agreement when the jackknifed and non-jackknifed measures of agreement are almost identical (as occurred in nine of the 14 cases considered). The jackknifed results for the non-grouped LAD and LSD prediction models are much worse for all 28 cases than the corresponding grouped LAD and LSD prediction model results. These findings suggest that LAD prediction models appear to be exceptionally stable for routine applications in the atmospheric sciences or any other branch of science for that matter. The fact that such problems occur with LSD (least squares) regression should not be surprising. Sheynin (1973) has shown that the choice of LSD over LAD regression by C.F. Gauss was a default choice due to the fact that he was unable to solve the linear programs he developed for LAD regression (good computational equipment did not exist at the time).

Acknowledgment

This research was supported by a climate grant from the National Science Foundation.

References

Gary, W. M., C. W. Landsea, P. W. Mielke and K. J. Berry, 1992a: Predicting Atlantic seasonal hurricane activity 6-11 months in advance. *Wea. and Forecast.*, 7, (in press).

Gary, W. M., C. W. Landsea, P.W. Mielke and K. J. Berry, 1992b: Predicting Atlantic basin seasonal tropical cyclone activity by 1 August. *Wea. and Forecast.*, (submitted).

Mielke, P. W., 1984: Meteorological applications of permutation techniques based on distance functions. In *Handbook of Statistics, Vol. 4: Nonparametric Methods*, eds. P. R. Krishnaiah, and P.K. Sen. Amsterdam: North-Holland Publishing Co., pp. 813-830.

Mielke, P. W., 1991: The application of multivariate permutation methods based on distance functions in the earth sciences. *Earth-Sci. Rev.*, 31, 55-71.

Sheynin, O. B., 1973: R. J. Boscovich's work on probability. *Arch. Hist. Exact Sci.*, 9, 306-324.

TABLE 1. MEASURE OF AGREEMENT FOR LAD AND LSD REGRESSIONS ASSOCIATED WITH NON-JACKKNIFED AND BOTH GROUPED (G) AND NON-GROUPED (NG) JACKKNIFED CASES.

	NS	NSD	H	HD	IH	IHD	HDP
LAD REGRESSION MODEL							
DECEMBER FORECAST RESULTS:							
NON-JACKKNIFED	.440	.514	.447	.478(.493)	.507(.498)	.451(.451)	.438(.457)
JACKKNIFED-G	.440	.514	.447	.450(.491)	.471(.498)	.446(.451)	.405(.447)
JACKKNIFED-NG	.411	.386	.279	.305	.376	.374	.235
AUGUST FORECAST RESULTS:							
NON-JACKKNIFED	.451	.612	.468	.516	.622	.612	.577
JACKKNIFED-G	.451	.611	.468	.505	.622	.611	.577
JACKKNIFED-NG	.203	.470	.219	.171	.530	.491	.314
LSD REGRESSION MODEL							
DECEMBER FORECAST RESULTS:							
NON-JACKKNIFED	.359	.407	.388	.400	.491	.450	.430
JACKKNIFED-G	.308	.359	.339	.355	.448	.403	.385
JACKKNIFED-NG	.244	.306	.280	.300	.408	.367	.338
AUGUST FORECAST RESULTS:							
NON-JACKKNIFED	.432	.534	.429	.490	.608	.576	.530
JACKKNIFED-G	.371	.484	.368	.433	.566	.525	.474
JACKKNIFED-NG	.266	.395	.248	.331	.494	.446	.377

TABLE 2. SQUARED PEARSON CORRELATION COEFFICIENTS FOR LAD AND LSD REGRESSIONS ASSOCIATED WITH NON-JACKKNIFED AND BOTH GROUPED (G) AND NON-GROUPED (NG) JACKKNIFED CASES.

	NS	NSD	H	HD	IH	IHD	HDP
LAD REGRESSION MODEL							
DECEMBER FORECAST RESULTS:							
NON-JACKKNIFED	.395	.488	.466	.518(.514)	.575(.581)	.542(.517)	.547(.544)
JACKKNIFED-G	.395	.488	.466	.497(.511)	.530(.581)	.528(.517)	.493(.527)
JACKKNIFED-NG	.368	.358	.280	.320	.415	.431	.276
AUGUST FORECAST RESULTS:							
NON-JACKKNIFED	.434	.592	.436	.573	.680	.615	.591
JACKKNIFED-G	.434	.591	.436	.559	.680	.615	.591
JACKKNIFED-NG	.180	.473	.168	.182	.571	.512	.317
LSD REGRESSION MODEL							
DECEMBER FORECAST RESULTS:							
NON-JACKKNIFED	.416	.523	.482	.538	.585	.543	.570
JACKKNIFED-G	.313	.434	.391	.464	.502	.454	.495
JACKKNIFED-NG	.202	.347	.299	.378	.433	.399	.423
AUGUST FORECAST RESULTS:							
NON-JACKKNIFED	.522	.653	.512	.598	.712	.666	.643
JACKKNIFED-G	.397	.569	.402	.506	.649	.574	.556
JACKKNIFED-NG	.247	.436	.163	.326	.533	.446	.398

J6.1

QUALITY/VALUE RELATIONSHIPS FOR FORECASTS OF AN AUTOCORRELATED CLIMATE VARIABLE

Richard W. Katz

Environmental and Societal Impacts Group
National Center for Atmospheric Research
Boulder, CO 80307

1. INTRODUCTION AND SUMMARY

Most studies of the economic value of weather and climate forecasts have tacitly assumed that the variable being forecast does not possess any temporal dependence. Nevertheless, it is well known that many climate variables are significantly autocorrelated over a wide range of time scales. If the decision-making situation being modeled were static, then the issue of autocorrelation could be treated in a relatively straightforward manner. Most decision-making problems, however, are dynamic in nature, with the action taken by a decision maker on the current occasion having an effect on any actions to be taken on future occasions. So the presence of autocorrelation creates a decision-making problem that is inherently more complex.

A previous paper (Katz, 1992) focused on the economic value of information about the temporal dependence of a climate variable. In the present paper, this analytical framework is extended to the case of imperfect forecasts for the next occasion being available to the decision maker. The dynamic cost-loss ratio decision-making model (Katz and Murphy, 1990) is again applied to the situation in which the temporal dependence of the climate variable is modeled by a Markov chain. For the infinite horizon, discounted version of this problem, it is still possible to obtain analytical expressions for the optimal policy and minimal expected expense associated with forecasts of an autocorrelated climate variable. Other references on the economic value of forecasts of autocorrelated climate variables include Epstein and Murphy (1988) and Wilks (1991).

In section 2, the notation and basic concepts are introduced for the dynamic cost-loss ratio decision-making problem. Particular care needs to be taken in making assumptions about the stochastic properties of the sequence of forecasts when the variable being forecast is itself serially correlated. The approach by which analytical results can be derived and numerical computations can be performed concerning the structure of the optimal policy and the corresponding minimal expected expenses for this decision-making model are presented in section 3. Illustrating some important features of these analytical results, section 4 contains some numerical examples. In particular, the quality/value curves still have a convex shape even with autocorrelation being present, like the case of the climate variable being independent (Katz and Murphy, 1990). How these curves do change in shape and magnitude under autocorrelation is also considered.

2. NOTATION AND CONCEPTS

2.1. Decision-making model

On each occasion, the decision maker must choose between two possible actions: either (i) *protect*; or (ii) *do not protect*. The climate variable Θ has two possible states: either (i) *adverse weather* ($\Theta = 1$); or (ii) *not adverse weather* ($\Theta = 0$). If protective action is taken, then the decision maker incurs a *cost* C. If protective action is not taken and adverse weather does occur, then a *loss* L is incurred, $0 < C < L$. The model is made dynamic by assuming that the loss L can be incurred at most once. The goal of the decision maker is to minimize the expected expense, totaled and discounted (with *discount factor* α, $0 < \alpha < 1$) over an infinite horizon.

2.2. Climatological information

The sequence of climate states $\{\Theta_t: t = 1, 2, ...\}$ is assumed to constitute a two-state, first-order Markov chain. The parameters of this model include the unconditional probability of adverse weather

$$p_\Theta = \Pr\{\Theta_t = 1\}, \quad 0 < p_\Theta < 1, \tag{1}$$

and the transition probabilities

$$P_{ij} = \Pr\{\Theta_t = j \,|\, \Theta_{t-1} = i\}, \quad i, j = 0, 1. \tag{2}$$

It is convenient to characterize the Markov chain in terms of the two parameters p_Θ and d, where

$$d = \mathrm{Corr}(\Theta_{t-1}, \Theta_t) = P_{11} - P_{01} \tag{3}$$

and "Corr" denotes the correlation coefficient. The transition probabilities can be expressed as:

$$P_{01} = p_\Theta(1 - d), \quad P_{10} = (1 - p_\Theta)(1 - d), \tag{4}$$

noting that $P_{00} = 1 - P_{01}$ and $P_{11} = 1 - P_{10}$. It is assumed that the autocorrelation parameter d is positive (i.e., $0 < d < 1$), as is typical for climate variables. The limiting case of $d = 0$ corresponds to an independent climate variable, whereas the other extreme of $d = 1$ corresponds to a perfectly persistent climate.

2.3. Forecast information

A forecast variable Z is introduced, with $Z_t = 0$ or 1 representing a "forecast" of the climate variable Θ on the tth occasion. It is assumed that $\{Z_t: t = 1, 2, \ldots\}$ has the same probabilistic structure as the climate variable; that is, a two-state Markov chain with the same parameters p_Θ and d. Because this requirement must hold in the two limiting situations of $d = 0$ and $d = 1$, it is a reasonable condition to impose in general.

It is necessary to make additional assumptions about the structure of the bivariate process of climate and forecast states $\{(\Theta_t, Z_t): t = 1, 2, \ldots\}$. It is assumed that the following properties hold: (i) the conditional distribution of Θ_t given Z_t, Z_{t-1}, ..., Θ_{t-1}, Θ_{t-2}, ... depends only on Z_t; and (ii) the conditional distribution of Z_{t+1} given Θ_t, Θ_{t-1}, ..., Z_t, Z_{t-1}, ... depends only on Θ_t. Condition (i) implies that the present forecast subsumes any predictive information contained in past climate and forecast states, whereas condition (ii) implies that the next forecast need only be based on the present climate state. In essence, these conditions amount to allowing the forecast variable to "lead" the climate variable by one occasion. Here the term "leading" is formally defined in the predictive sense of Granger and Newbold (1986).

The link from Z_t to Θ_t is determined by the conditional probabilities

$$p_{\Theta|Z}(i) = \Pr\{\Theta_t = 1 \,|\, Z_t = i\}, \quad i = 0, 1, \tag{5}$$

and the link from Θ_t to Z_{t+1} by

$$p_{Z|\Theta}(i) = \Pr\{Z_{t+1} = 1 \,|\, \Theta_t = i\}, \quad i = 0, 1. \tag{6}$$

Because of the conditions imposed, only one of these additional parameters in (5)-(6), say $p_{\Theta|Z}(1)$, need actually be specified to completely determine the joint stochastic behavior of Θ and Z. The other conditional probabilities in (5)-(6) are related by

$$p_{\Theta|Z}(0) = [p_\Theta/(1 - p_\Theta)][1 - p_{\Theta|Z}(1)], \tag{7}$$

$$p_{Z|\Theta}(1) = [P_{11} - p_{\Theta|Z}(0)]/[p_{\Theta|Z}(1) - p_{\Theta|Z}(0)], \tag{8}$$

$$p_{Z|\Theta}(0) = [p_\Theta/(1 - p_\Theta)][1 - p_{Z|\Theta}(1)]. \tag{9}$$

This conditional probability $p_{\Theta|Z}(1)$ can be viewed as specifying the accuracy of the forecasting system, ranging from a lower bound of $p_{\Theta|Z}(1) = P_{11}$ for purely climatological information to an upper bound of $p_{\Theta|Z}(1) = 1$ for perfect forecasts. It is convenient to transform this conditional probability into the equivalent *quality* measure

$$q = [p_{\Theta|Z}(1) - p_\Theta]/(1 - p_\Theta), \quad d < q < 1. \tag{10}$$

Here $q = \mathrm{Corr}(\Theta_t, Z_t)$ is a natural measure of forecast skill. The lower bound on q is d, rather than zero, because of the predictive capability of the Markov chain model itself, representing the persistence of the climate variable (see Katz, 1992). Finally, it is noted that both $p_{\Theta|Z}(1)$ and q are acceptable measures of quality, in the sense of satisfying the so-called "sufficiency relation" (see Ehrendorfer and Murphy, 1988).

2.4. Expected expense and economic value

The decision maker selects the action (protect or do not protect) on each occasion that minimizes the expected expense, totaled and discounted over the infinite future horizon. Here the expected expense is calculated with respect to whatever conditional probability distribution is available for the climate variable Θ. The minimal *forecast expected expense* is denoted by E_Z. If the forecasting system were not

available, the decision maker would still make use of the autocorrelation of the climate variable, obtaining a minimal *climatological expected expense* denoted by E_C [equivalent to E_Z in the limiting case of $p_{\Theta|Z}(1) = P_{11}$ (i.e., $q = d$)]. This expected expense for a persistent climatology was studied in Katz (1992).

The *economic value* V_Z of the forecasting system is defined as the difference in expected expenses

$$V_Z = E_C - E_Z. \tag{11}$$

In other words, it represents the reduction in expected expense incurred by the decision maker attributed to the forecasting system being available. A *quality/value curve* is obtained by plotting the economic value of the forecasts V_Z as a function of the quality q, $d < q < 1$.

3. APPROACH

Because of the assumed Markovian property of the forecast variable, it simplifies matters to condition on the initial forecast state. The expected expense E_Z can be expressed as

$$E_Z = (1 - p_\Theta)E_Z(0) + p_\Theta E_Z(1), \tag{12}$$

where $E_Z(i)$ denotes the minimal expected expense given an initial forecast of $Z_1 = i$, $i = 0, 1$. Using the idea of "backwards induction" (Ross, 1983), the conditional expected expenses satisfy a stochastic dynamic programming recursion

$$E_Z(i) = \min\{C + \alpha[P_{i0}E_Z(0) + P_{i1}E_Z(1)], p_{\Theta|Z}(i)L + \alpha[1 - p_{\Theta|Z}(i)]([1 - p_{Z|\Theta}(0)]E_Z(0) + p_{Z|\Theta}(0)E_Z(1))\}, \quad i = 0, 1. \tag{13}$$

Here the first (second) term on the right-hand side of (13) represents the expected expense if protective action is (is not) taken on the first occasion.

From this fundamental recursion, analytical expressions for the structure of the *optimal policy* (i.e., a rule specifying under which conditions a given action minimizes the expected expense) and the corresponding minimal expected expenses can be determined. In particular, (13) specifies a system of two equations in two unknowns, $E_Z(0)$ and $E_Z(1)$. Once these solutions are found, the desired expected expense E_Z can be obtained by substitution into (12). Both for simplicity and because of space limitations, these analytical results are not presented. Instead, the next section provides some numerical examples, results that can be obtained either from the analytical expressions or directly from the recursion (13).

It is worth noting that the fundamental recursion (13) encompasses several forms of decision-making model and of climatological and forecast information that have been previously treated in isolation. For instance, if the climate variable were not autocorrelated (i.e., $d = 0$), the recursion reduces to that studied in Katz and Murphy (1990). If the additional assumption is made that the discount factor $\alpha = 0$ (i.e., future expenses are completely ignored), then the static cost-loss ratio or "umbrella" problem is obtained (Katz and Murphy, 1987).

On the other hand, it would be straightforward to slightly generalize (13) to study other problems that have previously only been studied under the assumption of no autocorrelation. In particular, it would be feasible to treat the finite horion, undiscounted version of the dynamic cost-loss ratio decision-making model, as was done in Murphy *et al.* (1985) when $d = 0$ (see also Krzysztofowicz and Long, 1990). Finally, the recursion (13) could be extended to treat more complex forms of forecast information (e.g., having a forecasting system that provides forecasts two occasions ahead as well).

4. NUMERICAL EXAMPLES

In both of the following examples, the numerical values of the two economic parameters, the cost-loss ratio C/L and the discount factor α, and of the climate parameter, the unconditional probability of adverse weather p_Θ, are specified. Several values of the persistence parameter d are considered (i.e., $d = 0, 0.25, 0.5$). For each given value of d, a quality/value curve (as defined in section 2.4) is obtained, varying the conditional probability $p_{\Theta|Z}(1)$ between P_{11} and one (equivalently, $d < q < 1$). For the case of no autocorrelation (i.e., $d = 0$), Katz and Murphy (1990) have already established that the quality/value curve has a convex shape. The main purpose of these examples is to examine how the shape and magnitude of these curves change as a function of the autocorrelation parameter d.

Without loss in generality, it is assumed that the loss $L = 1$ (i.e., $C/L = C$). Hence, the minimal expected expenses and the economic value of forecasts can be viewed as scaled to fall within the unit interval [0,1].

4.1. Example 1 ($C/L = 0.15$, $\alpha = 0.9$, $p_\Theta = 0.2$)

This example concerns a situation in which the economic value of forecasts is relatively sensitive to the degree of autocorrelation of the climate variable. The numerical value of the discount factor α is a reasonable choice for an annual or longer time scale [i.e., a discount rate $r = (1 - \alpha)/\alpha \approx 0.11$].

Fig. 1 shows the quality/value curves for this example for the three numerical values of the persistence parameter d. It is evident that these curves remain convex for $d > 0$. For forecasts of equal quality q, the economic value is substantially reduced as d increases, with the reduction being roughly linear in d. Moreover, the curves rise at a less rapid rate the greater the degree of persistence.

4.2. Example 2 ($C/L = 0.01$, $\alpha = 0.98$, $p_\Theta = 0.025$)

This example concerns a situation in which the economic value of forecasts is less sensitive to the degree of autocorrelation of the climate variable. The numerical value of the discount factor α is a reasonable choice for a seasonal or longer time scale (i.e., a discount rate of $r \approx 0.02$).

Fig. 2 shows the quality/values curves for this example for the same three numerical values of the persistence parameter d. Again, the curves are convex for $d > 0$. But the curve does not change much unless the degree of persistence of relatively high. For forecasts of equal quality, the reduction in economic value is a highly nonlinear function of d. On the other hand, the curves rise at nearly the same rate, no matter what the value of d.

ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation (Division of Atmospheric Sciences) under Grant ATM-8714108. A portion of this research was conducted while the author was a sabbatical visitor at the University of California, Davis. The National Center for Atmospheric Research is sponsored by the National Science foundation.

REFERENCES

Ehrendorfer, M., and A.H. Murphy, 1988: Comparative evaluation of weather forecasting systems: Sufficiency, quality, and accuracy. *Monthly Weather Review*, **116**, 1757-1770.

Epstein, E.S., and A.H. Murphy, 1988: Use and value of multiple-period forecasts in a dynamic model of the cost-loss ratio situation. *Monthly Weather Review*, **116**, 746-761.

Granger, C.W.J., and P. Newbold, 1986: *Forecasting Economic Time Series* (second edition). Academic Press, Orlando, FL, 338 pp.

Katz, R.W., 1992: Dynamic cost-loss ratio decision-making model with an autocorrelated climate variable. *Journal of Climate*, **5** (in press).

Katz, R.W., and A.H. Murphy, 1987: Quality/value relationship for imperfect information in the umbrella problem. *American Statistician*, **41**, 187-189.

Katz, R.W., and A.H. Murphy, 1990: Quality/value relationships for imperfect weather forecasts in a prototype multistage decision-making model. *Journal of Forecasting*, **9**, 75-86.

Krzysztofowicz, R., and D. Long, 1990: To protect or not to protect: Bayes decisions with forecasts. *European Journal of Operational Research*, **44**, 319-330.

Murphy, A.H., R.W. Katz, R.L. Winkler and W.-R. Hsu, 1985: Repetitive decision making and the value of forecasts in the cost-loss ratio situation: A dynamic model. *Monthly Weather Review*, **113**, 801-813.

Ross, S.M., 1983: *Introduction to Stochastic Dynamic Programming*. Academic Press, New York, 164 pp.

Wilks, D.S., 1991: Representing serial correlation of meterological events and forecasts in dynamic decision-analytic models. *Monthly Weather Review*, **119**, 1640-1662.

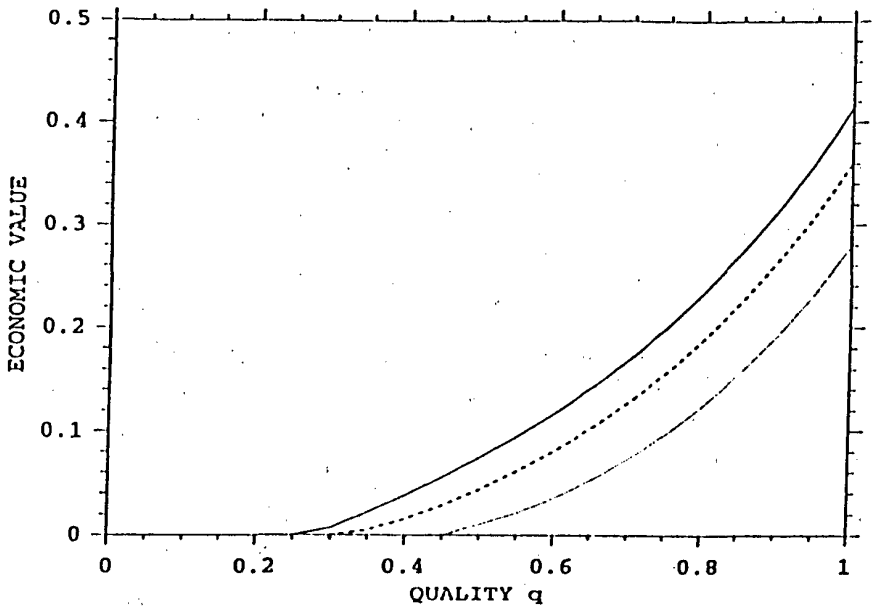


FIG. 1. Quality/value curves for Example 1 (C/L = 0.15, α = 0.9, p_Θ = 0.2) showing economic value V_Z as a function of forecast quality q. Solid line is for persistence parameter d = 0, dashed line for d = 0.25, and dotted line for d = 0.5.

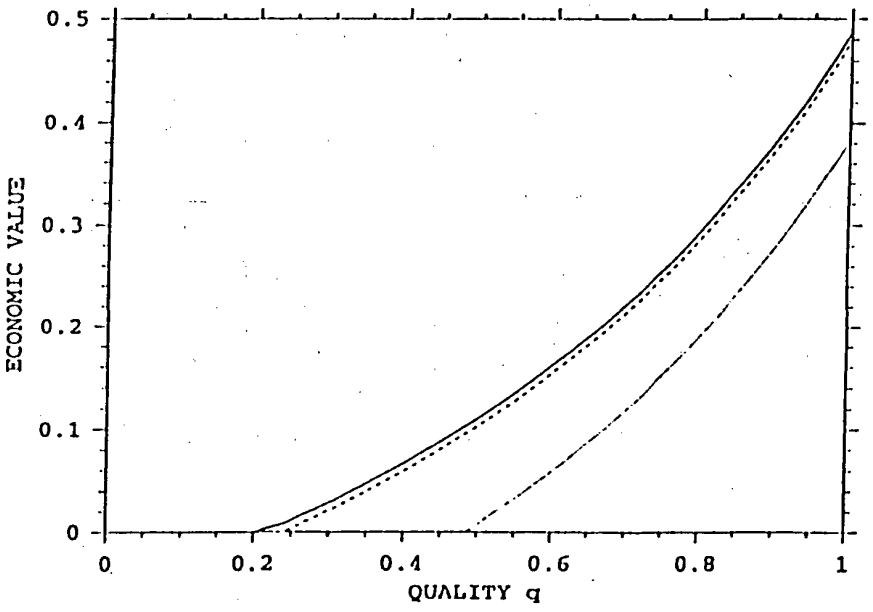


FIG. 2. Same as Fig. 1 for Example 2 (C/L = 0.01, α = 0.98, p_Θ = 0.025).

QUALITY/VALUE RELATIONSHIPS FOR WEATHER AND CLIMATE FORECASTING SYSTEMS

J6.3

Martin Ehrendorfer[1] and Allan H. Murphy[2]

[1]Institute for Meteorology and Geophysics, University of Vienna, A-1190 Vienna, Austria
[2]Department of Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331, U.S.A.

1. INTRODUCTION

The value of weather and climate forecasts is difficult to assess, whether these assessments are based on *prescriptive* studies of the ways in which the forecasts should be used or *descriptive* studies of the ways in which they are actually used. The former involve structuring, modeling, and analyzing idealized decision-making problems, whereas the latter involve investigating, analyzing, and describing the behavior of users in the real-world with regard to their information-processing and decision-making practices (see Katz and Murphy 1992; Stewart *et al.* 1984; Winkler and Murphy 1985). Regardless of which approach is taken, meteorologists and/or users generally neither possess the appropriate training nor the necessary experience to conduct and evaluate such studies.

For this (and other) reasons, meteorologists and/or users frequently use measures of the quality of forecasts as surrogates for measures of value. For example, it is frequently assumed, either explictly or implicitly, that increases in forecast accuracy necessarily lead to increases in forecast value. However, it is relatively easy to show that this assumption is, in general, *not* correct (see Section 2.3). What then is the nature of the relationship between forecast quality and forecast value? What general properties do quality/value relationships possess and what (if any) other properties do such relationships exhibit in specific decision-making problems? Under what conditions can increases in quality be unambiguously assumed to lead to increases in value?

These (and other related) questions have been addressed in a variety of recent studies related to the value of weather and climate forecasts in general and quality/value relationships for such forecasts in particular. The purposes of this paper are to describe the basic concepts and methods underlying the assessments of quality/value relationships, to summarize the results of recent studies investigating these relationships, and to discuss the methodological and practical implications of these results. In order to restrict the scope of this review to manageable portions, attention is focused on quality/value relationships in contexts in which a prescriptive approach is taken to forecast-value assessment.

In Section 2 some basic definitions and concepts regarding forecast quality and forecast value are introduced. The implications of these basic considerations for quality/value relationships are briefly discussed. Section 3 considers quality/value relationships in general, with particular reference to the sufficiency relation. First, this relation is defined and its properties are discussed, and then an example illustrating its application is presented. This section also summarizes recent studies involving the application of the sufficiency relation - and other closely related evaluation methodologies - to weather and climate forecasts. Section 4 describes recent work related to the study of quality/value relationships in the context of specific situations, including applications involving both prototypical and real-world decision-making problems. Section 5 contains a discussion of (a) the implications of currently available results regarding quality/value relationships for the practice of forecast verification; (b) major deficiencies in the current state of knowledge regarding quality/value relationships; and (c) the topics and issues that should be addressed in future studies of these relationships.

2. FORECAST QUALITY AND FORECAST VALUE: BASIC DEFINITIONS AND CONCEPTS

2.1 Forecast quality

Forecast quality is characterized by the joint distribution of forecasts and observations, $p(f, x)$, where f and x denote the forecasts and the observations, respectively (see, Murphy and Winkler 1987). This distribution contains all of the nontime-dependent information necessary to describe forecast quality *completely*. If, in addition, the time-dependent behavior of forecast quality is of interest, then the joint distribution $p(f, x, t)$ - in which t denotes time - must be considered.

Forecast quality is inherently multidimensional in nature, in the sense that, in general, more than one number is required to reconstruct the basic joint distribution described above. For example, it takes 3 $(= 2 \times 2 - 1)$ numbers to reconstruct $p(f, x)$ in dichotomous situations involving nonprobabilistic forecasts and it takes 21 $(= 11 \times 2 - 1)$ numbers to reconstruct $p(f, x)$ in dichotomous situations involving probabilistic forecasts with 11 distinct probability values. Dimensionality (in this sense) is a fundamental characteristic of forecast verification problems (Murphy 1991).

The multidimensional nature of forecast quality can also be understood from a different (but related) perspective. Forecast quality can be shown to possess several components (or attributes), including bias, accuracy, skill, reliability, resolution, and discrimination (Murphy and Winkler 1987). Moreover, these attributes are related to the various conditional and/or marginal distributions that can be derived by factorizing the basic joint distribution. For example, reliability and resolution are related to the conditional distributions of the observations given the forecasts, $p(x|f)$, whereas discrimination is related to the conditional distributions of the forecasts given the observations, $p(f|x)$.

These considerations suggest that traditional methods of forecast verification and evaluation, in which forecast quality is characterized by one or two measures of overall performance (e.g., mean square error, correlation coefficient, skill score), are fundamentally flawed. Specifically, traditional methods do not respect the true dimensionality of verification problems and, as a result, they are quite likely to overlook - or at least to measure inadequately and/or incompletely - various components of forecast quality. The nature and extent of these deficiencies depend on several factors, including the severity of the reduction in dimensionality, the properties of the verification methods themselves, and the statistical characteristics of the data (i.e., the pairs of forecasts and observations).

Comparative evaluation - which is concerned with the comparison of the quality of two or more sets of forecasts - suffers from many of these same deficiencies. As presently practiced, it is also based largely on summary measures of overall performance such as mean square errors, (anomaly) correlation coefficients, and skill scores. As in the case of a single set of forecasts, in order to assess the relative quality of different sets of forecasts and to draw proper

inferences regarding their value, it is necessary to consider quality in its full dimensionality. In particular, the conditional and marginal distributions associated with the respective joint distributions must be evaluated - and compared - in order to make definitive statements regarding relative forecast quality and forecast value (see also Murphy 1991).

2.2 Forecast Value

Forecasts possess no intrinsic value; instead, they acquire value through their use by decision makers. In the context of the prescriptive approach to decision making and the value of information inherent in forecasts, the basic determinants of forecast value are (1) the alternatives (or admissible actions, possible decisions) available to the decision maker; (2) the payoff structure associated with the decision-making problem; (3) the quality of information on which decisions are based in the absence of forecasts; and (4) the quality of the forecasts themselves (see Hilton 1981). If the alternatives available to the decision maker change (e.g., an action is added or deleted), the decision-making problem itself is changed, and such a change generally leads to changes in payoff structure and forecast value.

The payoff structure specifies a loss or gain for each possible combination of alternative and event (in this paper, the events are defined in terms of weather and/or climate variables). These losses or gains can be expressed in many different ways; for example, in terms of monetary losses/gains, lives lost/saved, etc. Here we assume that all losses or gains have been expressed in monetary terms and that these monetary payoffs reflect the true worth of these outcomes to the decision maker. In addition, it is assumed that the decision maker chooses the alternative that minimizes (maximizes) his/her expected loss (gain). In effect, these assumptions imply that the decision maker's utility function is linear in monetary payoff (see also Raiffa 1968; Winkler and Murphy 1985).

Determinants (3) and (4), taken together, indicate that forecast value depends on both the quality of the forecasts and the quality of the information on which decisions are based in the absence of the forecasts. In particular, if the quality of the forecasts is such that the decision maker makes the same decisions with and without the forecasts, then the forecasts are of no value. It should also be noted that the assumption that the decision maker possesses a *linear* utility function simplifies the assessment of forecast value. Under this assumption, the value of the forecasts is simply the difference between the decision maker's expected payoffs when his/her decisions are made with and without the forecasts.

It is also important to recognize that forecast value in general depends on forecast quality in its full dimensionality; that is, in order to assess forecast value the joint distribution $p(f, x)$ (or its basic factorizations) must be known. In the prescriptive approach, expressions for forecast value generally involve the conditional distributions of the observations given the forecasts, $p(x|f)$, and the marginal distribution of the forecasts, $p(f)$ (e.g. Winkler *et al.* 1983; Murphy 1985). Thus, this approach measures forecast value in an *ex ante* sense rather than in an *ex post* sense.

2.3 Implications for Quality/Value Relationships

Since forecast quality is inherently multidimensional in nature and forecast value depends *(inter alia)* on forecast quality in its full dimensionality, the relationship between forecast quality and forecast value is necessarily complex. In addition, the prescriptive approach to decision making itself, in which the decision maker (under the linear utility assumption) chooses the alternative that minimizes (maximizes) his/her expected loss (gain), dictates that this relationship is inherently nonlinear.

It is in general true that forecast value increases as forecast quality (in its full dimensionality) increases. However, the multidimensional nature of forecast quality implies that increases or decreases in *components* of quality (e.g., accuracy) do not necessarily imply concurrent increases or decreaes in value. For example, Murphy and Ehrendorfer (1987) have shown that increases in forecast accuracy can actually result in decreases in forecast value. Such quality-value reversals can occur when one-dimensional scores that measure particular components of quality are used as surrogates for multidimensional measures of quality itself. In such situations, changes in the basic components of the underlying joint distribution of forecasts and observations can yield a better one-dimensional score at the same time that they prescribe that a user take courses of action that lead to less desirable outcomes. In fact, these reversals can occur whenever the dimensionality of forecast quality is reduced in an arbitrary manner (i.e., in a manner that fails to take into account the user's decision-making problem). Only in those situations in which forecast quality is (or is assumed to be) one-dimensional does a one-to-one monotonic relationship exist between forecast accuracy (which is then equivalent to forecast quality) and forecast value.

In order to investigate in greater detail the relationship between quality and value, it is therefore quite natural to identify the conditions - to be satisfied by the joint distributions of forecasting systems - such that increases in quality lead unambiguously to increases in value. These conditions are embodied in the sufficiency relation (Blackwell 1953; DeGroot and Fienberg 1986), which explicitly accounts for the multidimensional nature of forecast quality. The applicability of this relation in the context of comparative evaluation of weather and climate forecasting systems has been explored in several recent studies (see Section 3). Due to the potential importance of the sufficiency relation in assessing the quality and value of forecasting systems, as well as the nature of the quality/value relationship, it is considered in detail in the following section.

3. QUALITY/VALUE RELATIONSHIPS IN GENERAL: THE SUFFICIENCY RELATION

3.1 Description of the Sufficiency Relation

In the context of comparative evaluation, the sufficiency relation accounts for the multidimensional nature of forecast quality by considering the two joint distributions of forecasts and observations. This multidimensional comparison establishes whether or not system A is *sufficient* for system B. In brief, system A is sufficient for system B if B's joint distribution (or a factorization thereof) can be obtained through a stochastic transformation of A's joint distribution (for a formal definition, see e.g. Ehrendorfer and Murphy 1988). The stochastic transformation represents an auxiliary randomization that introduces uncertainty into B's forecasts that was not present in A's forecasts.

The conditions for the existence of a stochastic transformation are rather stringent. In any case, given that system A is sufficient for system B, two important consequences follow: (i) system A's forecasts are of higher quality than system B's forecasts; and (ii) system A's forecasts possess greater value than system B's forecasts independent of any reference to a specific user (or payoff structure). Thus, if sufficiency can be established (which is not always possible; see below), this relation orders the forecasting systems in both quality and value. It is evident, then, that the sufficiency relation is a potentially useful tool in investigating the general nature of the relationship between forecast quality and forecast value.

It is important to understand that the sufficiency relation establishes only a quasi-order on forecasting systems. That is, it is not always possible to show that system A is sufficient for system B (or vice versa); in such cases, no stochastic transformation exists and the two systems are said to be *insufficient* for each other. The frequency with which - and the conditions under which - insufficiency is encountered in the real world are issues of considerable practical importance and warrant careful investigation.

3.2 Application of the Sufficiency Relation: An Example

In this subsection the usefulness of the sufficiency relation for the investigation of the relationship between the quality and value of forecasting systems is illustrated through a prototypical example (for details, see Ehrendorfer and Murphy 1992). The sufficiency relation is applied in the context of the comparative evaluation of prototypical climate forecasting systems that are restricted to three distinct probability forecasts, each specifying probabilities for three possible states of nature.

The forecasting systems under consideration are prototypical in the sense that the verification problem is essentially two-dimensional. The first parameter, δ, indicates the "distance" of the two non-climatological forecasts from the climatological forecast, whereas the second parameter, π, is the frequency of use of the non-climatological forecasts. This setup (which is in addition highly symmetric) makes it possible to describe the forecasting systems under consideration in a two-dimensional diagram.

In this situation, application of the sufficiency relation yields a separation of the two-dimensional parameter space into three different kinds of regions: (i) region S containing the systems B that are sufficient for the given reference system A; (ii) region S' containing the systems B for which the reference system A is sufficient; (iii) region I containing the systems B that are insufficient for A. An example of the corresponding *sufficiency diagram* is shown in Fig. 1. The reference system A is indicated in this figure by a large dot and has the parameter values $\delta = -0.10$ and $\pi = 0.15$. From this figure it is evident that systems B are sufficient for A if they use a more extreme non-climatological forecast than A (expressed through larger values of δ). However, this result holds only if π^B is at least as large as π^A; that is, the more extreme non-climatological forecast must be issued by system B with a frequency that is at least as large as the frequency of use of the non-climatological forecast of system A. Otherwise, δ^B must become substantially larger as π^B decreases. Still, to a limited degree smaller π^B can be offset by larger (in absolute value) δ^B. However, the converse does not hold; if δ^B is smaller than δ^A it can be seen from Fig. 1 that this deficiency with regard to sufficiency cannot be offset even by large values of π^B.

In order to illustrate the consequences of the sufficiency relation for the relationship between the quality and value of the forecasting systems under consideration here, isopleths of the expected ranked probability score (ERPS; solid lines) as well as of the value of the forecasts (VF; dashed lines) are included in the diagram. Note that both types of isopleths are drawn at unequal intervals. The ERPS (for a definition of the RPS, see Epstein 1969) is chosen to serve as representative of a one-dimensional surrogate of forecast quality. The VF is assessed within the framework of the prototypical decision-making problem considered by Murphy (1985); in this problem, three possible alternatives (protective actions) are available to the decision maker (see also Section 4.1). In the context of Fig. 1, the basic expense matrix - describing the expenses (payoff structure) associated with specific combinations of actions and events (see also Section 2.2) - is modeled through the cost-loss ratio C/L. Here this ratio takes on the value 0.3; this number indicates that the cost of full protection against adverse conditions is 30% of the total loss.

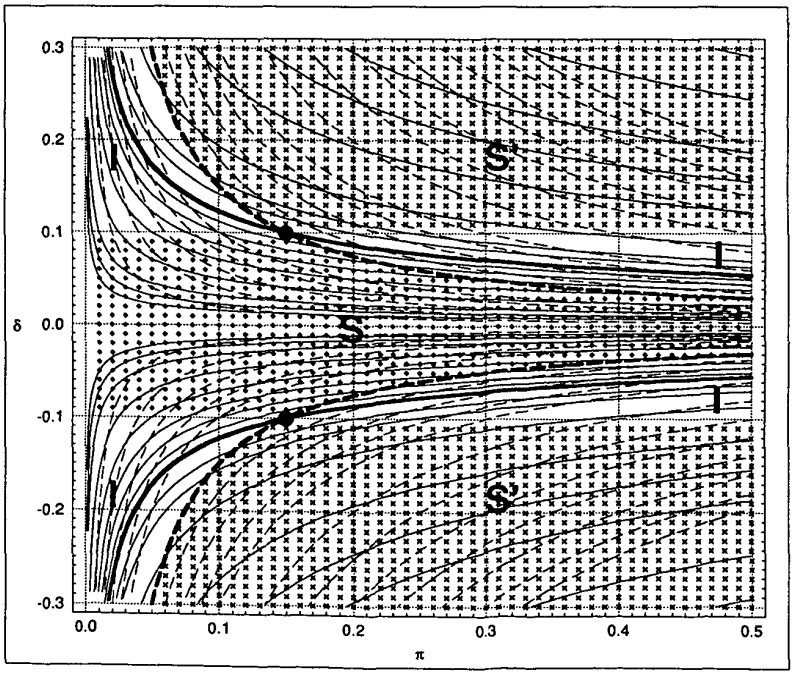


FIG. 1. Example of a sufficiency diagram (see Section 3.2). For a given reference system A (denoted by a large dot) the diagram identifies (i) the alternative systems B for which the reference system is sufficient (diamonds, region denoted by S), (ii) the alternative systems B that are sufficient for the reference system (crosses, region denoted by S'), and (iii) insufficient systems (blank region denoted by I). The identification of the regions shown is achieved through explicit application of the sufficiency relation. Lines included in the diagram are isopleths of ERPS (solid) and VF (dashed; assessed as described in Section 3.2 for $C/L = 0.3$). These isopleths are drawn at unequal intervals for the numerical values of ERPS (VF) of 0.25, 0.30, 0.35, 0.372, 0.39, 0.405, 0.41025, 0.412, 0.414, 0.415, 0.416, 0.417, 0.418, 0.419, 0.4196, 0.4198, 0.4199 (0.001, 0.002, 0.003, 0.004, 0.005, 0.0075, 0.009, 0.011, 0.013, 0.015, 0.020, 0.025, 0.030, 0.035, 0.040, 0.045, 0.049), with increasing (decreasing) numerical values from the lower- and upper-right corners toward the middle of the diagram (note the symmetry around the horizontal line $\delta = 0$). The bold isopleths denote the ERPS and VF of the reference system; namely, ERPS = 0.414 and VF = 0.0075. For additional details, see Section 3.2.

From the general discussion of the implications of the sufficiency relation for forecast quality (see Section 3.1), the results in Fig. 1 are clear; all systems lying in S' possess better expected scores than that of the reference system. This result shows that ERPS is to some degree consistent with the results derived from the sufficiency relation, in the sense that a better score can never be achieved by a system for which A is sufficient (i.e., the ERPS-isopleths are convex curves in this diagram). However, it also becomes clear that some aspects (specifically, one component) of quality are lost if only the score itself is considered; it is impossible to infer from ERPS alone whether system B is superior to A in all dimensions of quality (i.e., is sufficient for A) or merely insufficient for A. This deficiency of the ERPS - common to all one-dimensional scores in such situations - is due to the fact that ERPS-isopleths exist that cross both regions S' and I.

Considering the VF-isopleths, it is evident that higher quality (as indicated by larger values of δ and π) implies higher value as the VF-isopleths are again convex curves. Thus, sufficiency implies higher quality as well as higher value (see also Section 3.1). The shape of the VF-isopleths plotted for VF=0.0075 (i.e., VF of the reference system A) is rather remarkable. Obviously, this isopleth represents the boundary between region I and regions S' (for $\pi \leq \pi^A = 0.15$) and S (for $\pi \geq \pi^A = 0.15$). The special nature of this isopleth can be further interpreted as follows: consider an alternative system B with the same VF as the reference system. Then, a marginal improvement in quality will lead to a system that is sufficient for the reference system (given that δ^B is larger than δ^A; otherwise, the improved system is merely insufficient for A). Note that this property is rather special and does not apply to systems that possess the same ERPS as the reference system; that is, marginal improvements in quality when both systems exhibit the same ERPS will *not* lead to a sufficient system, since the ERPS-isopleth of the reference system does not represent a boundary in the sufficiency diagram.

Next, examining the behavior of the ERPS- and the VF-isopleths together indicates the reason that a multi-valued relationship exists between ERPS and VF in this simple example. Specifically, the two types of isopleths intersect; for example, while traveling along a solid curve (i.e., holding quality constant) a number of dashed curves are intersected (i.e., value is changing). Further, the possibility of a quality-value reversal is obvious; a better score may be associated with a decrease in value. For example, consider traveling from a point (e.g., in S') where both isopleths intersect along a dashed curve towards the next intersection (i.e., ERPS improves while VF is held constant) and then moving along a solid curve to the next intersection (i.e., ERPS is constant while VF decreases). This represents a quality-value reversal, because an improved score is associated with a decrease in value. However, it can also be seen from Fig. 1 that if the multidimensional nature of quality is considered in the sense that an increase in quality is denoted by improving on (or holding constant) both π and δ (and not improving one and decreasing the other as in the case of a quality-value reversal), then such an increase is *always* accompanied by greater value due to the convexity of the VF-isopleths.

In the diagram in Fig. 1 the implications of the sufficiency relation for value have been considered for a specific class of users (i.e., those users for whom the cost-loss ratio decision-making model is appropriate with $C/L = 0.3$). As a further illustration of the implications of the sufficiency relation for the value of the forecasting systems under consideration, VF is presented for a larger class of users (i.e., C/L is allowed to vary), in Fig. 2. In this case, the VF of four selected forecasting systems is shown as a function of C/L. The parameters determining the quality of these systems (their respective positions in the sufficiency diagram) are given in the legend. Specifically, the reference system A is considered as well as a system B that is insufficient for A, but has the same value of ERPS; further, a second insufficient system C with a better ERPS is considered; finally, two systems, D and E, are considered that are both sufficient for the reference system A, where system D possesses the property of using more extreme forecasts less often (i.e., $|\delta^D| > |\delta^A|$, but $\pi^D < \pi^A$) (see also Fig. 1).

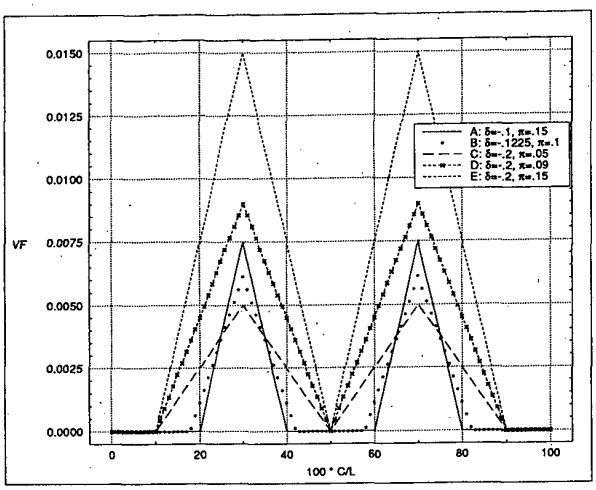


FIG. 2. The value VF of selected forecasting systems of the type considered in Section 3.2 as a function of the payoff structure parameterized through the cost-loss ratio C/L. Performance characteristics π and δ are given in the legend. Systems A, B, and C are chosen such that they are insufficient for each other, whereas systems D and E are sufficient for the reference system A (see Fig. 1). Note that the relative magnitude of VF for the insufficient systems depends on the specific payoff structure identified through the numerical value of C/L, whereas systems D and E have greater VF than A over the whole range of C/L as they have been found to be sufficient for A.

First, considering $C/L = 0.3$ in Fig. 2 (recall that $C/L = 0.3$ has been chosen to compute VF in Fig. 1), it can be seen that for this specific payoff structure the reference system A has larger VF (namely, 0.0075) than the two systems B and C that are insufficient for A. However, the VF of A is substantially smaller than the VF of systems D and E that are sufficient for A (namely, 0.009 and 0.015, respectively). Second, from the consideration of different users (modeled through different cost-loss ratios), it can be seen from Fig. 2 that the systems that are sufficient for A always have larger values of VF than A regardless of the specific payoff structure under consideration. The latter is an illustration of one of the powerful consequences of being able to show that one forecasting system is sufficient for another forecasting system. Third, it is also evident from Fig. 2 that, in the case of insufficient systems, the answer to the question whether one system is more valuable than another system depends on the specific value of the cost-loss ratio. For example, for a user with $C/L = 0.2$ the VF of the system B is larger than the VF of system A (0.001127 versus 0.0), with the opposite relationship holding for, for example, $C/L = 0.3$ (0.0075 versus 0.006127). This result illustrates clearly the interpretation that can be given to insufficient systems: if two systems are insufficient for each other, the payoff structure of the user determines which one is more valuable; that is, the alternative system is found to have greater value for some users and less value for other users (relative to the reference system).

3.3 Review of Recent Applications of the Sufficiency Relation

The sufficiency relation - originally developed in the context of the comparison of statistical experiments by Blackwell (1951, 1953) - was introduced into the forecasting literature by DeGroot and Fienberg (1982, 1983, 1986). It was introduced into the meteorological literature by Ehrendorfer and Murphy (1988) who explored its application to primitive probabilistic forecasting systems. They showed that it is possible to identify conditions that basic performance characteristics (i.e., elements of the joint distribution) must satisfy in order to ensure that one forecasting system is sufficient for another forecasting system. In addition, they verified that, if A's forecasts are sufficient for B's forecasts in such contexts, the value of the former is greater than that of the latter irrespective of the user's payoff structure.

The sufficiency relation has been formulated in a different manner by Krzysztofowicz and Long (1991a), based on a theorem by Blackwell and Girshick (1954). They identified a forecast sufficiency characteristic (FSC) which employs simple conditions to determine whether one forecasting system is sufficient for another forecasting system. This approach offers advantages over the direct application of the definition of sufficiency, in that it avoids explicit computation of stochastic transformations. The fact that the search for stochastic transformations becomes increasingly laborious as the dimensionality of the joint distribution increases makes an approach based on the FSC particularly attractive.

Comparative evaluation of objective and subjective precipitation probability forecasts has been undertaken using FSCs by Murphy and Ye (1990a). For these highly competitive forecasts, they found that the respective FSCs seldom satisfied the conditions for sufficiency. Situations that might be characterized as "almost sufficient" were found on numerous occasions, but considerable care must be excercised in drawing conclusions regarding relative quality and value in such situations. Krzysztofowicz and Long (1991b) also applied FSCs to comparative evaluation of precipitation probability forecasts, but they modeled the predictive probabilities using beta distributions (instead of using empirical relative frequencies). The use of beta distributions tends to smooth the empirical data, and they found that the conditions for sufficiency were satisfied on most occasions with smoothed data.

Recently, Krzysztofowicz (1992) used the sufficiency relation as a basis for formulating a measure of forecast skill that, under certain conditions, allows direct inferences regarding sufficiency when comparing two forecasting systems. Under these conditions, the Bayesian correlation score represents a one-dimensional measure of forecasting performance that incorporates all relevant aspects of forecast quality. This study demonstrates the importance of the sufficiency relation as a theoretical framework within which meaningful and useful verification scores can be developed.

4. QUALITY/VALUE RELATIONSHIPS IN SPECIFIC SITUATIONS

The relationship between the quality and value of weather and climate forecasts has been investigated in a variety of specific situations. Two general types of weather/climate-information-sensitive situations can be identified: (a) prototypical situations and (b) real-world situations. The former represent idealized decision-making problems, whereas the latter represent decision-making problems that actually arise in the real world. Most studies of quality/value relationships in these situations have been based on a prescriptive, decision-analytic approach to decision making and to assessing the value of information (see Winkler and Murphy 1985; Katz and Murphy 1992). In implementing this approach, simplifying assumptions are frequently made about various features of the underlying situations. These assumptions relate to such features as the structure of the problem (e.g., static or dynamic; i.e., past decisions do not or do affect future decisions), the alternatives available to the decision maker, the weather/climate events, the payoff structure (e.g., costs, losses), the format and number of distinct forecasts, and the decision criterion. This section briefly reviews a representative set of studies of quality/value relationships in prototypical (Section 4.1) and real-world (Section 4.2) contexts.

4.1 Prototypical Situations

The most widely studied prototypical decision-making problem is the well-known cost-loss ratio situation (e.g., Thompson 1962; Murphy 1977). Relationships between forecast quality and forecast value in static versions of this problem have been investigated by Chen *et al.* (1987), Katz and Murphy (1987), Murphy (1985), Murphy and Ehrendorfer (1987), and Murphy and Ye (1990b). Moreover, quality/value relationships in dynamic versions of the cost-loss ratio situation have been explored by Katz and Murphy (1990) and Murphy *et al.* (1985). Recently, studies of the latter type have been extended to include autocorrelated forecasts and/or observations (Epstein and Murphy 1988; Wilks 1991).

Two important characteristics of quality/value relationships identified in such studies are: (a) their inherent nonlinearity and (b) their multi-valued nature when quality is not measured in its full dimensionality. With regard to the nonlinearity of the quality/value relationship, the latter is often characterized by the existence of a quality threshold below which forecasts are of no value (see also Section 3.2). Above the quality threshold, value generally increases as a nonlinear function of increasing quality. It is also interesting to note that the quality/value "curve" is convex, in the sense that the *sensitivity* of forecast value to changes in forecast quality increases as forecast quality increases.

Quality/value relationships become multi-valued when the multidimensional nature of quality is not respected (e.g., when quality is measured in terms of a one-dimensional measure of accuracy). In such circumstances, as illustrated in Section 3.2, reversals can occur in the usual quality/value relationship, in the sense that (for example) value can decrease as accuracy increases for at least some users (see Chen *et al.* 1987; Murphy and Ehrendorfer 1987). In a different vein, consideration of the autocorrelation in forecasts and/or observations reduces the value of forecasts (because the zero point on the value scale is raised), but it does not alter the above-mentioned general characteristics of quality/value relationships.

Other idealized situations include the generic choice-of-crop problem investigated by Winkler *et al.* (1983). Quality/value relationships in these situations have not been subjected to the same intensive study as that directed towards these relationships in the cost-loss ratio situation.

4.2 Real-World Situations

Studies of the value of weather/climate forecasts - and quality/value relationships - in real-world situations usually involve the formulation of models that prescribe the user's decision-making and information-processing procedures. For example, the so-called "impact functions" that translate the effects of the weather/climate variables into economic or other payoffs to the user must be specified. In addition, statistical models are frequently used to characterize the quality of the forecasts under consideration. Considerable effort is frequently required (in the areas of data acquisition and analysis, as

well as in model development and refinement) to ensure that these models represent reasonably realistic descriptions of the relevant procedures and relationships.

Most real-world studies of quality/value relationships have been conducted within the framework of agricultural decision-making problems. For example, studies of the static decision-making problems include the haying/pasturing situation (Wilks and Murphy 1985) and a specific choice-of-crop problem (Wilks and Murphy 1986). Dynamic decision-analytic studies include the corn-production problem (Sonka *et al.* 1986, 1987), the fruit-frost problem (Katz *et al.* 1982), the fallowing-planting problem (Brown *et al.* 1986, Katz *et al.* 1987), and the harvest-scheduling problem (Wilks *et al.* 1992).

With regard to quality/value relationships, the results of these real-world studies support the conclusion that these relationships are inherently nonlinear. Moreover, most of these studies reveal quality thresholds below which the forecasts are of no value. With regard to forecast value itself, studies involving short-range weather forecasts indicate that such forecasts can be of considerable value, achieving 50% of the value of perfect forecasts (e.g., Katz *et al.* 1982). On the other hand, current long-range forecasts appear to be of relatively little value overall, although even modest increases in quality could lead to significant increases in their value (e.g., Brown *et al.* 1986; Wilks and Murphy 1985). Moreover, these results also demonstrate that, above the quality threshold, the relationships between forecast quality and forecast value are generally nonlinear.

5. DISCUSSION AND CONCLUSION

The quality of forecasting systems is an inherently multidimensional concept requiring for its complete description consideration of the joint distribution of forecasts and observations. Thus, by its very nature, quality cannot be fully described by one-dimensional scores. Such one-dimensional descriptions of forecasting performance are only adequate for measuring certain components of forecast quality (e.g., accuracy) and are, in general, incapable of characterizing forecast quality in its full dimensionality.

The value of forecasting systems is acquired through the use of their forecasts by decision makers involved in weather-sensitive decision-making problems. Here, the value of a forecast can be thought of as the expenses saved by using the forecast. The value of a forecasting system is strongly related to the quality of the forecasting system.

The present paper has concentrated on various aspects of the relationship between the quality and the value of forecasting systems. This relationship is known to be of a complex, and inherently non-linear, nature (see Section 2.3). Specifically, no simple way exists to infer value from quality (and vice versa). Moreover, the relationship between scores - as surrogates of forecast quality - and forecast value is, in general, a multi-valued functional relationship in the sense that forecast value ranges between certain limits for a particular score. Specific situations do exist when quality and value stand in a one-to-one monotonic relationship (this relationship is, in general, nonlinear). These situations are those in which quality is (or is assumed to be) one-dimensional.

In the context of the study of quality/value relationships it is important to distinguish between the nature of the results regarding forecast value - and quality/value relationships - that are obtained from prototypical and real-world studies (see Section 4). In addition to the fact that prototypical studies do not relate to any specific weather/climate-information-sensitive situation, such studies provide value estimates (either directly or indirectly through quality/value relationships) in relative terms. For example, these estimates may indicate the value of the forecasts relative to the value of perfect forecasts, but not the magnitude of their economic value. Thus, studies of forecast value - and quality/value relationships - in real-world contexts are especially important because these studies provide the only source of information regarding the (absolute) magnitudes of forecast-value estimates.

The consequences of the complicated natue of the relationship between quality and value are manifold. First, the current practice of automatically identifying a forecast with a better score (e.g., a more accurate forecast) with a forecast that is also more useful (i.e., of higher value) is, evidently, fraught with some danger. Such a conclusion is not always valid and even the reverse relationship may hold (see Sections 2.3 and 3.2).

Second, from a more theoretical point of view it is necessary to explore and clarify the nature and characteristics of the relationship between forecast quality and forecast value. Such studies have been conducted to date primarily in the context of specific situations (see Section 4) without the identification of *general* properties of the relationship between quality and value. For example, it would be of considerable interest to be able to describe in general terms the shape of the multi-valued quality/value envelope when quality is measured by a given score (e.g., the Brier score) or to identify the conditions that must be satisfied by the elements of the joint distribution such that several constraints (e.g., impossibility of quality/value reversals, limitation to one-to-one functional relationship) concerning the quality/value relationship are satisfied. In this context it seems to be of particular importance to identify a framework within which quality/value relationships can be studied in a systematic and coherent manner. To some degree, the sufficiency relation represents part of such a framework (see Section 3), since it naturally accounts for the multidimensional character of forecast quality. However, this framework must be complemented by an appropriate decision-analytic framework.

The study of the relationship between quality and value of forecasts not only provides quantitative assessments of the benefits of meteorological forecasts, but also yields useful insights into several related issues. For example, it may be possible to identify classes of users for whom less complex quality/value relationships exist or for whom monotonic quality/value relationships exist under less restrictive conditions. It may also lead to the identification of essential components of forecast quality in the sense that such components summarize effectively most of the information contained in the joint distribution. Further, the identification of a coherent framework to study quality/value relationships should be extremely valuable for the methodological aspects of forecast verification. As such, the study of quality/value relationships must be considered to represent an integral part of the complex problem of forecast evaluation.

Acknowledgments. This work was supported in part by the National Science Foundation under Grant SES-9106440. The availability of computing time on the IBM 3090-400E at the University of Vienna is gratefully acknowledged.

REFERENCES

[1] Blackwell, D., 1951: Comparison of experiments. *Proc. Second Berkeley Symp. on Mathematical Statistics and Probability*, J. Neyman, Ed., University of California Press, 93-102.

[2] Blackwell, D., 1953: Equivalent comparisons of experiments. *Annals of Mathematical Statistics*, **24**, 265-272.

[3] Blackwell, D., and M.A. Girshick, 1954: *Theory of Games and Statistical Decisions*. Wiley, 355 pp.

[4] Brown, B.G., R.W. Katz, and A.H. Murphy, 1986: On the

economic value of seasonal precipitation forecasts: The fallowing/planting problem. *Bull. Amer. Meteor. Soc.*, **67**, 833-841.

[5] Chen, Y.-S., M. Ehrendorfer, and A.H. Murphy, 1987: On the relationship between the quality and value of forecasts in the generalized cost-loss ratio situation. *Mon. Wea. Rev.*, **115**, 1534-1541.

[6] DeGroot, M.H., and S.E. Fienberg, 1982: Assessing probability assessors: calibration and refinement. *Statistical Decision Theory and Related Topics III, Volume 1*, S.S. Gupta and J.O. Berger, Eds., Academic Press, 291-314.

[7] DeGroot, M.H., and S.E. Fienberg, 1983: The comparison and evaluation of forecasters. *The Statistician*, **32**, 12-22.

[8] DeGroot, M.H., and S.E. Fienberg, 1986: Comparing probability forecasters: basic binary concepts and multivariate extensions. *Bayesian Inference and Decision Techniques*, P. Goel and A. Zellner, Eds., Elsevier, 247-264.

[9] Ehrendorfer, M., and A.H. Murphy, 1988: Comparative evaluation of weather forecasting systems: sufficiency, quality, and accuracy. *Mon. Wea. Rev.*, **116**, 1757-1770.

[10] Ehrendorfer, M., and A.H. Murphy, 1992: Evaluation of prototypical climate forecasts: the sufficiency relation. *J. Climate*, **5**, in press.

[11] Epstein, E.S., 1969: A scoring system for probabilities of ranked categories. *Journal of Applied Meteorology*, **8**, 985-987.

[12] Epstein, E.S., and A.H. Murphy, 1988: Use and value of multiple-period forecasts in a dynamic model of the cost-loss ratio situation. *Mon. Wea. Rev.*, **116**, 746-761.

[13] Hilton, R.W., 1981: The determinants of information value: Synthesizing some general results. *Management Science*, **27**, 57-64.

[14] Katz, R.W., B.G. Brown, and A.H. Murphy, 1987: Decision-analytic assessment of the economic value of weather forecasts: The fallowing/planting problem. *Journal of Forecasting*, **6**, 77-89.

[15] Katz, R.W., and A.H. Murphy, 1987: Quality/value relationship for imperfect information in the umbrella problem. *The American Statistician*, **41**, 187-189.

[16] Katz, R.W., and A.H. Murphy, 1990: Quality/value relationships for imperfect weather forecasts in a prototype multistage decision-making model. *Journal of Forecasting*, **9**, 75-86.

[17] Katz, R.W., and A.H. Murphy, Eds., 1992: *Economic Value of Weather and Climate Forecasts*. Cambridge University Press, in preparation.

[18] Katz, R.W., A.H. Murphy, and R.L. Winkler, 1982: Assessing the value of frost forecasts to orchadists: A dynamic decision-making approach. *J. Appl. Meteorol.*, **21**, 518-531.

[19] Krzysztofowicz, R., 1992: Bayesian correlation score: a utilitarian measure of forecast skill. *Mon. Wea. Rev.*, **120**, 208-219.

[20] Krzysztofowicz, R., and D. Long, 1991a: Forecast sufficiency characteristic: construction and application. *International Journal of Forecasting*, **7**, 39-45.

[21] Krzysztofowicz, R., and D. Long, 1991b: Beta likelihood models of probabilistic forecasts. *International Journal of Forecasting*, **7**, 47-55.

[22] Murphy, A.H., 1977: The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation. *Mon. Wea. Rev.*, **105**, 803-816.

[23] Murphy, A.H., 1985: Decision making and the value of forecasts in a generalized model of the cost-loss ratio situation. *Mon. Wea. Rev.*, **113**, 362-369.

[24] Murphy, A.H., 1991: Forecast verification: its complexity and dimensionality. *Mon. Wea. Rev.*, **119**, 1590-1601.

[25] Murphy, A.H., and M. Ehrendorfer, 1987: On the relationship between the accuracy and value of forecasts in the cost-loss ratio situation. *Weather and Forecasting*, **2**, 243-251.

[26] Murphy, A.H., R.W. Katz, R.L. Winkler, and W.-R. Hsu, 1985: Repetitive decision making and the value of forecasts in the cost-loss ratio ratio situation: A dynamic model. *Mon. Wea. Rev.*, **113**, 801-813.

[27] Murphy, A.H., and R.L. Winkler, 1987: A general framework for forecast verification. *Monthly Weather Review*, **115**, 1330-1338.

[28] Murphy, A.H., and Q. Ye, 1990a: Comparison of objective and subjective precipitation probability forecasts: the sufficiency relation. *Monthly Weather Review*, **118**, 1783-1792.

[29] Murphy, A.H., and Q. Ye, 1990b: Optimal decision making and the value of information in a time-dependent version of the cost-loss ratio situation. *Mon. Wea. Rev.*, **118**, 939-949.

[30] Raiffa, H., 1968: *Decision analysis: Introductory lectures on choices under uncertainty*. Addison-Wesley, 309 pp.

[31] Sonka, S.T., P.J. Lamb, S.E. Hollinger, and J.W. Mjelde, 1986: Economic use of weather and climate information: concepts and an agricultural example. *Journal of Climatology*, **6**, 447-457.

[32] Sonka, S.T., J.W. Mjelde, P.J. Lamb, S.E. Hollinger, and B.L. Dixon, 1987: Valuing climate forecast information. *Journal of Climate and Applied Meteorology*, **26**, 1080-1091.

[33] Stewart, T.R., R.W. Katz, and A.H. Murphy, 1984: Value of weather information: a descriptive study of the fruit-frost problem. *Bull. Amer. Meteor. Soc.*, **65**, 126-137.

[34] Thompson, J.C., 1962: Economic gains from scientific advances and operational improvements in meteorological prediction. *J. Appl. Meteorol.*, **1**, 13-17.

[35] Wilks, D.S., 1991: Representing serial correlation of meteorological events and forecasts in dynamic decision-analytic models. *Mon. Wea. Rev.*, **119**, 1640-1662.

[36] Wilks, D.S., and A.H. Murphy, 1985: The value of seasonal precipitation forecasts in a haying/pasturing problem in western Oregon. *Mon. Wea. Rev.*, **113**, 1738-1745.

[37] Wilks, D.S., and A.H. Murphy, 1986: A decision-analytic study of the joint value of seasonal precipitation and temperature forecasts in a choice-of- crop problem. *Atmosphere-Ocean*, **24**, 353-368.

[38] Wilks, D.S., R.E. Pitt, and G.W. Fick, 1992: Modeling optimal alfalfa harvest scheduling using short-range weather forecasts. Submitted to *Agricultural Systems*.

[39] Winkler, R.L., and A.H. Murphy, 1985: Decision analysis. *Probability, Statistics, and Decision Making in the Atmospheric Sciences*, A.H. Murphy and R.W. Katz, Eds., Westview Press, 493-524.

[40] Winkler, R.L., A.H. Murphy, and R.W. Katz, 1983: The value of climate information: A decision-analytic approach. *Journal of Climatology*, **3**, 187-197.

AN EXAMINATION OF METHODOLOGICAL ISSUES IN CLUSTERING NORTH AMERICAN PRECIPITATION

Xiaofeng Gong
Weather Bureau of Liaoning Province
Shenyang, People's Republic of China

and

Michael B. Richman
Cooperative Institute for Mesoscale Meteorological Studies
The University of Oklahoma
Norman, OK 73019

1. INTRODUCTION

In geophysical studies, investigators often need to classify the variables under investigation into relatively homogeneous groups, for indentification of the underlying physics and to assess the potential for predictability. Cluster analysis (CA) is one of a number of multivariate techniques that can be used to accomplish such classification. A key to a successful clustering algorithm is the ability to define clusters (groups of similar variables or observations) which exhibit two properties: external isolation and internal cohesion (Cormack, 1971). External isolation requires that variables (or observations) within one cluster be separated from those in another cluster by a fairly empty space. Internal cohesion requires the variables (or observations) within one cluster be similar to one another. CA imposes a characteristic structure on the data for exploratory purposes. It does so in several analysis steps -- by calculating dissimilarity amongst the variables and by using various formulae for admission of a variable into a cluster.

Interestingly, there have been only limited comparisons of the various methodological considerations for CA research. The majority of them have appeared in the psychological and biological literature (i.e., Gower, 1967; Johnson, 1967; Kuiper and Fisher, 1975; Blashfield, 1976; Milligan, 1980). While these papers do provide some critical guidance for geophysical work, CA results are dependent upon the data structure and none of the cited studies incorporated an exhaustive list of the available algorithms (or even a large subset). Several published meteorological studies have applied CA to precipitation data. Kikuchihara (1984) used complete linkage CA to establish homogeneous precipitation regions in Japan. A number of researchers used Ward's method successfully (Goosens, 1985; Winkler, 1985; Reich, 1986; Easterling, 1989) for their precipitation studies. These are all hierarchical methods. Non-hierarchical methods have been applied less often to geophysical studies. Gadgil and Joshi (1983) used K-means CA to establish a precipitation subdivision of India, as did Winkler (1985) for heavy precipitation in the central and eastern U.S. Ronberg and Wang (1987) employed an algorithm known as the nucleated agglomerative method to regionalize Chinese precipitation.

The purpose of this work is to review a large range of available clustering techniques which have already been used on an individual basis in the geophysical science research. The analyses were designed to shed light on the sensitivities of CA methods to various aspects of the analysis to provide a general guide in selecting appropriate algorithms. This paper is drawn from an extensive review and Monte Carlo comparison of CA methodology (Gong and Richman, 1992).

2. CLUSTER ANALYSIS METHODOLOGY

Consider $\underline{X}$ a p x n matrix of raw data, where p references observations and n references entities or variables. The first required step is to transform $\underline{X}$ into a new n x n (or p x p) matrix of either similarities or dissimilarities. This is the fundamental stage before invoking a clustering algorithm. Hierarchical clustering methods (i.e., CA which creates a nested sequence of partitions from the dissimilarity matrix by successive mergers) require specific measurements of dissimilarity in order to merge the most similar variables. Since there is no general agreement on a best similarity/dissimilarity measure, we therefore chose to adopt three dissimilarity coefficients to maximize the utility of our work. These are discussed below.

2.1 *Distance measures*

a. Euclidean Distance (ED)

If one considers a data matrix $\underline{X}$, the Euclidean distance between variables x_i and x_j is given by

$$d_{ij} = \left[(x_i - x_j)^T (x_i - x_j)\right]^{1/2}$$

assuming the p-observations are independent. Since observations are rarely independent in most data sets, a Principal Component (PC) transformation is applied to $\underline{X}$ and the PC scores are used to calculate the distances.

b. Theta Angle between Variables (T)

The angle between variables has also been used as a dissimilarity coefficient. The value ranges from 0° for identical vectors to 90° for vectors with no common elements. The angle between x_i and x_j is given by

$$\Theta = \text{Arc cos}\left[\frac{x_i^T x_j}{|x_i||x_j|}\right]$$

c. Inverse Correlation Coefficient (IC)

Since the correlation coefficient has historically been applied to CA, we computed the inverse correlation to form a dissimilarity coefficient

$$r_{ij}^{-1} = \left|\frac{x_i^T x_j}{(x_i^T x_i)^{1/2}(x_j^T x_j)^{1/2}}\right|^{-1}$$

The values range from 1 for identical entities to $+\infty$ for those with no similarity.

2.2 *Seed Points*

Since we are also examining non-heirarchical routines (CA which does not form a nested sequence of partitions), the *a priori* choice of partition is a prerequisite. These seed points are the estimates of cluster centroids, around which the clusters may be formed. There is no widely-accepted best method and the effects of choosing a different method are not well documented. We chose three different methods in the hope of determining if one was more accurate and documenting possible sensitivities of each cluster algorithm to individual partition methods.

a. Sequential (S)
The first k data entities in matrix X are chosen as the seed points for k clusters.

b. Randomly (R)
The k seed points are chosen via a pseudo-random number generator.

c. Gridded (G)
An equal area grid is overlaid on a map of the spatial domain and the k variables closest to each grid box centroid is used.

2.3 *Clustering Algorithms*

The first group tested were hierarchical methods. These are all agglomorative algorithms which start with each variable in its own cluster and then combine the two least dissimilar variables. The combination process continues in discrete steps until only one cluster remains (with all n variables). The analyst decides at which level the CA should be interpreted. All methods yield "hard clusters" where every variable is assigned to only one of k clusters. Distance formulae for the 5 hierarchical methods employed will now be presented.

a. Single Linkage (SL) - under this method, a cluster is defined as a group of variables such that every member of the cluster is more similar to at least one member of the same cluster than it is to any member of another cluster. If d_{ij} is the distance between clusters i and j, and the clusters are to be merged, the distance between the new cluster k and any other cluster m will be

$$d_{km} = \min(d_{im}, d_{jm})$$

Hence, SL CA joins clusters by the shortest link between them. This can sometimes lead to "chaining", where single variables are more likely to join an existing cluster than to act as the nucleus of a new cluster.

b. Complete Linkage (CL) - under this method, the distance between clusters is determined by the distance between two entities, one from each cluster, that are most distant. The distance between cluster k, merged by i and j, and other cluster m will be

$$d_{km} = \max(d_{im}, d_{jm})$$

In this case d_{im} and d_{jm} are distances between the most distant members of clusters i and m, and clusters j and m, respectively. The bias in CL CA tends to run counter to SL, as the chance of a cluster obtaining a new member becomes smaller as the size of the cluster increases.

c. Average Linkage between Merged Groups (AL1) - this technique treats the distance between two clusters as the average distance between all pairs of variables where one member of a pair belongs to each cluster. This distance between cluster k, merged by clusters i and j, and another cluster m is determined by

$$d_{km} = \frac{N_i}{N_i+N_j} d_{im} + \frac{N_j}{N_i+N_j} d_{jm}$$

where N_i and N_j are the number of entities in those clusters.

d. Average Linkage within the New Group (AL2) - this technique evaluates not only the similarity between all pairs of variables that belong to two different clusters (as AL1 does), but also the similarity of all pairs of variables which belong to the same cluster. The key concept is to use this in order to characterize a cluster by the average of all links within it. The sum of pairwise distances between cluster k, merged by cluster i and j, and another cluster m is:

$$d_{km} = d_{im} + d_{jm}$$

e. Ward's Method (W) - this hierarchical method was designed to generate clusters in such a way that mergers at each stage are chosen so as to minimize the within-group sum of squares (hence optimizing an objective statistic). At each stage it combines two clusters, k and m, that result in the least increase in the function

$$W_{km} = \frac{N_k N_m}{N_k + N_m} (\bar{x}_k - \bar{x}_m)^T (\bar{x}_k - \bar{x}_m)$$

where $\bar{x}_k$ and $\bar{x}_m$ denote the centroids k and m.

We have also applied two types of non-hierarchical hard clustering methods previously used in atmospheric sciences. Non-hierarchical methods must have their cluster centroids chosen, *a priori*, with "seed points". The two methods used alter cluster membership so as to obtain a new partition according to the nearest centroid sorting algorithm.

f. K-Means clustering (KM) - this method begins with the prespecification of k seed points as the centroids for k clusters. The Euclidean distance is then computed between the variables and the centroids and each variable is assigned to the initial partition. The centroids are then recomputed according to the nearest centroid sorting formula. The nearest centroid (Y_{mj}) is given by:

$$Y_{mj} = \frac{1}{N_m} \sum_{i=1}^{N_m} x_{ij} \qquad (j = 1,2,...,p;\ m = 1,2,...,k)$$

where N_m represents the number of variables assigned to cluster m. Variables are reassigned to the new clusters until convergence is achieved.

g. Nucleated Agglomorative clustering (NA) - this method is based on the same sorting algorithm as shown above for KM. The key difference is that the procedure performs as agglomorative clustering at a range from k_{max} (a number of initial clusters) to a predetermined number k_{min}. This makes NA useful for examining a range of values for the number of clusters. Moreover, the seed points only need to be calculated once.

3. DATA

The data set used for this study was the enlarged Richman-Lamb precipitation network for North America (Richman *et al.*, 1991). The data used for this study were the daily precipitation totals for 766 locations in the eastern two-thirds of the U.S. and Canada for the May-August periods of 1949-1987. The grid-like station location of the network is displayed in Fig. 1. It is bounded on the west and east by the Rocky Mountains and Atlantic Ocean, to the south by the Gulf of Mexico, and to the north by the limit of agriculture or the availability of precipitation stations with at least moderate density. The daily values were totaled to 7-day values (i.e., 1-7 May, 8-14 May, ..., 20-27 August) prior to the clustering. Hence the clusters should be indicitive of weekly precipitation over the period of observation (39 years). There were a total of 663 observations.

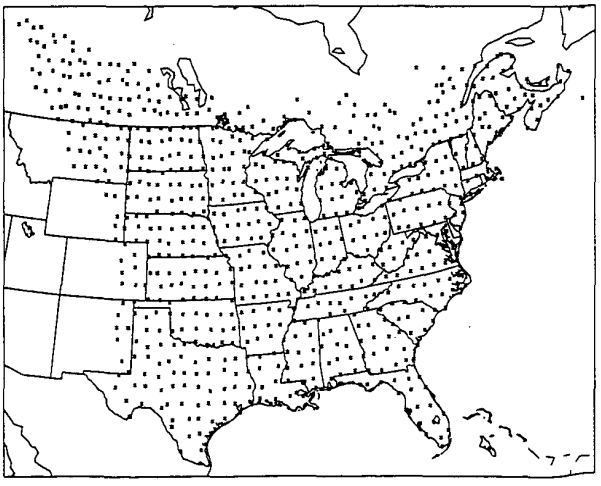


Figure 1. Precipitation station network (766 locations).

4. RESULTS

4.1 *Results of full analysis*

In order to objectively evaluate the aforementioned methods, we needed to keep as many parameters as possible fixed. One decision made in CA is the determination of the number of clusters to be retained. Rather than selecting a single value, we chose to keep a range of 8 to 25 to reflect all of the significant precipitation variability without losing signal. This allowed both direct comparisons between methods for a fixed cluster number and evaluation of the effects of changing the number of clusters on each method. Hence we calculated separate solutions for 8,9,...,25 clusters for a total of 18 analyses. For the hierarchical techniques outlined under sections 2a-2e, each of the 766 variables starts as its own cluster and is then merged until we reach 25 clusters. The analysis is then repeated at 24 clusters, etc. Under the KM method (2f), 8 to 25 seed points are chosen for separate clustering, whereas for the NA method (2g) $k_{max} = 25$ and $k_{min} = 8$ in one step. Since each of the hierarchical techniques is repeated for the three dissimilarity coefficients (and the non-hierarchical ones for three initial partition methods), we end up with 21 clustering methods (7x3), each one of which contain 18 different solutions (8-25 clusters) for a total of 378 analyses. These will be summarized by the use of boxplots and by spatially mapping the clusters on the precipitation network for a representative solution.

The evaluation of the relative accuracy of each clustering method was based on matching the CA spatial pattern against its corresponding interstation correlation field. In order to accomplish this, each cluster's geometric centroid was selected and this served as the base point (1.0) for the correlation analysis. Variables (stations) inside a cluster were assigned values of 1.0, whereas those outside the cluster were assigned values of 0.0, in accordance with the usual practice for hard clustering methods. This means we are correlating a binary value (stations in the cluster versus those outside the cluster) with a continuous variable (precipitation correlation field), and this is known as a biserial correlation. Milligan (1980) successfully used the biserial correlation to match synthetic data. For an 8 cluster solution, 8 correlation matching coefficients are obtained and averaged to represent a single value for the 8 cluster match. This is repeated for 9 clusters, etc., until all 18 average matches are obtained. These are then summarized in the boxplots (Fig. 2).

Figure 2 also contains orthogonally (OR) and obliquely (OB) rotated principal component (PC) solutions as a basepoint comparison for the previous work of Richman and Lamb (1985; 1987). The PCs had their loadings "binarized" or made into hard clusters at the +0.2 loading isopleth. This hardening of the PC loadings reduced their accuracy considerably, from biserial match of close to +0.90 to approximately +0.75. However, we considered this essential to try to make comparisons between the methods. It does point to a penalty incurred when one uses hard clusters to portray precipitation.

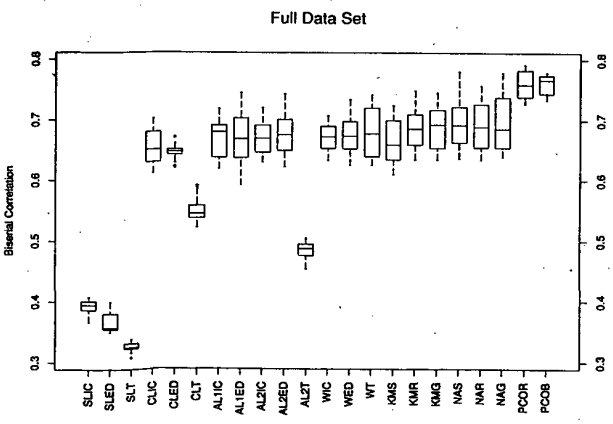


Figure 2. Box and whisker plots for the matches for each of the cluster/PC analysis methods to the corresponding interstation correlations (see text for details). Box indicates median value and interquartile ranges, while whiskers show 1.5 interquartile range and any outliers (asterisks).

The key results of this work (Fig. 2) indicate single linkage (SL) methods all performed poorly with biserial correlations of less than 0.4. In order to determine why the matches were low, we present a representative solution mapped to the domain for 16 clusters (mid-way between 8 and 25). These CA results are shown in Fig. 3 and illustrate the nature of the regionalization for each method. Under all three SL methods, the chaining solutions are evident, with one large cluster and 15 much smaller ones (Figs. 3a-3c). This is obviously a pathological result of the SL algorithm as it bears no resemblance to the known patterns of interstation covariation. The complete linkage (CL) methods are summarized in Fig. 2 and Fig. 3d-f. All three CL methods are more accurate than SL. The inverse correlation (IC) and Euclidean distance (ED) solutions have considerably higher correlations than the theta angle (T), with medians being in the 0.65 range for the former. The spatial maps (Fig. 3d-f) show less chaining than SL, although CLT had 8 very large small clusters. The first average linkage (AL1) method could only be mapped for the IC and ED measures, as T led to severe fragmenting (stations from different areas of the domain would be included in a single cluster). The median values shown in the boxplots are close to +0.7. The spatial patterns now show fairly equally sized clusters (Fig. 3g,h), yet there is not much agreement between the two AL1 regionalization's boundaries. The second average linkage (AL2) technique had similar accuracy levels to AL1 for the IC and ED coefficients (Fig. 2). The AL2T solution was now stable enough to be mapped, yet the biserial correlation was much lower having boxplots with no overlap between AL2T and AL2IC (or AL2ED). Inspection of AL2T (Fig. 3k) shows some chaining in south Texas, whereas AL2IC and AL2ED (Fig. 3i,j) have more equal sized precipitation regions. The final hierarchical CA method tested, Wards (W), showed the least range in WIC and WED (Fig. 2) and slightly more for WT, as well as a fairly high correlation. The spatial regionalizations provided by W (Fig. 3l-n) show similar sized clusters to those of AL1 and AL2, where the western regions are slightly larger than those to the east. However, the regional boundaries of all of the methods vary considerably.

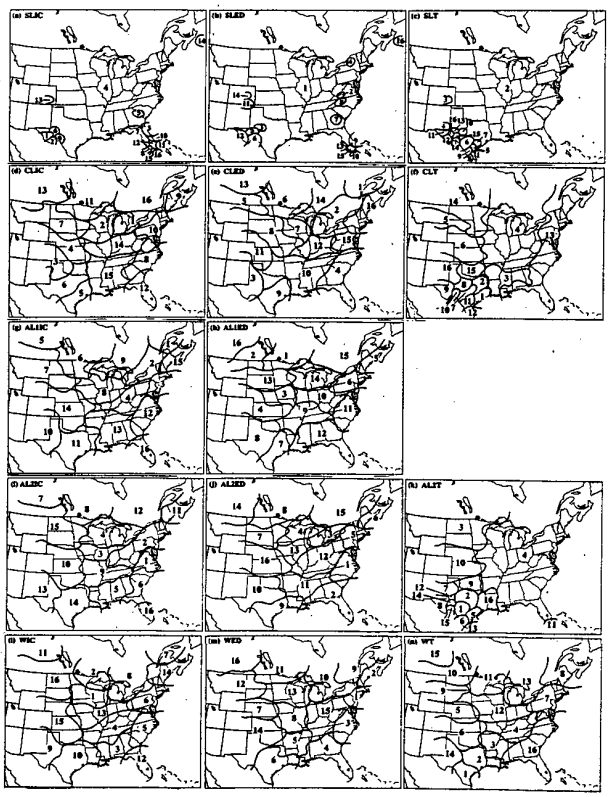


Figure 3. Regionalization of the eastern two-thirds of central North American growing season precipitation by various clustering methods and rotated PCA. Number of clusters/PCs is 16 in all analyses.

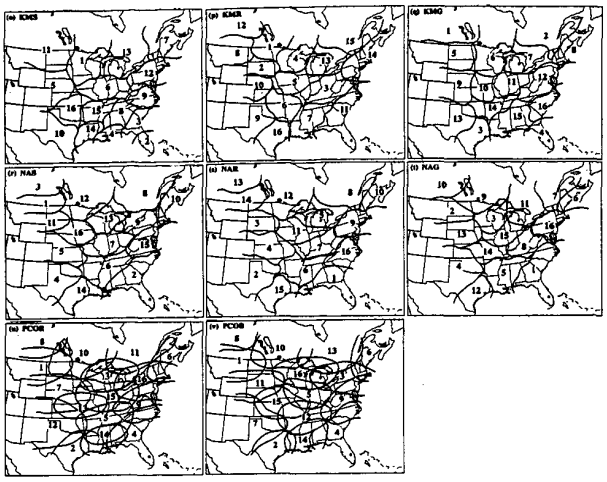


Figure 3. Continued.

Analysis of the non-hierarchical techniques revealed slightly higher median values than the hierarchical ones overall (Fig. 2). Interestingly, a bias in the initial partition method for KMS was uncovered. Since the first partition was based on variables 1-8 to 1-25 (all in the southern edge of the domain), the regionalization (Fig. 3o) has a distinctly smaller set of clusters in the south. This is an important finding as KM CA applications are growing in popularity in the atmospheric sciences. KMR and KMG do not appear to show this bias (Fig. 3p,q), although the random selection of seed points would be suseptible to this initial partition influence and the analyst would not necessarily be able to uncover it. There is still considerable variability in the regional boundaries of the various methods, both within CA method and within any given dissimilarity coefficient. The second non-hierarchical method, nucleated agglomorative (NA), was notable as it showed no sensitivity to partitioning method (Fig. 2). It also seemed to have more stable regions (Fig. 3r,s,t) despite several areas in the midwest and east being variable. However, due to the high biserial correlations, fairly low range, no apparent sensitivity to the seed method, and increased stability, we feel that the NA method is the one of choice for hard clustering of precipitation data.

The rotated PCs had very high biserial correlations, in excess of 0.75 (recall they were in excess of 0.85 prior to binarization) and lower ranges than any of the CA methods. The agreement among the two rotational method's (*Varimax versus Harris-Kaiser B'B*) regions is striking. This may help shed some light on the degree of inter-rotational variability in the context of another regionalization method. However, a direct comparison between PCA and CA is difficult as PCA does not yield hard clusters with no overlap. Contrarily, the orthogonally rotated PCs have a high degree of overlap at the loading level (+0.2) chosen to maximize the biserial correlation.

4.2 *Results of resampling*

In order to apply statistical tests, to compare specific clustering algorithms, we needed to run multiple samples of each method. Monte Carlo bootstrapping was used to sample with replacement for sample sizes 50, 100, 250, and 500. This is in the range of most published studies encountered and permits a controlled examination of the strengths and weaknesses of each algorithm. Two statistics were used to compare the biserial correlation derived matches -- pairwise t-tests (not shown here) and the MSD test for unplanned comparisons (Sokal and Rohlf, 1981). We also examined fragmentation of rainfall regions. Since the underlying correlations were always coherent ellipses, any algorithm which had disjoint, fragmented subregions was considered undesirable. These statistics are summarized in Table 1 where two noteworthy features emerge: (i) fragmentation is most serious at the smaller sample sizes (i.e., 50, 100), where CL, AL1, AL2IC and WIC are most effected and, (ii) non-hierarchical solutions rarely fragment. Further investigation revealed that most fragmentation of CA occurred at the lower end of the number of clusters retained ($\leq$ 12). This is just the opposite of what Richman and Lamb (1987, p. 157) indicate for PCA, where retaining too many eigenvectors leads to splintering or fragmenting.

At the smallest sample size (50) SL method were all very poor, regardless of the input dissimilarity matrix (Fig. 4a). The MSD test (Table 2) showed no significant differences among the techniques. CL was statistically superior to SL. The remaining hierarchical methods were all more accurate than SL techniques, though the MSD test generally placed CL in an equivalent category. The exceptions were AL2T and WIC being significantly less accurate than WED. The non-hierarchical KM methods were not significantly better than WED, largely due to the aforementioned sensitivity to seed point location. The NAS results were significantly more accurate than KMS, but NAR and NAG were numerically indistinguishable from KMR and KMG, respectively. The PC results were not better than the NA method, pointing to stringent sampling requirements needed to get an accurate PC regionalization.

When the sample size was increased to 100 (Fig. 4b) the median biserial correlations increased from about 0.5 to 0.6 for the better methods. The same general patterns tended to arise: CL was clearly superior to SL, W was the most accurate hierarchical method and NA was the most accurate non-hierarchical one. WED and WT were more accurate than WIC according to the MSD test (Table 2). Again, all NA methods were more accurate than KMS. Both the orthogonally and obliquely rotated (PCOR, PCOB) PC methods had the highest medians, at this sample size, yet the MSD test did not separate them from the NA CA methods. At the 250 sample level (Fig. 4c), all interquartile ranges narrowed considerably, suggesting more uniform results, and the median of the better methods rose to ~0.65. This emphasizes the advantages of adequate sampling for all of these methods. Moreover, the degree of fragmentation decreased sharply (Table 1). Noteworthy statistical differences distinguished by the MSD test (Table 2) were AL1 techniques being more accurate than their CL counterparts. Once again the W methods had the highest medians for hierarchical methods, whereas the NA had the highest medians for non-hierarchical techniques tested, yet the MSD test indicated NA methods superior to KMS only. The PC biserial matches were significantly more accurate than the NA ones for this sample size.

These trends continue at 500 bootstrapped observations (Fig. 4d). The median biserial correlation match continued to rise to 0.65-0.70 for the better CA algorithms. Fragmenting further declined (Table 1), being serious for CLIC. The CL and AL techniques were all on the same approximate plane of accuracy according to the MSD-test. However, W algorithms were superior to CL methods. The NA were once again the most accurate non-hierarchical solutions, but statistically superior to only KMS, when using the MSD test (Table 2) to distinguish them. PCOR and PCOB were quite accurate, having statistically higher matches than any of the CA methods.

Table 1. Fragmenting statistics for each cluster method as a function of sample size.

Size	SLIC	SLED	SLT	CLIC	CLED	CLT	AL1IC	AL1ED	AL2IC	AL2ED	AL2T	WIC	WED	WT	KMS	KMR	KMG	NAS	NAR	NAG
FULL				5	7							2								
50				85	37	8	35	11	42			64	4					1	2	2
100				71	26	11	36	7	21			48								
250				70	21	5	12		7			24								
500				43	5	4	4		4			20								

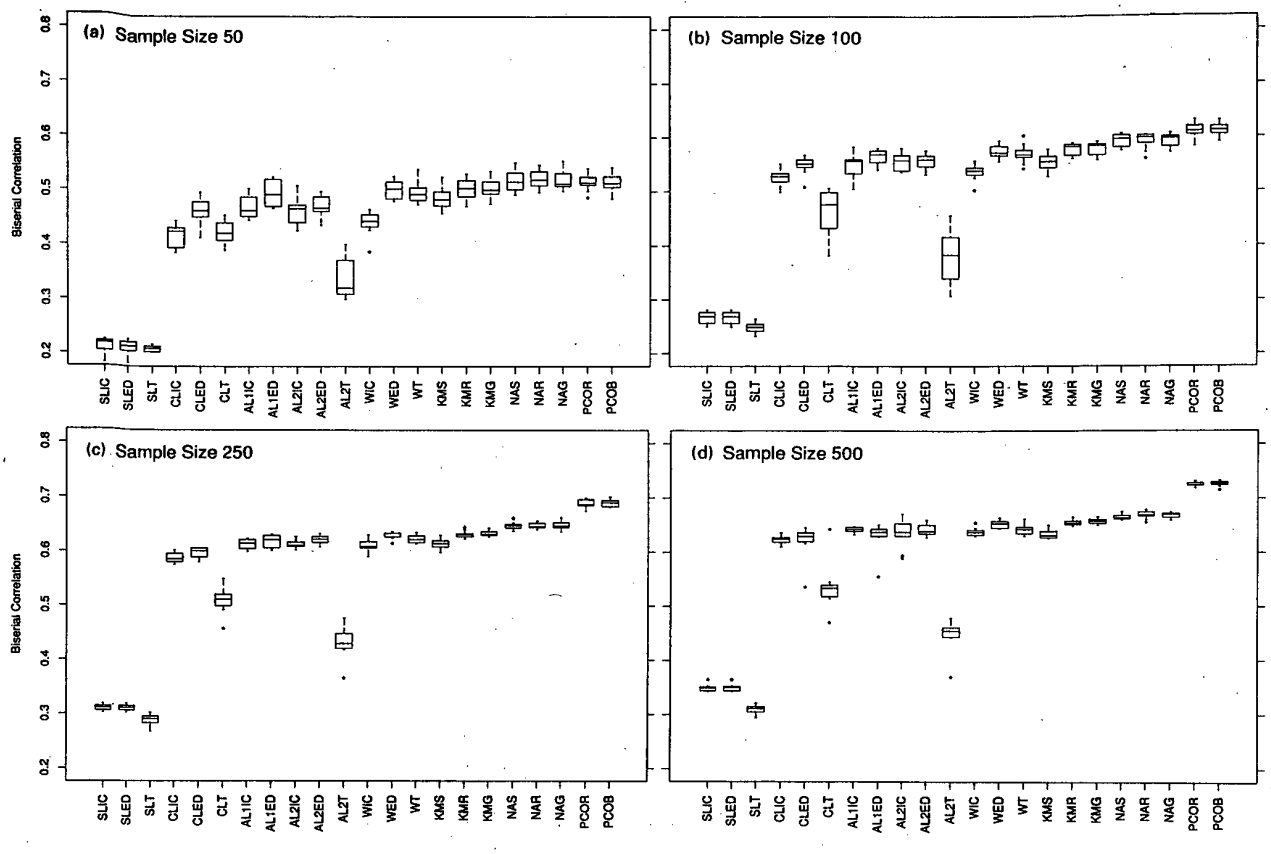


Figure 4. Same as Fig. 2 for bootstrapped resampling. Panels a, b, c, and d correspond to 50, 100, 250, and 500 observations, respectively.

Table 2. MSD test results at the $\alpha = 0.05$ level. l_t and u_t refer to the lower and upper confidence limits, respectively.

Method	SAMPLE SIZE 50			SAMPLE SIZE 100			SAMPLE SIZE 250			SAMPLE SIZE 500		
	$\bar{Y}_t$	l_t	u_t	$\bar{Y}_t$	l_t	u_t	$\bar{Y}_t$	l_t	u_t	$\bar{Y}_t$	l_t	u_t
SLIC	.2118	.1953	.2283	.2659	.2504	.2813	.3102	.3013	.3192	.3489	.3356	.3621
SLED	.2094	.1929	.2259	.2659	.2504	.2813	.3102	.3013	.3192	.3496	.3364	.3628
SLT	.2088	.1923	.2253	.2485	.2330	.2640	.2886	.2796	.2976	.3096	.2964	.3228
CLIC	.4182	.4017	.4347	.5264	.5109	.5419	.5875	.5786	.5965	.6229	.6097	.6361
CLED	.4641	.4476	.4806	.5481	.5326	.5636	.5979	.5889	.6069	.6208	.6076	.6340
CLT	.4236	.4071	.4401	.4633	.4478	.4788	.5092	.5002	.5181	.5360	.5228	.5492
AL1IC	.4699	.4534	.4864	.5487	.5332	.5642	.6140	.6050	.6229	.6410	.6278	.6543
AL1ED	.4967	.4802	.5132	.5655	.5500	.5810	.6189	.6099	.6279	.6292	.6160	.6424
AL2IC	.4641	.4476	.4806	.5559	.5404	.5714	.6138	.6049	.6228	.6352	.6220	.6485
AL2ED	.4734	.4569	.4899	.5552	.5398	.5707	.6225	.6135	.6315	.6416	.6284	.6548
AL2T	.3367	.3202	.3532	.3858	.3703	.4013	.4304	.4214	.4393	.4487	.4355	.4620
WIC	.4476	.4311	.4641	.5356	.5201	.5511	.6108	.6018	.6198	.6385	.6252	.6517
WED	.5037	.4872	.5202	.5743	.5588	.5897	.6285	.6196	.6375	.6518	.6386	.6650
WT	.4992	.4827	.5157	.5691	.5537	.5846	.6233	.6143	.6323	.6431	.6299	.6563
KMS	.5204	.5039	.5369	.5541	.5386	.5695	.6152	.6062	.6242	.6335	.6202	.6467
KMR	.5234	.5069	.5399	.5775	.5620	.5930	.6318	.6228	.6407	.6551	.6419	.6684
KMG	.5213	.5048	.5378	.5778	.5623	.5932	.6341	.6252	.6431	.6576	.6444	.6708
NAS	.4869	.4704	.5034	.5922	.5767	.6077	.6479	.6389	.6568	.6655	.6523	.6788
NAR	.5048	.4883	.5213	.5919	.5764	.6074	.6479	.6390	.6569	.6688	.6556	.6820
NAG	.5068	.4903	.5233	.5925	.5771	.6080	.6483	.6393	.6573	.6683	.6551	.6815
PCOR	.5184	.5019	.5349	.6106	.5951	.6261	.6893	.6803	.6983	.7553	.7121	.7385
PCOB	.5168	.5003	.5333	.6118	.5963	.6273	.6902	.6812	.6991	.7262	.7130	.7395

5. CONCLUSIONS

The intercomparison of various clustering techniques is of interest to geophysical researchers. This work used a well-studied North American growing season precipitation data set, to investigate the capability of various clustering methods to recover that input data structure and test the applicability of a large range of clustering algorithms for regionalization. Overall, the findings largely confirmed what was known about SL techniques -- they lead to chaining solutions. The CL methods also yielded inferior regionalizations. In general, the theta angle (T) was found to negatively impact SL, CL, AL1, and AL2 methods, yet it was equally accurate with the IC and ED coefficients tested for W. Both the IC and ED dissimilarity measures were about as accurate for all hierarchical techniques. However, the nonhierarchical techniques tested (KM and NA) generally outperformed the top hierarchical ones (AL1, AL2, W). Overall, the NA were thought to be the best methods as they showed no sensitivity to initial partition, as did KM.

Bootstrapped resampling at 50, 100, 250, and 500 observations yield an interesting picture of these techniques. At the fewest samples, several techniques had problems with fragmenting of coherent precipitation regions, notably the CL, AL1, AL2IC, and WIC methods. These instances decreased as a function of sample size. Moreover, all techniques medians rose dramatically as more observations were admitted. Use of the MSD test allowed us to separate those solutions that had significantly higher biserial correlation matches than others, confirming the findings for the full data set.

The PC results are included, yet we hesitate to draw too many comparisons between them and the CA methods. Isoplethed loadings of PCs have overlapping classifications where some stations may be assigned to multiple regions and others may not be classified to any regions. The CA methods tested herein all classify each station to only one of the regions and there are no unclassified stations. As a result, the regions tend to be irregularly shaped under CA. These do not always match the interstation correlations well and, hence, lead to lower biserial correlations. This points to the potential for fuzzy clustering methods which yield clusters with gradients (similar to PCA). The regionalizations provided by the various CA methods tend to be different, both for any one algorithm and for any one dissimilarity coefficient. This instability leads to a problematic physical interpretation as one needs to consider which one is most meaningful? Based on the relatively larger stability of NA methods, we recommend them for precipitation regionalization where hard clusters are desired. In cases where an overlapping regionalization is acceptable, PCA is recommended.

ACKNOWLEDGEMENTS

This research was supported by NSF grant ATM 89-08545 and EPA Cooperative Agreement CR-816318-02-0. Some of the computational support was provided by the National Center for Supercomputing Applications. We wish to thank Ginger Rollins for typing the manuscript.

REFERENCES

Blashfield, R.K., 1976: Mixture model tests of cluster analysis: accuracy of four agglomerative hierarchical methods. *Psychological Bulletin, 83*, 377-388.

Cormack, R.M., 1971: A review of classification. *J. Royal Statistical Soc. (Series A), 134*, 321-367.

Easterling, D.R., 1989: Regionalization of thunderstorm rainfall in the contiguous United States. *Int. J. Climatol., 9*, 567-579.

Gadgil, S and N.V. Joshi, 1983: Climatic clusters of the Indian region. *J. Climatol., 3*, 47-63.

Gong, X.-F. and M.B. Richman, 1992: An examination of cluster analysis methodology in geophysical reserach. Part I: precipitation data. Submitted to *J. Climate.*

Goosens, C., 1985: Principal component analysis of Mediterranean rainfall. *J. Climatol., 5*, 379-388.

Gower, J.A., 1967: A comparison of some methods of cluster analysis. *Biometrics, 23*, 623-627.

Johnson, S.C., 1967: Hierarchical clustering schemes. *Psychometrika, 32*, 261-274.

Kikuchihara, H., 1984: The regional division of Japan by precipitation. *The Geophysical Magazine, 41*, 61-157.

Kuiper, F.K. and L.A. Fisher, 1975: A Monte Carlo comparison of six clustering procedures. *Biometrics, 31*, 777-783.

Milligan, G.W., 1980: An examination of the effect of six types of error perturbation of 15 clustering algorithms. *Psychometrika, 45*, 325-342.

Reich, V.T., 1986: Die raumliche struktur des niederschlagsfeldezs hierachische gruppierungs verfahren und dendrogramme. *Zeitschrift fur Meteorologie, 36*, 293-333.

Richman, M.B. and P.J. Lamb, 1987: Pattern analysis of growing season precipitation in southern Canada. *Atmosphere-Ocean, 25*, 137-158.

Richman, M.B. and P.J. Lamb, 1985: Climate pattern analysis of three- and seven-day summer rainfall in the central United States: Some methodological considerations and a regionalization. *J. Climate Applied Meteor., 24*, 1325-1343.

Richman, M.B., P.J. Lamb, and J.R. Angel, 1991: Relationships between monthly precipitation over central and eastern North America and Southern Oscillation. *Proceedings of the Fifteenth Annual Climate Diagnostics Workshop*, U.S. Dept. of Commerce, Washington, D.C., 373-383.

Ronberg, B. and W.-C. Wang, 1987: Climate patterns derived from Chinese proxy precipitation records: An evaluation of the station networks and statistical techniques. *J. Climatol., 7*, 391-416.

Sokal, R.R. and F.J. Rohlf, 1981: *Biometry; The Principles and Practice of Statistics in Biological Research*, San Francisco, W.H. Freeman, 776 pp.

Winkler, J.A., 1985: Regionalization of the diurnal distribution of summertime heavy precipitation. *Preprints, Sixth Conf. on Hydrometeorology*, American Meteorological Society, Boston, MA, 9-16.

J7.3

CLUSTER ANALYSES OF RADIANCE DATA MEASURED BY SATELLITE AND COMPUTED FROM FORECAST PROFILES

Dongsoo Kim
Cooperative Institute for Research in Environmental Sciences (CIRES)
University of Colorado/NOAA, Boulder, CO 80309-0449

1. INTRODUCTION

The retrieval of temperature and humidity from radiance data measured by satellite can be expressed in a highly simplified form,

$$T = T^f + H * (R - R^f), \qquad (1)$$

where T is the temperature/humidity profile (hereafter called profile), R is the radiance (converted to brightness temperature), T^f is the first guess, R^f is the radiance computed from T^f, and H is a retrieval operator that translates the radiance increment to a vertically distributed profile increment. Several options for the first guess include climatology, field average, or a numerical weather prediction (NWP) model forecast.

Equation (1) is a typical expression used in optimum interpolation (OI), where H depends upon an empirically determined first guess error covariance; H interpolates values at observation points to grid points. The empirical covariance is usually determined by fitting an analytic function to a large data set. It is also common to stratify the data by season or month. The coefficients of the fitting function vary with the data sample used.

Satellites measure upwelling radiance from the ground through the entire atmosphere; therefore, the radiance is an integrated quantity for each spectral wavelength. In discretized form, we can write

$$R = K * T, \qquad (2)$$

where K converts profiles to radiances at a given spectral wavelength. The process described in (2) is called the forward computation, where the forward operator K contains the vertical weighting values whose peak weight level varies depending on which channel is selected.

Recent developments in satellite data retrieval have focused on new approaches in obtaining the first guess of either a profile or a radiance. For example, researchers at the French Laboratoire de Meteorologie Dynamique (LMD) demonstrated air mass classifications (Scott et al., 1991). NOAA's National Environmental Satellite, Data, and Information Service (NESDIS) is also planning an operational classification-based regression method (Bloom et al., 1992). These developments have attracted the interest of researchers at the Forecast Systems Laboratory (FSL), who have successfully assimilated observations from a variety of sources such as wind profiler and aircraft data; and satellite data in the future (Benjamin et al., 1991).

A new effort to incorporate polar orbiting TIROS Operational Vertical Sounder (TOVS) data at FSL is based on the direct use of radiance data rather than retrieved data. Therefore, understanding the error characteristics of measured radiances and radiances computed from first guess profiles is a first step toward successful assimilation by means of statistical analysis.

This study investigates a method to update a model forecast with real-time TOVS radiance data. The method is based on the cluster analysis of radiance increments (measured minus computed values from a NWP model forecast).

2. COLLOCATED DATA SET

FSL operates both a mesoscale numerical model with a 3-hour assimilation cycle called the Mesoscale Analysis and Prediction System (MAPS) and a direct readout TIROS ground receiving system. This environment allows TOVS data to be assimilated by constructing an empirical covariance H from a locally processed data base. With this purpose in mind, FSL began collecting data valid at 0000 UTC and 1200 UTC during March 1991. The data set comprises:

(1) NOAA-10 TOVS measured radiances within 20 km of RAOB sites,

(2) radiances computed from RAOB profiles,

(3) radiances computed from MAPS 3-hour forecast profiles at the RAOB sites,

(4) RAOB profiles at mandatory levels, and

(5) MAPS 3-hour forecasts at the mandatory levels.

Items (4) and (5) can verify MAPS forecast errors, items (1) and (3) are used to update model forecasts, and items (2) and (3) can be used to evaluate errors of the forward model. Thus, items (1) through (5) will help construct the retrieval operator H, which translates radiances into vertical profiles.

The data set includes 459 *objects*; each has 11 *variables* consisting of High-resolution InfraRed Sounder (HIRS) channels 4, 5, 6, 10, 11, 12, 13, 14 and 15, and Microwave Sounding Unit (MSU) channels 2 and 3. The variable characteristics are: HIRS channels 4, 5 and 6 sense temperatures near the 400, 600 and 800 hPa levels, respectively, in the 15 μm region; HIRS channels 10, 11 and 12 sense humidity near the 900, 700 and 500 hPa levels; HIRS channels 13, 14 and 15 sense temperature near the 1000, 950 and 700 hPa in the 4.3 μm region, and MSU channels 2 and 3 sense temperature near 700 and 300 hPa levels in the mircrowave region, which is not contaminated by cloud. The peak weight levels of HIRS channels are not constant when clouds are present. Satellites measure radiances at and above the cloud top surface. Variables with peak weights above 150 hPa are not considered because MAPS does not forecast above this level even though TOVS measured radiances are available. When analyzing these data, one should keep in mind error sources such as: (1) sampling errors due to the data window, +2 hours and 20 km, (2) cloud contamination, (3) an imperfect forward model, (4) errors in the radiance and RAOB measurements, and (5) deficiencies in the MAPS forecast model.

3. DATA ANALYSES

3.1 Preparation

As a part of the forward computation, cloud parameters (cloud top pressure and cloud amount) are estimated by fitting computed HIRS channels 7 and 8 from MAPS forecast profiles to measured radiances of HIRS channels 7 and 8 (Eyre and Menzel, 1989). Any object whose measured HIRS channel 8 radiance differs from MAPS-computed radiance for the channel by 2^o K is deleted from the data base. Out of 459, 443 objects passed the quality control check. The second column of Table 1 shows standard deviations of radiance increment (measured minus MAPS-computed) of 443 objects for each variable. The humidity channels (HIRS 10, 11 and 12) show large values, indicating that humidity profiles need to be improved more rigorously than temperature profiles.

3.2 Cluster analysis

The purpose of clustering is to group objects that have internal homogeneity, so that objects in other groups are dissimilar. The preliminary analysis of the data set (e.g., multivariate scatter plots) showed no distinctive or mutually exclusive clusters. In order to avoid results dominated by a few variables (e.g., HIRS 11, 12), all variables were standardized. Then, the *distance* to measure dissimilarities was computed for all pairs of objects. The distance is defined by

$$d_{ij} = \left[\sum_{k=1}^{11} (R_{ik} - R_{jk})^p\right]^{1/p}, i,j = 1,...,n \qquad (3)$$

where n is number of objects and R is the radiance increment (measured radiance minus MAPS-computed radiance). p is set to 2, thus making d_{ij} a Euclidean distance.

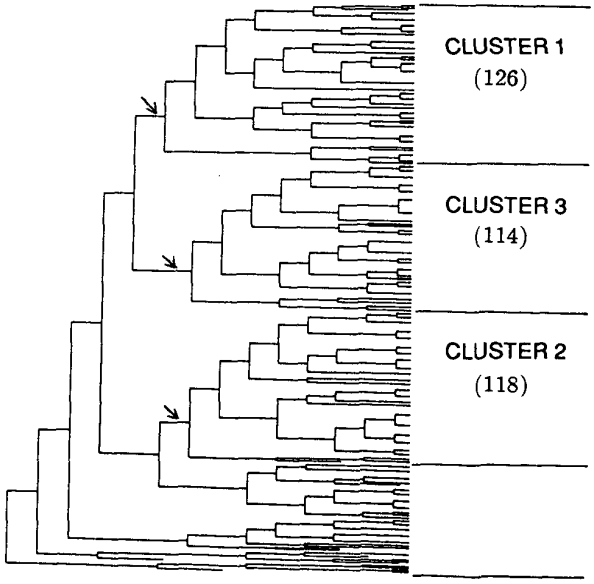


Fig. 1. Dendrogram of hierarchical clustering method with complete-linkage. Arrows represent snipped branches which produce clusters. Biggest three clusters which make up 81 % of all objects are considered. Numbers in parentheses are sample size.

In hierarchical clustering of n objects, there are $n-1$ mergers and 2^{n-1} possible orderings for the leaves in a cluster tree. The method of merging each object in this study is *complete-linkage*, in which distance between groups is determined from the most remote pair of objects (Gordon, 1981). Figure 1 shows the dendrogram of hierarchical clustering. Clusters with a sub-sample size larger than 100, which make up 81 % of the 443 objects, are considered further. Heights of the cutting trees are indicated by arrows; the rest are regarded as unrepresentative.

The clustering results are very much dependent on the methodology. One worthwhile exercise is to look at the distribution of objects in each cluster. Since we have RAOB profiles of corresponding objects, profile increments (RAOB minus MAPS) are compared with radiance increments (measured minus MAPS-computed). Figure 2 shows vertical distributions of temperature (Fig. 2a) and the mixing ratio (Fig. 2b) relative to the level means of the 443 objects. The biases in the profiles are consistent with known weighting functions of radiance channels. For example, HIRS 4, which peaks near 400 hPa, is positive in cluster 3 (in Fig. 2c), and the temperature profile for the 300-500 hPa layer is also positive (in Fig. 2a). A second example applies to cluster 2.

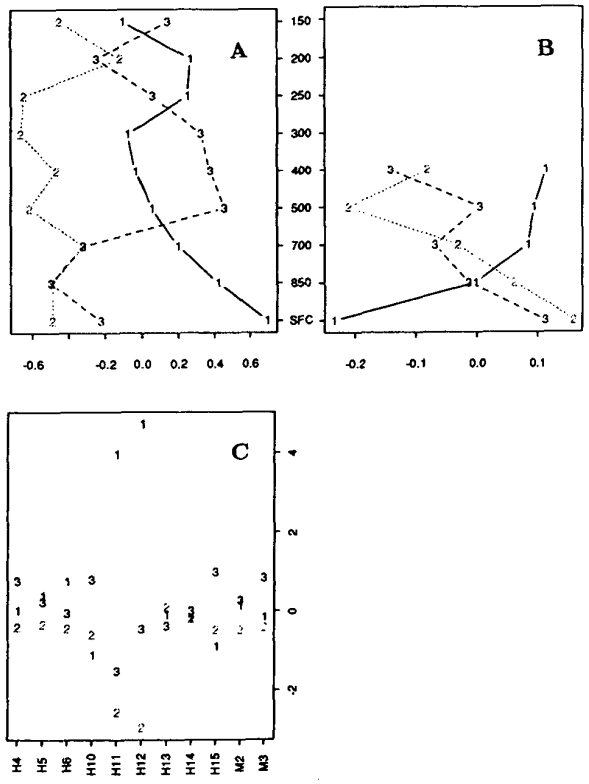


Fig. 2. Vertical distribution of mean temperature (2a) and mixing ratio (2b) increments (RAOB minus MAPS) for each cluster relative to the level means of all objects. The unit of temperature is degree and mixing ratio is log(g/Kg). Fig. 2c is mean radiance increments (measured minus MAPS-computed) for each cluster relative to the channel mean of all objects. Numbers correspond to cluster identification.

Increments for HIRS 4, 5 and 6 which peak in the troposphere, are negative in Fig. 2c. Corresponingly, the temperature increment in Fig. 2a, is also negative in the troposphere. As a final example, both HIRS-11 and 12 are moisture channels with peaks near 700 hPa and 500 hPa. The large positive bias in cluster 1 (Fig. 2c) is consistent with the positive bias in the humidity profile for cluster 1 in Fig. 2b. This consistency suggests that vertically correlated temperature and humidity errors in MAPS forecasts can be updated by satellite-measured radiance data. However, this is valid only when clusters for new data sets can be properly identified.

Table 1 shows the standard deviations of the variables for each cluster. Each cluster has smaller standard deviations than those of all objects in most variables. Exceptions are HIRS channel 10 and 13 of cluster 3 and HIRS channel 14 of cluster 1.

Table 1

Standard deviations of radiance increment (K) for each cluster as determined in Sec. 3.2

Variable	All(443)	C1(126)	C2(118)	C3(114)
H4	0.839	0.498	0.535	0.668
H5	0.766	0.675	0.557	0.541
H6	0.866	0.765	0.553	0.531
H10	1.403	0.639	1.212	1.587
H11	3.707	3.129	2.249	1.613
H12	4.792	4.156	3.889	2.414
H13	0.839	0.672	0.597	0.949
H14	0.724	0.753	0.606	0.695
H15	1.176	0.644	0.980	1.017
M2	0.831	0.827	0.778	0.766
M3	0.910	0.778	0.697	0.772

3.3 Assignment

Now, we define a mesaure $S(k)$ which can be used to assign new sample data tp previously determined clusters.

$$S(k) = \sum_{i,j} (R_i - \overline{R}_i) C_{ij}^{-1}(k) (R_j - \overline{R}_j), i, j = 1, .., 11. \quad (4)$$

The index k takes on the values 1, 2, 3 corrsponding to the clusters just defined or it takes the value zero for the historical sample of 443. Indices i and j take on the values from 1 to 11, corresponding to the 11 radiometric channels. For each cluster (k = 0, 1, 2, 3), a covariance matrix $C(k)$ is computed ahead of time, relating statistically the radiance increments for every pair of channels. $C_{ij}^{-1}(k)$ is an element of the inverse covariance matrix for the k^{th} cluster. R_i is the radiance increment in the i^{th} channel for a particular sounding in the new data sample. $\overline{R_i}$ is the average radiance increment for all members of the new sample, which might consist of a swath of soundings over a fairly large area, as depicted in Fig. 3. After computation of $S(0)$ through $S(3)$, the sounding in question is assigned to that cluster which produces the smallest value of S.

The assignment of each cluster is applied to a test set of radiances from a TIROS pass over the midwestern United States at 0035 UTC on 25 March 1991 and the MAPS 3-hour forecast valid at 0000 UTC. At each sounding location the measure $S(k), k = 0, 1, 2, 3$ is evaluated, from which the minimizer is assigned to the cluster identification. Figure 3a shows the HIRS-11 radiance increments (measured minus MAPS-computed) with field mean subtracted. The soundings with bias errors larger than ± 1.5 K are identified as non-representative clusters and are marked with a $\emptyset$ in Fig. 3b. The two figures do not match exactly because Fig. 3b is the result of 11 variables. However, spatial coherence of cluster indices is an interesting feature.

3.4 Prediction

The eventual goal of TOVS data assimilation is to obtain a profile increment ($T - T^f$ in Eq.(1)) from the radiance increment (measured minus MAPS-computed) and retrieval operator H. The retrieval operator can be either an empirical cross-covariance or a set of parameters of a multivariate regression model. In either case, the prediction procedure should be robust against assignment of new soundings to the wrong cluster. To guard against wrong assignments, one can constrain the parameter space with prior information. The uncertainty of the forward model, i.e., covariance of RAOB-computed radiance minus MAPS-computed radiance, can serve as the constraint.

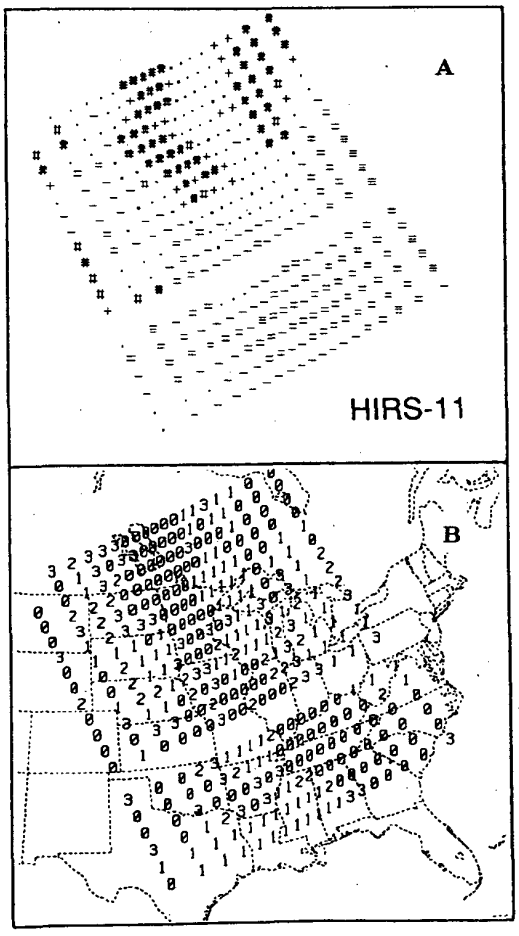


Fig. 3. Radiance increment (measured minus MAPS-computed) field of HIRS channel 11 (3a) in degree Kelvin. The symbols are: # > 1.5, 1.0 > # ≥ 1.5, 0.5 > + ≥ 1.0, −0.5 ≥ · ≥ 0.5, −1.0 ≥ − > 0.5, −1.5 ≥=> 1.0, −1.5 >≡, and Fig.3b shows cluster identification.

4. CONCLUSION

Data assimilation involves the blending of a NWP model and observations, since neither model nor observations are error-free. Remotely sensed satellite data have poor vertical resolution but excellent horizontal resolution, which must utilized. Vertically correlated errors in the forecast profiles lead to errors in the computed radiances for those channels whose peak weight levels are nearby. If these profile errors are horizontally correlated, this spatial coherence will be detected by satellite observations.

From the collocated data set collected in March 1991, three clusters based on the radiance increments (measured minus MAPS-computed) showed consistent profile increments (RAOB minus MAPS). The cause of vertically correlated errors in the model will be investigated, but it currently provides a basis for updating model forecasts by satellite radiance data.

More refinement of the clustering method is needed. Inconsistencies between the Euclidean distance and assignment measure occur when subtracting the mean. Time series of 443 objects (not shown) indicated a non-stationary trend in HIRS channel 4 and MSU channels 2 and 3 (i.e., mid-troposheric layers). This trend was due to a cold outbreak during the middle of March 1991. When computing Eculidean distance, variables must be detrended before standardization. Similarly, test data sets must be detrended, if necessary, (e.g., for latitudinal dependency) before computing the assignment measure.

The NWP model continues to be improved. Even though the evaluation of model performance is usually done through reference to the RAOB network, this does not show meso-scale horizontal variability of model errors because of poor horizontal resolution. However, satellite data provide better resolution for the analysis of meso-scale model error which has spatial coherence when viewed from the satellite perspective.

5. ACKNOWLEDGMENTS

All statistical analyses are done using the S-PLUS by Statistical Sciences, Inc. The author thanks Tom Schlatter, John Snook and Nita Fullerton of FSL for their careful review and editorial assistance.

6. REFERENCES

Benjamin, S., T. Smith, P. Miller, D. Kim, T. Schlatter and R. Bleck, 1991: Recent improvements in the MAPS isentropic-sigma data assimilation system. *Preprints, Ninth AMS Conference on Numerical Weather Prediction.* 14-18, October 1991, Denver, CO, 118-121.

Bloom, H., Wilson, L., Reale, A. and L. McMillin, 1992: Operational TOVS soundings using a radiance classification retrieval approach. *Preprints, Sixth AMS Conference on Satellite Meteorology and Oceanography. 5-10*, January 1992, Atlanta, GA, 48-51.

Eyre, J., and P. Menzel, 1989: Retrieval of cloud parameters from satellite sounder data: a simulation study, *Mon. Wea. Rev.*, **28**, 267-275.

Gordon, A. D., 1981: *Classification.* Chapman and Hall, 193 pp.

Scott, N., A. Chedin, V. Achard, B. Bonnet, F. Cheruy, C. Claud, J. Escobar, N. Husson, H. Rieu, Y. Tahani and B. Tournier, 1991: Recent advances in the 3D thermodynamic analysis of the Earth system through the "3I" algorithm. Extension to the second generation vertical sounders. *Preprints, Sixth International TOVS Study Conference*, 1-6 May 1991, Airlie, VA, 425-467.

CLUSTER ANALYSIS IN CLIMATE VARIABILITY RESEARCH - RECENT DEVELOPMENTS

Klaus Wolter

CIRES, University of Colorado
Boulder, Colorado

Dramatic increases in greenhouse gases have caused great concern about global climate change. Therefore, the need for empirical analyses that extract the large-scale signal from existing *noisy* but long climate records is more pressing than ever. This paper will discuss the application of cluster analysis of variables (CAV) to this problem area, focusing on recent developments (see Wolter, 1989, for previous review). While CAV partitions data according to second-moment statistics, the related technique of cluster analysis of cases (CAC) classifies data using first-moment statistics. Since we are interested in monitoring common areas of climatic change, analyzing correlation structures (CAV) seems more appropriate than CAC.

Examples presented here stem from the Comprehensive Ocean-Atmosphere Data Set (COADS; Woodruff et al., 1987), but most of the arguments easily transfer to other climate data sets. Two main areas of interest are pursued here, namely a comparison of CAV with other signal detection techniques, and improvements in technical aspects of CAV.

Point-correlation maps, although tedious to analyze for large grids, reveal the *anisotropic* nature of many climate fields. A given iso-correlation contour often traces the outline of a cluster, even if the reference point is not in the center of that cluster. Strong clusters are therefore easily verified in such analyses. Principal component analysis (PCA) has often been used to analyze large-scale climate data sets directly. Loading centers of the leading PC's (especially for rotated solutions) regularly coincide with clusters. Tandem analyses of PCA on clusters show that resultant PC's are virtually indistinguishable from directly analyzed PC's.

Three different aspects of recent improvements in CAV will be discussed in this talk. First, cross-validation of cluster fields has become feasible with the recent increase in computer power. Stability of cluster solutions with respect to slight alterations in sample size and methodology thus serves to increase our confidence in the analyzed clusters. By the same token, it is easier to denote the limits of applicability of CAV where this stability breaks down.

Second, theoretical work on the optimal number of clusters (e.g., Milligan and Cooper, 1985) may help to reduce the subjectivity of choosing the proper number of clusters for the task at hand. Akin to the problem of selecting the right number of PC's to retain for rotation, there is no single criterion available that delivers this decision.

Finally, since CAV maximizes the internal correlation within each cluster, it increases the reliability of climate indices based on such clusters. Signal enhancement of these indices is accomplished not only by the superior averaging within each cluster, but also by the reduced impact of missing data (ref. Katz and Glantz, 1986). Assuming stationarity, sparse data sets such as 19th century COADS can thus be analyzed longer back in time than otherwise possible.

Katz, R. W., and M. H. Glantz, 1986: Anatomy of a rainfall index. *Mon. Wea. Rev.*, **114**, 764 - 771.

Milligan, G. W., and M. C. Cooper, 1985: An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, **50**, 159 - 179.

Wolter, K., 1989: Cluster analysis of variables in climate research. *Proceedings of the 4th International Meeting on Statistical Climatology*, Rotorua, New Zealand, March 1989, pp. 198 - 202.

Woodruff, S. D., Slutz, R. J., Jenne, R. L., and P. M. Steurer, 1987: A comprehensive ocean-atmosphere data set. *Bull. Amer. Meteor. Soc.*, **68**, 1239 - 1250.

J6.2

RELATIVE OPERATING CHARACTERISTIC, SUFFICIENCY, AND VALUE OF FORECASTS

Roman Krzysztofowicz

University of Virginia
Charlottesville, Virginia

ABSTRACT

There is a growing quest to connect statistical measures of forecast quality, which provide direct feedback to meteorologists and climatologists, with the ex ante economic value of forecasts -- a measure which is of ultimate interest to rational decision makers. This talk will explore such connections in a situation wherein forecasts of some future state of weather or climate are used to make a binary decision (e.g., take action or do not take action).

Consider two alternative forecast systems, say A and B, producing forecasts of the same state. From the viewpoint of a rational user, system A is at least as preferred as system B, if the ex ante economic value of forecasts produced by A is no smaller than the ex ante economic value of forecasts produced by B. From the collective viewpoint of all rational users, A is at least as preferred as B, if this is the preference order from the viewpoint of each individual user. In such a case, system A is said to dominate system B.

From the viewpoint of a Bayesian statistician, forecast systems can be compared in terms of a sufficiency relation, or in terms of their Relative Operating Characteristics (ROC) -- the concept widely used in signal detection theory. Our discussion will concentrate on the equivalence of the following three statements: (i) system A dominates system B; (ii) system A is sufficient for system B; and (iii) ROC of system A is superior to ROC of system B. Proper statistical methods of constructing ROC's for weather and climate forecast systems will be highlighted.

J4.1 THE ROLE OF DECISION MAKER EXPECTATIONS IN VALUING CLIMATE INFORMATION

Michael A. Mazzocco
University of Illinois

Bruce J. Sherrick
University of Illinois

Steven T. Sonka
University of Illinois

Peter J. Lamb
University of Oklahoma

1. INTRODUCTION

Information technology has been developing rapidly for more than a decade. The usefulness or value of information is sometimes questioned to help guide the development of information systems. This paper reports on a stream of research emanating from Lamb's (1981, p. 1001) call for expanded research on "how climate forecast schemes could/should be used." At issue has been the identification of factors affecting the value of climate information. Prior work, including that of others, has focused on the economic value of climate information in agricultural production (Mjelde et al., 1988; Katz, Murphy and Winkler, 1982; Byerlee and Anderson, 1982). Similar efforts continue to investigate which decisions are benefitted by climate forecasts and the characteristics of the forecasts that give rise to their value.

The methods used to value climate information are based on Hilton (1984). They describe the difference in the expected value of outcomes with and without information. That is, the value of information set p_k is described as:

$$V(p_k) = \underset{x_k}{\text{Max}} \int \pi(x_k,\theta)p(\theta|k)d\theta - \underset{x}{\text{Max}} \int \pi(x,\theta)p(\theta)d\theta \quad (1)$$

where the outcome π is dependent upon a decision set (x_k or x), and a stochastic event (θ). The decision maker's perception of the probability distribution of θ is altered by the information set p_k, or, more specifically, by the prediction k. The outcome in this model, π, represents profit or return above variable costs (Mjelde et al.). Thus, the information only has value to the extent it alters the optimal decision set such that the solution to the first term is not equal to the solution to the second term, i.e. $x_k^* \neq x^*$. The value of the information system is the expected value of equation (1), where the expectation is taken over possible predictions coming from the information system. This model is easily adaptable to agricultural production, where climate is the stochastic event and decisions relate to the use of variable inputs. This model may also be written in terms of the utility of π, which is not done here for simplicity.

Most research along this vein has endowed the decision maker with perfect knowledge of historic climate probabilities. Thus, historic climate probabilities are used to establish the probability distribution of θ in the "no information" scenario. This assumption of perfect historical prior probabilities is challenged here. The theory behind the use of alternative prior probabilities is reviewed in the next section, with a discussion of venture theory from the literature of decision sciences. This is followed by a comparison of information value when alternative prior probabilities are used. Then a report on an elicitation of climate expectations from a sample of agricultural producers is given, comparing their prior beliefs to historic probabilities. The last section describes an application of calibration theory, in which an assessment of the differences between true and subjective probabilities provides support for adopting the assumptions of venture theory.

2. VENTURE THEORY: DECISION WEIGHTS INSTEAD OF PROBABILITY

The process of decision making under uncertainty has a rich history in the literature of psychology, economics, and management science. Most models have used mathematical probabilities to prescribe behavior for risky choice. However, the use of probabilities breaks down in *describing* observed behavior of individuals making risky choices (Allais, 1953; Ellsberg, 1961). Where risk characterizes a situation if the decision maker knows the relative chance of each outcome, ignorance describes a situation if the decision maker has "no basis whatsoever on which to judge the relative likelihood of potential outcomes of [a] decision (Yates and Zukowski, 1976, p. 19)." Alternatively, ambiguous decisions lie between these two extremes; there is some basis for assessing relative probabilities, but not with precise confidence.

Some behavioral scientists have chosen to think of ambiguous decisions as problems in second order probabilities (e.g. Becker and Brownson, 1964). With this characterization, the solution to an ambiguous decision can be derived from compound expectations. However, Yates and Zukowski clearly demonstrate that observed behavior is not compatible with the notion of second order probabilities.

Alternatively, Einhorn and Hogarth (1985, 1987) described a model of *decision making* under ambiguity which employs an "anchoring-and-adjustment strategy in which an initial probability is used as the anchor (or starting point) and adjustments are made for ambiguity (1987, p. 46)." The source of the initial probability may be any information, historical or otherwise, available to the decision maker. The subjective probability used in decision making, S(p), is given by:

$$S(p) = p_A + k \quad (2)$$
$$k = k_g + k_s \quad (3)$$

where p_A is the anchor probability and k is the adjustment. The adjustment is made from a mental evaluation of higher and lower values of p, where k_g is the effect of simulating values of p greater than p_A and k_s is the effect of simulating smaller values of p.

Hogarth and Einhorn (1990) further expanded this theory to include other factors that affect the adjustment. Because these factors affect the assessment of subjective probabilities, they refer to the adjusted probabilities as decision weights, which are arrived at by "venturing" or mentally simulating outcomes. Adjustment factors include outcome uncertainty, the degree of ambiguity, the context of the decision, and the sign and size of payoffs.

Outcome uncertainty refers to the number of outcomes a decision maker anticipates experiencing. Consider a midwestern U.S. agricultural producer making a decision on the timing of fertilizer application. If he applies fertilizer in the fall there is some chance winter precipitation will be sufficiently heavy to leach nitrogen from the soil, rendering his expenditure

Table 1. Expected Value of Perfect Forecast vs. Historical and Ambiguous Priors, East Central Illinois Corn ($)

Prior Belief	320 Acres	Per acre
15-yr historical probability	1,000	3.12
Ambiguous, most recent 3 years	3,084	9.64
Ambiguous, 100 lbs. spring N	3,133	9.79
Ambiguous, 150 lbs. fall N	5,237	16.37

Source: Mazzocco (1989)

ineffective. In this situation he experiences one outcome from the decision. Say the fertilizer application cost is $10,000 and that there is a 50 percent chance that precipitation will be above some specified level that will cause the $10,000 expenditure to be revisited in the spring. It is of little value to know that the expected outcome is -$5,000. Contrast this to a gambler at a slot machine with 1,000 coins, where the expected value of the payoff is more meaningful because of the number of plays; the net outcome is likely to be closer to the expected the value. Thus, there is less outcome uncertainty.

The maintained hypothesis is that for individuals exhibiting cautious behavior (akin to risk aversion), the greater the outcome uncertainty, the more mental simulation takes place that results in the overweighting of losses and the underweighting of gains. The same is true for the degree of ambiguity; the less confident a decision maker is in his assessment of anchoring probabilities, the more mental simulation he performs. The degree of underweighting or overweighting is affected by the size of the payoff. The degree of caution exercised in the process of mental simulation is a function of the decision context. It is interesting that unlike probabilities, the decision weights need not sum to one, which is consistent with the Allais and Ellsberg decision paradoxes.

A few observations of this construction of a decision problem are noteworthy. First, it allows for the initial decision weights to come from any source, including past experience, historical data, or an expert opinion. Second, it is not inconsistent with expected utility theory, but more encompassing. Third, it indicates that individuals need not respond to information in the same manner. Different individuals will possess different degrees of ambiguity on different problems. Thus, the decision process will be characterized by different amounts of simulation, resulting in different degrees of underweighting gains and overweighting losses in response to information designed to lower ambiguity.

3. CLIMATE INFORMATION VALUE UNDER ALTERNATIVE PRIORS[1]

Recall from equation (1) that the value of information fundamentally depends on the prior probabilities (θ) and the adjusted probabilities given the forecast ($\theta|k$). To illustrate the differences in climate information value, consider an average farm in east central Illinois, growing 320 acres of corn. Further consider that the climate parameters of interest to the decision maker are 1) the amount of winter precipitation (as described earlier), 2) July rainfall, and 3) the general summer growing conditions (a composite of precipitation, solar radiation, temperature, and pan evaporation). Using historic climate probabilities, this decision maker would plan to apply 150 pounds of nitrogen fertilizer per acre in the spring. In fact, this is the observed average application rate in this area, as well as a reasonable approximation of average corn acreage per farm (Illinois Department of Agriculture).

However, other practices are observed. These include applications of 150 pounds of nitrogen in the fall and only 100 pounds of nitrogen in the spring. Wise and Yotopoulos (1968) ascribe similar observations in developing countries to tradition. However, extremely minor modifications in the historic probabilities of the three climate parameters of interest cast these practices as optimal decisions. Furthermore, it has been suggested that recent events affect decision weights more than distant events, indicating an analysis of priors built on recent history. The value of climate information in these cases will depend upon how it affects the decision weights defined by venture theory.

Table 1 contains the results of a model simulating corn production over actual climate from 1971 to 1985. A decision maker endowed with the full range of pertinent historical climate information (15-year prior) would expect to receive a maximum average annual benefit of $1,000 for adopting the information contained in a perfect forecast of the relevant events described above. Alternatively, a decision maker whose decision weights were based on the most recent three years' experience would expect to receive a maximum average annual benefit from the forecast of $3,084, or $9.64 per acre of corn. Again, this assumes that the forecast completely replaces any prior notion of the decision weights. For a farmer whose ambiguity leads to 150 lbs. per acre of fall applied nitrogen, the maximum average annual value of the forecast rises to $5,237, or $16.37 per acre.

The value of changing the ambiguous decision weights to coincide with historic climate probabilities is clear. For the farm modeled in this exercise, if rational behavior leads to fall nitrogen application there is a $4,237 annual expected benefit simply from understanding the probabilities of winter precipitation. Thus, there appears to be a substantial benefit in changing farmers' perceptions of prior decision weights, bringing them more in tune with historical probability. To ascertain the frequency and magnitude of these discrepancies, research was conducted to elicit farmers' assessment of the probability of climate events that affect growing crops. The next section reports on the design and results of that survey.

4. DESCRIBING FARMER EXPECTATIONS OF CLIMATE VARIABLES

A survey was conducted to better understand the formation of farmers' climate expectations and observe their consistency with venture theory. Participating farmers were selected for their 1) cooperation with a farm record keeping association, 2) their close proximity to a single weather reporting station (to mitigate the potential effects of widely differing experiences), 3) relatively large cash grain operation, and 4) understanding of probability concepts. Personal interviewers elicited 5 to 7 fractile breaks (i.e. the level of rainfall at which the 1, 10, 25, 50, 75, 90, and 99 cumulative percentiles occurred) through a series of

questions posed in both the CDF framework and inverse CDF framework. After checking for internal consistencies, fifty-one useable surveys were collected. (Copies of the survey document are available from the authors upon request.) Specific variables of interest are April rainfall and July rainfall.[2] A large amount of April rainfall is considered a negative outcome as it delays planting. A large amount of July rainfall is considered a positive outcome as it supports crop growth. Thus, venture theory would suggest that probabilities of large April rainfall would be overweighted by respondents, while probabilities of large July rainfall would be underweighted.

Decision weights or subjective beliefs (PDF^S) are compared to the "objective" or "true" probability measure for each variable (PDF^O).[3] A variant of the Burr-3 distribution was used to model both PDF^O and each farmer's PDF^S. The Burr has zero support (negative levels are disallowed), may take on a wide range of skewness and kurtosis, and can be used to fit almost any set of unimodal data (Tadikamalla, 1980). The Burr PDF and CDF for rainfall, Y, with parameters λ, and τ are:

$$f_{BR}(Y|\lambda,\tau) = \lambda\tau Y^{-\tau-1}(1+Y^{-\tau})^{-\lambda-1} \quad \text{for } \lambda, \tau, Y > 0, \tag{4}$$

$$F_{BR}(Y|\lambda,\tau) = (1+Y^{-\tau})^{-\lambda} = (Y^{\tau}/(1+Y^{\tau}))^{\lambda} \tag{5}$$

This distribution has been used extensively in various forms to model precipitation amounts (Mielke, 1973; Mielke and Johnson, 1973), as a function for business losses (Lomax, 1954; Dubey, 1966) and by the insurance industry as a candidate for loss distributions.

4.1 Results

Historic weather data from 1903 to 1990 were used to estimate the parameters of the "true"distributions of April and July rainfall. Nonlinear least squares was used to estimate the parameters of each farmers' subjective distributions for both April and July rainfall.[4] The findings are both summarized across farmers and in terms of each farmers fit to the historic probability function. Figure 1 gives a sample of the farmers' subjective beliefs about April rainfall along with PDF^O.

Table 2 examines the cross section of farmer responses. For each fractile break elicited, the collective responses were summarized and compared to the actual. For example under April rainfall, at the 25% level (the level at which there is a 25% chance of observing less rainfall and 75% chance of observing more) the precipitation level corresponding to the true distribution is 2.34

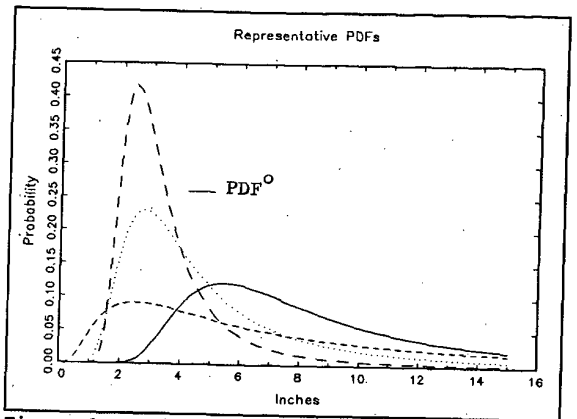


Figure 1. April Rainfall: PDF^O and Sample PDF^S

inches. In other words, there is a 75% chance of observing at least 2.34 inches of rain in the month of April. Of the farmers surveyed, 63% replied with a higher number (i.e. expected more rainfall at the 25% level), the average of all responses was 2.785 inches and the standard deviation of the response to that question was 1.003. Notice that the average expected rainfall is greater than the true at all percentile levels greater than the 10 percentile level. Further, the percentage of farmers overstating expected precipitation is greatest at the 50th percentile of the true CDF. Clearly, this "negative outcome" is overweighted in subjective probability.

However, farmers, on average, understate the incidence of rainfall in July. For example, at the 90th percentile, the true level of July rainfall is 6.74 inches, but the average of the farmers' responses was 6.49 inches with a standard deviation of 2.23. Two-thirds of the farmers understated the 50th percentile level of rainfall, thus underweighting the probability of a positive outcome.

As was discussed earlier, for individuals, the value of climate information depends upon the entire prior distribution and the process of revising expectations through the adoption of information. Further, observed decisions indicate that much of the value of climate information (as it pertains to midwestern crop production) is contained in the historic distribution. Hence we need a method to conveniently represent the differences between expectations and objective measures at all levels of the CDF.

Table 2. Summary of Farmers' Subjective PDFs on Climate Events

Climate Event	Distribution Attribute	Percentile				
		10%	25%	50%	75%	90%
APRIL Rainfall	% of farmers >= actual	31.5	63	85.1	74.1	61.1
	True (inches)	1.57	2.34	3.47	4.93	6.54
	Average response (inches)	1.415	2.785	4.444	5.806	7.461
	Std. Dev.	0.551	1.003	1.219	1.426	2.035
JULY Rainfall	% of farmers <= actual	85.2	77.8	66.7	55.6	57.4
	True (inches)	1.27	2.05	3.27	4.89	6.74
	Average response (inches)	0.813	1.781	3.008	4.614	6.489
	Std. Dev.	0.528	0.704	0.821	1.164	2.23
Snowfall	% of farmers <= actual	93	85	87	83	87
	True (inches)	11.3	14.7	20.7	29.8	37
	Average response (inches)	4.961	9.098	14.618	20.137	26.765
	Std. Dev.	3.654	5.013	6.367	7.776	9.951

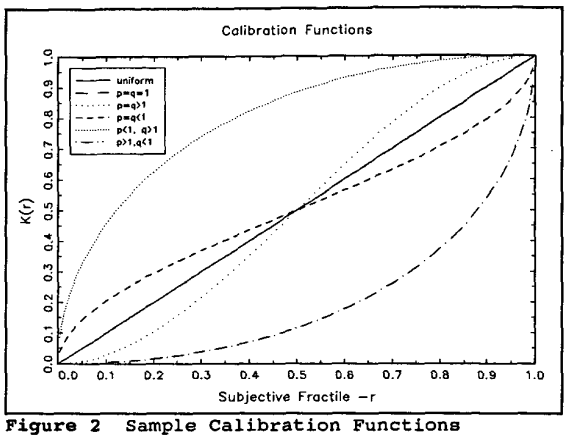


Figure 2 Sample Calibration Functions

4.2 Calibration Tests

Calibration refers to the correspondence between a predicted and an actual event. In terms of distributions, calibration describes how close the predicted and resulting functions are. Heuristically, if there were a reason for the individual expectations to yield estimated parameters that required an adjustment to correspond to the "true" parameters, then this adjustment is termed the calibration function. Specifically, if the true parameters of a distribution are $\phi(x)$ and the estimates are $F(x)$, then $K(F(x)) = \phi(x)$ implicitly defines a transformation $K(\bullet)$ of F to generate estimates, $K(F(x))$, that are well calibrated or reliable (Sherrick et al., 1990). The function $K(\bullet)$ is called the calibration function.

A simple test for calibration may be performed by testing the uniformity of K, for if $F(\bullet)$ is already well calibrated, K is simply a uniform mapping. Regions of $K(\bullet)$ with slope greater than one correspond to regions of the CDF^S that need to have mass added and regions of $K(\bullet)$ with slope less than one correspond to regions of the forecasted distribution that have too much mass and need to be decreased. A parametric form can be chosen for the calibration function and estimated using standard methods. The parameters of the estimated function can shed light into the nature of the mis-calibration (Fackler and King, 1990). For the purposes of this study, the calibration function is based on the beta distribution with density

$$K(x) = x^{p-1}(1-x)^{q-1}/\beta(p,q), \qquad (6)$$

where $\beta(p,q)$ is the beta function with parameters p and q. As noted in Fackler and King, the Beta distribution is well known, flexible and contains the uniform distribution as a special case when p=q=1, implying perfect calibration. Thus a likelihood ratio statistic is easily constructed for the hypothesis that the calibration function is uniform. Other shapes of the fitted calibration curve indicate the "reweighting" of the estimated distributions needed to correspond to those subsequently observed. At least 5 general shapes for the calibration function emerge that serve well to summarize the nature of the miscalibration displayed by each individual. Figure 2 displays the sample calibration functions corresponding to the following cases: (1) well calibrated or uniform, p=q=1; (2) underconfidence or an overstatement of dispersion, p>1, q>1; (3) overconfidence or an understatement of dispersion, p<1, q<1; (4) overstatement of location, p<1, q>1; and (5) understatement of location, p>1, q<1. The slope of the calibration function reflects the reweighting of the subjective distribution that is needed to make it correspond to the true distribution.

Calibration functions were estimated for each sample farmer's CDF^S for both April and July rainfall. The results are given in table 3. With respect to April rainfall, the propensity is to understate variability and also understate location. The estimated July calibration functions suggest a different result -- that nearly half the farmers had subjective distributions that overstated variability. Further, the location tended to be understated reflecting pessimism about the likelihood for favorable July precipitation. Again, these observations are consistent with venture theory.

5. CONCLUSIONS

Interpreting the results of this study is quite clear; what to do in response is not so clear. The value of climate information in crop production has been shown to increase dramatically when decision makers' subjective assessment of probabilities is in need of calibration. A sample of farmers showed their subjective distributions of April and July rainfall were consistent with the adjustments predicted by venture theory, based both on the sign of the payoff and their ambiguity level. Calibration functions indicate the alterations necessary to equate their subjective beliefs with historic probability. This indicates that a mere educational program alone might not be sufficient to make the distributions equal. Rather, it indicates that some exageration of climate forecast probabilities may be necessary to counter expected underweighting and overweighting.

(This research was partially supported by NSF Grant ATM 89-13023)

Table 3. Calibration Functions Summary

Parameter Groups	April Rainfall (No. of farmers)	July Rainfall (No. of farmers)	Interpretation
P >1, Q >1	12	23	Over dispersed, under confident
P <1, Q <1	21	12	Over confident, under dispersed
P >1, Q <1	3	12	Overstated location
P <1, Q >1	15	4	Understated location
Uniform	*	**	Well calibrated

* 5 had pseudo P-values on the likelihood ratio test > .05 and could be considered well calibrated. However, the variance of the estimator is biased toward low values due to the procedures used to estimate subjective PDFs.

** 8 had pseudo P-values on the likelihood ratio test > .05 and could be considered well calibrated.

Notes

1. This section reports on results of decision models whose development is beyond the scope of this paper. For more information on the construction of these models see Mjelde et al. and Mazzocco (1989) and related references therein.

2. Data were also collected on winter snowfall, expected prices, interest rates, and other variables affecting financial success.

3. We take the historic distributions to be an adequate characterization of the objective or true PDF of weather events. The possibility of climate change is thus not addressed.

4. The *Gauss* programming language was used on an IBM compatible microcomputer. Briefly, a nonlinear optimization routine with a quadratic loss function was used to recover parameters. Given the flexibility of the distribution function, the fit was typically very good in terms of squared distance between farmer responses and levels on the estimated CDFs. The procedure has unknown power though due to the prespecified interval breaks at which respondents were polled.

References

Allais, M. 1953. "Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'Ecole Americaine." Econometrica, 21:503-546.

Becker, J.W. and F.O. Brownson. 1964. "What Price Ambiguity? Or the Role of Ambiguity in Decision Making." Journal of Political Economy, 72:62-73.

Byerlee, Derek and Jock R. Anderson. 1982. "Risk, Utility and the Value of Information." Review of Marketing and Agricultural Economics, 50:231-246.

Dubey, S.D. 1966. "Characterization Theorems for Several Distributions and Their Applications." Journal of Industrial Mathematics Society, 16:1-22.

Ellsberg, D. 1961. "Risk, Ambiguity and the Savage Axioms." Quarterly Journal of Economics, 75:643-699.

Einhorn, Hillel J. and Robin M. Hogarth. 1985. "Ambiguity and Uncertainty in Probabilistic Inference." Psychology Review, 92:433-461.

Einhorn, Hillel J. and Robin M. Hogarth. 1987. "Decision Making Under Ambiguity." in Rational Choice, Robin M. Hogarth and Melvin W. Reder, eds. University of Chicago Press, Chicago, IL.

Fackler, P. F. and R. P. King. 1990. "Calibration of Option-Based Probability Assessments in Agricultural Commodity Markets." American Journal of Agricultural Economics, 72:73-83.

Hilton, R.W. 1981. "Determinants of Information Value." Management Science, 27:57-64.

Hogarth, Robin M., and Hillel J. Einhorn. 1990. "Venture Theory: A Model of Decision Weights." Management Science, 36:780-803.

Illinois Dept. of Agriculture. Illinois Agricultural Statistics. Springfield, IL. Various Issues.

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.

Lamb, Peter J. 1981. "Do We Know What We Should Be Trying To Forecast Climatically?" Bulletin of the American Meteorological Society, 62:1000-1001.

Lomax, K.S. 1954. "Business Failure: Another Example of the Analysis of the Failure Data." Journal of the American Statistical Association, 49:847-852.

Mazzocco, Michael A. 1989. "Valuing Climate Forecasts For Midwestern Grain Producers." SWS Contract Report 460, Climate and Meteorology Section, Illinois State Water Survey, Champaign, IL.

Mielke, P.W., Jr. 1973. "Another Family of Distributions for Describing and Analyzing Precipitation Data", Journal of Applied Meteorology, 12:275-280.

Mielke, P.W., Jr., and E.S. Johnson. 1973. "Three Parameter Kappa Distribution: Maximum Likelihood Estimates and Likelihood Ratio Tests." Monthly Weather Review 101:701-707.

Mjelde, James W., S.T. Sonka, B.L. Dixon and P.J. Lamb. 1988. "Valuing Forecast Characteristics in a Dynamic Agricultural Production System." American Journal of Agricultural Economics, 70:674-684.

Sherrick, B.J., S.H. Irwin and D.L. Forster. 1990. "Expected Soybean Futures Price Distributions: Option-Based Assessments", The Review of Futures Markets, 9:386-409.

Tadikamalla, P. R.. 1980. "A Look at the Burr and Related Distributions", International Statistical Rev., 48:337-44.

Wise, John and Pan A. Yotopoulos. 1968. "A Test of the Hypothesis of Economic Rationality in a Less Developed Economy: An Abstract." American Journal of Agricultural Economics, 50:395-397.

Yates, J. Frank and Lisa G. Zukowski. 1976. "Characterization of Ambiguity in Decision Making." Behavioral Science, 21:19-25.

UTILIZATION OF SNOW AND ICE DATA FOR CLIMATE CHANGE DETECTION

B.E. Goodison
Climate Research Branch
Canadian Climate Centre
4905 Dufferin Street
Downsview Ontario
Canada

ABSTRACT

Many of the assessments of future global change are intimately related to climate and hydrology. Monitoring of the climate system through systematic observation and analysis of climate related variables on a global, national and regional scale is one contribution to improving our predictive capability. Identification of a significant change in a climate variable is an integral part of this monitoring. One challenge is to identify what climate elements should be assessed; a second is to determine the significance that may be ascribed to any change in the element being assessed.

One of the highest priority science questions is to determine how global and regional storages and fluxes of water, including snow cover in northern latitudes, will interact with a change in climate. The IPCC noted that cryospheric elements (snow cover, sea ice, lake ice, glaciers and ice sheets, and permafrost) are often sensitive indicators and integrators of basic climate elements, such as precipitation, temperature and solar radiation, and are key variables in the global climate system; accurate information on cryospheric changes is essential for full understanding of this system.

Key elements must be defined which should be routinely monitored on a regional or global basis. Data sets to be used in the analysis must be collated. Ideally both conventional and remotely sensed data will be used. This paper will present the potential use of cryospheric elements, particularly snow cover and lake ice data as possible indicators of change. A challenge to the statistical community on how we might properly assess variability or detect change in these elements naturally evolves.

Current cryospheric data sets are a mix of ground measurements and information derived from satellite remote sensing. For change analysis the frequency of collection, archiving and processing of conventional or satellite data into product information may be daily, weekly, monthly or be limited to a set date. Ground data may be collected at specific sites on set dates every year or be collected at an irregular network of stations on varying dates; stations come and go. Yet a time series of consistent and compatible data must be prepared for assessment of variability and change. The use of spatially averaged information derived from remote sensing further challenges the analysis of change. Techniques have been developed to convert satellite data into geophysical products which are used to monitor spatial and temporal variability. Given that the remote sensing record is relatively short, one will have to produce a compatible data time series which combines information from both ground and satellite sources. Development of a strategy for using such cryospheric information as an indicator of climate variability and change is a fundamental challenge for the climate, cryospheric and statistical communities.

METHODOLOGICAL CHOICES TO CLUSTERING PRECIPITATION DATASERIES AND A CASE STUDY FOR HUNGARY

Judit Bartholy

Department of Meteorology, Eötvös Loránd University, Budapest, Hungary

Introduction

Several diverse considerations have recently focused attention on the classification of meteorological fields, mapping somehow the probability distributions of the climate parameters. Different methods are available for this purpose: rotated empirical orthogonal function analysis (Horel, 1981; Barnston – Livezey, 1986), teleconnection analysis (Wallace – Gutzler, 1981; Esbensen, 1984), cluster analysis (Hartman, 1984; Ambrózy – Bartholy – Gulyás, 1984) and some combined techniques (Petzold, 1983; Bartholy, 1989). In the last decades many experiments were carried out for the quantitative and structural analysis of the precipitation field of different areas using each of the above methods (Stooksbury – Michaels, 1991; Lamb – Richman, 1982; Ehrendorfer, 1987).

In this study we don't concentrate on the evaluation of the advantages or disadvantages of those algorithms – it was already done in several analitical study recently (Richman, 1985; Lund, 1963; Kalkstein, 1987; Livezey, 1989). We focus our interest just on clusteranalysis and its methodology. We made this limitation because in this case we don't want to loose the chance to "watch and understand" directly the results what we will get. This work was carried out mostly to identify geographical regions on which the chosen parameters of the precipitation are near equal or at least "very close" to each other; with other words: we are expecting homogeneous final groups.

Formulating the problems

The questions to be answered

In the recently initiated study we try to answer the following questions of basic importance:

1./ Among the various clustering algorithms which one is the most appropriate to classify the precipitation fields or a regions represented with separate station data? Is it really true that some of the hierarchical clustering algorithms give better results in statistical climatological studies? (Recently this hypothesis has been exposed in several papers but a convincing proof is not get available.) Do there exist at all objective, measurable parameters that allow us to compare the various classifying methods and to decide which one is the "optimal" among them?

2./ How can we determine the optimal number of the final groups used in the clustering process? Using different numbers of clusters what can we say about the similarity, between the resulting cluster systems? If we use the same number of clusters in two different cluster algorithms, how can we relate the two cluster systems, how strong compability can be expected?

3./ Is it justifiable and does it have a real advantage that we treat the precipitation as a continuous field by transforming the data measured at discrete points in space and time using interpolation, weigting, areafactors, ... etc? What is the minimum density of the precipitation stations necessary for the above procedure of reconstruction of a precipitation field to work? What is the upper limit of the areal gradient of the orographic changes, which allows us yet to use the field-representation (of the precipitation) on a subarea?

4./ We would like to separate and solve the two following tasks:

a./ Comparing and clustering the precipitation timeseries of evry single station of the chosen geographical area. (In this case we compare and examine the similarity of the individual stations.)

b./ Clustering the "precipitation fields", which are represented with few hundred stations. (In this case we will get as results the few tipical precipitation amount field, for the whole examined region; or with other words: we will get few spatial precipitation distributions wich occure with high probability.)

5./ We should decide: how many parameters are necessary to represent the climatology of the precipitation? [Just the average(?), the average and the standard deviation(?), the extreme values are important too(?), how to choose the timescale(?), ... etc.]

6./ Somehow we should manage the outliers problem too: We are using robust estimate with α-trimmed mean and variances, so we can make less sensitive our results to the outliers. (With different α values we try to get more information about the "average behavior" of the precipitation fields and about the extreme events on different α levels.)

Beyond the above general questions the present work is especially timely now in Hungary. On the one hand because we have the opportunity at first to look over the data base containing the daily and monthly precipitation observations made at more than 900 stations in Hungary, since puting the data to magnetic types has only recently been finisched. Taking into account the number of available parameters and the length of the timeseries we have given the possibility to carry out new investigations, in deepness and complexity not imagineable earlier. These investigations may stimulate and will give the basis for further works on weather forecast and detection of climate change. On the other hand, the serious economical difficulties of Hungary raise the question whether all the 905 precipitation stations are necessary in order to represent the precipitation field with satisfactory precision.

Of course, we will not be able to answer all the above questions – for some of them an unambiguous answer might even not exist –, but we emphasize the

importance of the precise formulation of problem because it influents the choice of the database, the timescale, the geographic region and determines the methods of investigation too.

This work is on the one hand a methodological study because it's runing through and examining a number of cluster techniques. But on the other hand it is in the same time a case study, in the sence: we choose a specific geographic area (the territory of the Charpathian-basin) with the related database to determine the "optimal", or the few best clustersystems of precipitation for this region.

The possible cluster techniques

We put on plenty questions of which we would like to get appropriate answers in this work, but what could we expect using cluster analysis, what they are good for?

Ball (1971) and later Romesburg (1984) recommend a lot of posible uses of cluster techniques, for us could be important a few of them: prediction based on groups, data exploration, data reduction, hypothesis generating and testing, to search for natural groupings or regions in the data, to simplify the description of large set of data (Jardine-Sibson, 1971).

We didn't want to exclude in advance any of the clustering methods (although of course there exist a few techniques in which cases there are no sence to count the results for precipitation data). Otherwise the results should show the incompetence of an algorithm.

The cluster methods we considered with in this study (after Everit, 1980 and MacQueen, 1967):

A./ Hierarchical techniques: Agglomerative and divisive methods (but mostly agglomerative - nearest neighbour, centriod, Ward's and linkage methods)

B./ Non-hierarchical techniques: Partitional and overlapping methods (but mostly partitional methods - MacQueen K-mean, H-mean and a few more variant of them.)

Acknowledgement

This research was partially supported by the Hungarian Academy of Sciences under Grants OTKA/443, OTKA/T4196 and OTKA/268.

References

Ambrózy, P. - J. Bartholy - O. Gulyás, 1984: A system of seasonal macrocirculation patterns for the Atlantic-European region. Időjárás, Vol. 88. No. 3. pp. 121-133.

Barnston, A. G. - Livezey R.E., 1985: High resolution rotated EOF analysis of northern hemisphere 700 mb heights for predictive purposes. Proc. of 9th Conf. on Prob. and Stat. in Atm. Sci., Virginia Beach, Am. Met. Soc. pp. 290-298.

Bartholy, J., 1989: Determination of seasonal macrosynoptical types using cluster analysis and rotated EOF analysis. Acta Climatologica, Tomus XXI-XXIII. Fasc. 1-4. pp. 23-33.

Ehrendorfer, M. 1987: A regionalization of Austria's precipitation climate using principal component analysis. Journ. of Clim., Vol. 7., pp. 71-89.

Esbensen, S. K., 1984: A comparison of intermonthly and interannual teleconnections in the 700 mb geopotential height field during the northern hemisphere winter. Mon. Wea. Rev. Vol. 112., pp. 2016-2032.

Everit, B., 1980: Cluster analysis. Halsted Press, a division of John Wiley and Sons, New York, p. 133.

Hartman, D. L., 1984: Some implication of the mesoscale circ. in trop. cloud clusters for large-scale dyn. and Clim. J. of the Atm. Sci., Vol. 41. No. 1. pp. 113-121.

Horel, J. D., 1981: A rotated principal component analysis of the interannual variability of the northern hemisphere 500 mb height field. Mon. Weather Rev., 109. pp. 784-812.

Jardine, N. - R. Sibson, 1971: Mathematical taxonomy. John Wiley and Sons Ltd. p. 275.

Kalkstein, L. S. - Guauri Tan - J. A. Skindlov, 1987: An evaluation of three clustering procedures for use in synoptic climatological classification. Journ. Clim. Appl. Met. Vol. 26., pp. 717-730.

Klaus, D., 1978: Spatial distribution and periodicity of mean annual precipitation south of the Sahara. Arch. Meteor. Geoph. Biokl., Serie B, No. 26., pp. 17-26.

Lamb, P. J. - M. S. Richman, 1982: Regionalization of central United States for short-period summer rainfall. Proceedings of the 7th Clim. Diagn. Workshop. Boulder, Colorado pp. 180-188.

Livezey, R. E., 1989: A critical review of two decades of teleconnection studies in the United States and Europe. Proceedings of the Fourth Int. Meeting on Statistical Clim. 27-31. March, 1989. Rotorua, New Zealand.

Lund, I., 1963: Map-pattern classification by statistical methods. J. Appl. Meteor., No. 2., pp. 56-65.

MacQueen, J., 1967: Some methods for classification and analysis of multivariate observations. 5th Berkeley Symposium on Mathematics, Statistics and Probability. Vol. 1. pp. 281-298.

Petzold, D. E., 1983: Weather map pattern classification: Refinements of a correlation method. Proc. of the 2nd Int. Meeting on Stat. Clim. Sep.26-30. 1983. 1983. Lisboa, Portugal, pp. 4.6.1-4.6.2.

Richman, B. M., 1985: Rotation of principal components. (Rewiev Paper), Illinois State Water Survey. Climate and Met. Sec. p. 73.

Romesburg, H. C., 1984: Cluster analysis for researchers. Lifetime Learning Publ., Belmont, California, p. 333.

Stooksbury, D. E. - P. J. Michaels, 1991: Cluster analysis of southeastern US climate stations. Theor. Appl. Climatol., Vol. 44., pp. 143-151.

Wallace, J. M. - D. S. Gutzler, 1981: Teleconnection in the geopotential height field during the northern hemisphere winter. Mon. Weather Rev., Vol. 109. pp. 784-812.

Date Due

BRODART, INC. Cat. No. 23 233 Printed in U.S.A.