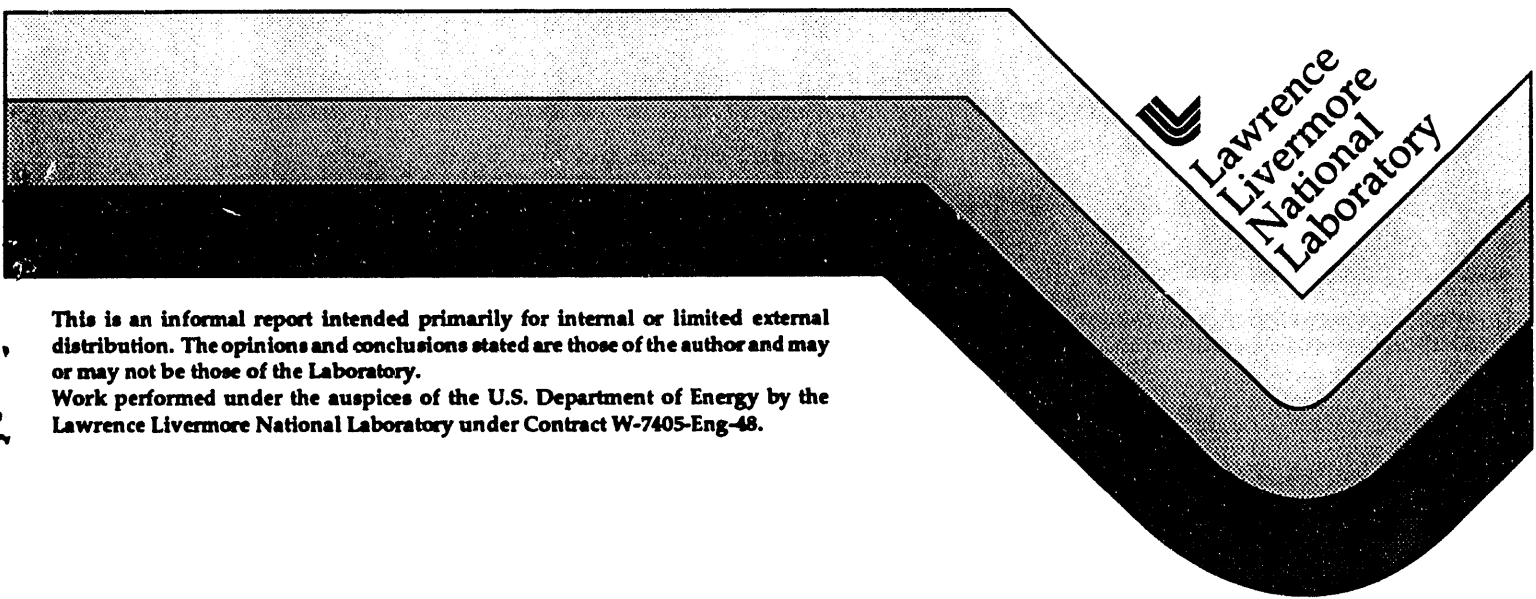UCRL-ID-114222

Statistical Intercomparison of Global Climate Models: A Common Principal Component Approach with Application to GCM Data

Sailes K. Sengupta
James S. Boyle

May 1993

Lawrence Livermore National Laboratory

STATISTICAL INTERCOMPARISON OF GLOBAL CLIMATE MODELS: A COMMON PRINCIPAL COMPONENT APPROACH WITH APPLICATION TO GCM DATA

Sailes K. Sengupta
James S. Boyle

Program for Climate Model Diagnosis and Intercomparison
Lawrence Livermore National Laboratory
Livermore, CA, USA

May 1993

Abstract

Variables describing atmospheric circulation and other climate parameters derived from various GCMs and obtained from observations can be represented on a spatio-temporal grid (lattice) structure. The primary objective of this paper is to explore existing as well as some new statistical methods to analyze such data structures for the purpose of model diagnostics and intercomparison from a statistical perspective. Among the several statistical methods considered here, a new method based on common principal components appears most promising for the purpose of intercomparison of spatio-temporal data structures arising in the task of model/model and model/data intercomparison. A complete strategy for such an intercomparison is outlined. The strategy includes two steps. First, the commonalty of spatial structures in two (or more) fields is captured in the common principal vectors. Second, the corresponding principal components obtained as time series are then compared on the basis of similarities in their temporal evolution.

1. Introduction.

The Atmospheric Model Intercomparison Project (AMIP), of the World Climate Research Program's Working Group on Numerical Experimentation (WGNE) is an ambitious attempt to comprehensively intercompare General Circulation Models (GCMs). The participants in AMIP will simulate the global atmosphere for the decade 1979 to 1988. All the modeling groups have agreed on a common solar constant and CO_2 concentration, and will use a common monthly averaged sea surface temperature (SST) and sea ice data set. This project provides an unprecedented opportunity for realistic and detailed validation and intercomparison of current GCMs. An overview of the AMIP is provided by Gates (1992). If the amount of data generated by a single GCM integration for a ten year period is massive, the output of thirty such integrations is overwhelming. In this work we attempt to present a statistical framework to begin the difficult task of model intercomparison and verification.

To begin we attempt to summarize the aspects of the task of intercomparison:

- We are required to compare a large number of models. (About 30 modeling groups are participating in AMIP)

- The model output in each case is a multivariate vector of geophysical variables (temperature, wind, water vapor, etc.) with a large number of components. (The standard output of AMIP specifies over 15 variables.)

- Each component is defined over a spatial grid and hence is expected to have spatial autocorrelations of varying magnitudes. This issue is further complicated by the fact that the various models have different horizontal and vertical grids and thus may have different underlying correlation structures.

- The gridded data in each case has a temporal evolution based on the underlying physical processes and will in general, have pronounced temporal autocorrelations.

Once these problems are addressed and overcome, at least in some fashion, then an effective intercomparison/validation methodology must first devise a parsimonious representation of the spatio-temporal process(es) described above providing a frame-

work for intercomparison/validation, as well as a large number of exploratory data analysis procedures accompanied by, whenever possible, confirming procedures for the intercomparison/validation.

A significant amount of effort has been expended by climatologists in the past to address this dual task. Data reduction by Karhunen-Loeve expansion (also known as principal components analysis or PCA among statisticians and as empirical orthogonal functions or EOF's by geophysicists) was pioneered by Lorenz (1956). EOF based analyses have a long and rich history of several successful applications. EOF's or principal vectors are based on the covariance structure of a data set. A recent article by Bretherton et al. (1992) provided an intercomparison of several methods of analysis of the covariance structures of meteorological fields in search of coupled patterns. This study, however, does not address the problem in a manner that involves an explicit space-time structure in the analysis.

The importance of building a spatio-temporal framework for detecting coupled patterns was first emphasized by Preisendorfer and Mobley (1982 a–c). They provided a comprehensive theory of data intercomparison by splitting the spatio-temporal 'distance' between two data sets into three parts: 'SITES', describing the separation of the centroids, 'SPRED' representing the difference in radial spread of the data sets and 'SHAPE', describing the difference in the spatial and temporal configurations of the data sets. Preisendorfer and Mobley recognized a serious disadvantage in their approach, especially in the use of the single 'SHAPE' statistic for an intercomparison of the spatial/temporal correlation structure. The proposed correlation statistic compresses all the spatio-temporal information into a single statistic and in the process loses the detailed information regarding the spatial pattern, temporal evolution and distribution of the variance in the data sets. Nevertheless, the use of these statistics and some extensions thereof (Livezy, 1985, Willmott et al., 1985, Zwiers and Thibeaux, 1987 and Zwiers, 1987, Wigley and Santer, 1990) have been suggested for routine use in the quantitative comparison of meteorological fields. Whatever the merits of these space-time statistics are, the need remains for an effective spatial compression of data for the purpose of any quantitative study of the temporal evolution of spatial fields. Accordingly, much of this study is focussed on an effective representation of the spatial fields in a common framework.

Temporal evolution of a global grid can be represented in two different ways:

- As a multiple time series (MTS) $f(x,t)$ where x is a vector representing the physical location on the grid and t is the time.

- A space-time stochastic process (STSP) $f(X)$ with the index $X=(x,t)$, defined over the space-time continuum.

Section 2 will describe the general structure of MTS in a GCM and some problems associated with it in handling the problem of model diagnosis and intercomparison. Section 3 will provide a brief summary of the structure of STSP and its role in GCM output analysis. In an MTS representation the number of components for a global grid or even a regional subgrid is inordinately large. A parsimonious description of such an MTS is essential for meaningful statistical inference procedures. Section 4 will address this issue leading to a Reduced Multiple Time Series (RMTS) based on the use of a few principal vectors. In discussing PCA for continuous domains, Preisendorfer and Mobley (1988) introduce the concept of rotation of the empirical eigenvectors to bring them to dynamically meaningful configurations and use the 'degree' of rotation needed as a basis for a diagnostic test for the statistical link between the observation and the model. In this work we take a somewhat different approach based on the idea of Common Principal Components (CPC) introduced by Flury (1984). This will provide a common frame of reference for the purpose of model/data and model/ model intercomparison.

Section 5 will provide the details of the method of Common Principal Vectors (CPV) introduced in section 4. Once the CPV fields are obtained to provide a common frame of reference, one can get the corresponding CPC's as time series by projection onto the respective data and model fields. These in turn can be analyzed by the methods of time series analysis. A crucial step in this analysis is time series model identification and intercomparison. This has been dealt with in section 6. In addition to the standard auto-regressive moving average (ARMA) identification procedure, a new procedure based on pattern recognition (PR) methods is indicated there. For the intercomparison of two time series one would like to see how closely the identified time evolution pattern in one series can be used to predict the other time series data. This task of prediction is normally handled by traditional methods which are linear and

parametric in nature and are good as first approximations only. A nonlinear and non-parametric time series model can be formulated based on Artificial Neural Networks (ANN) as described by Elsner and Tsonis (1992) for example. A brief introduction to the use of ANN for prediction in the context of model validation is also provided in section 6. The next section provides the outline of a strategy for intercomparison of space-time fields. In section 8, we show an intercomparison using the strategy outlined in section 7. Finally, directions of future research are considered in section 9.

2. Structure of multiple time series (MTS) in GCM analysis

An MTS is regarded as a finite part of the realization of a vector stochastic process. Consider an (M x N) grid sampled at time t from a field F with the data arranged as a one-dimensional array Y_t of length K = M x N. The time evolution of Y_t in discrete times is:

$$t \rightarrow Y_t(w) \in R^K$$

where $Y_t(w)$ is a realization of the field of values in the K-dimensional Euclidean space. As is well known, several useful structures can be imposed on an MTS:

- A vector auto-regressive (AR) process of lag p,

$$VAR(p): \qquad Y_t = \sum_{i=1}^{p} A_i Y_{t-1} + a + \varepsilon_t \qquad (1)$$

where a is a K - vector of constants, A is a K x K matrix of coefficients and 'errors' ε_t, $t \in T$ are independent with $E(\varepsilon_t) = 0$ and with a non-singular matrix of covariance. Here T is the index set for time t representing a set of integers.

- A vector auto-regressive moving average (MA) process with the AR parameter p and the MA parameter q, VARMA(p,q).

- VAR(p) with $p \rightarrow \infty$ also denoted by VAR(∞).

In the last two cases an upper bound on p is unknown, a situation not infrequent in climate studies. See Lutkepohl, 1991 for a comprehensive study of MTS models.

Also well known are the methods of estimating a and A using either the least squares or the maximum likelihood techniques. Statistical properties of these estimates have also been investigated. Further, reliable criteria for selection of the orders p, q of the model exist, as given, for example, by Akaike (1974), Hannan and Quinn (1979) and Hannan and Rissanen (1982).

For GCM applications, a global MTS has two distinct disadvantages. First, for the full global grid even the simplest time series model becomes unmanageable. Take, for example, a simple MTS model VAR(1) given by

$$Y_t = A_1 Y_{t-1} + a + \varepsilon_t \tag{2}$$

The number of parameters in this model on a K x K grid over a time span 1 to T is (2 K^2+ K). With K = 64, this becomes $2(64)^2 + 64$ or 8,256. Even if we restrict the grid to a region of moderate size, the number of parameters to be estimated remains excessively large.

Secondly, it should be borne in mind that GCM outputs are obtained on a global grid. Observations in one cell are strongly related to those in the neighboring cells. An MTS based approach does not explicitly take this spatial structure into account. As a result, information related to spatial structure in the data set is obscured, making it difficult to provide a physical interpretation of the results derived from an MTS based analysis (Katz, 1992).

3. Including a Spatial Structure in the Model: A space-time stochastic process approach

Whittle (1954) introduced a 2d spatial analog of an AR time series model. Later, Bartlett (1971) considered a spatio-temporal Markov model with a given type of spectral density function and Cox (1974) emphasized the need to consider a spatio-temporal model on a spatial lattice for studying the instantaneous spatial structure of a field. Whittle's model can be briefly described as follows.

Let Y_{ij} be the observed value of a process at location (i,j) on a spatial grid. Let L_x, L_y denote respectively the 'lag' operator along the horizontal and vertical directions. Thus,

$$L_x \, Y_{i,j} = Y_{i-1,j} \text{ and } L_y \, Y_{i,j} = Y_{i,j-1}.$$

We shall also denote by L_t the time lag operation.

Let $L_k^1 = L_k \, L_k \dots L_k$ denote the result of applying the operator L_k (k=x or y) l times successively. Then the spatial AR model can be defined as

$$F\,(L_x, L_y)\,Y_{i,j} = e_{i,j} \tag{3}$$

where F is a polynomial in the operators L_x and L_y:

$$F\,(L_x, L_y) = \sum_k \sum_l c_{kl} L_x^k L_y^l \tag{4}$$

where k, l are integers positive or negative and ε_{ij} 's for different pairs (i,j) are independently distributed with zero mean and a common covariance. In a first order spatial AR model k+l =1, and the spatial dependence is limited to a lag of 1.

Since the exponents k and l in the expression (4) can be both positive and negative, the dependence on neighboring cells is bidirectional, and standard time series methods of analyses based on unidirectional dependence are not directly extendable. Tjøtsheim (1983) has introduced the notion of unilateral lattice process models of the causal type (dependent on a quadrant or a half-space) with interesting applications in waveform recognition (Tjøtsheim, 1978) and image processing (Tjøtsheim, 1981). An important alternative to this type of models is provided by Besag (1974). He used a first order conditional model with a (spatial) transition probability structure:

$$P(Y_{i,j} \,|\text{rest of the sample}) = P(Y_{i,j} \mid Y_{i-1,\,j},\, Y_{i,\,j-1})$$

depending only on the immediate neighbors. Oshumi (1988) provided the most general spatial random field model as a basis of some meteorological applications. He considers a non homogeneous field with second order increments. The method in principle can be extended to spatio-temporal random fields. However at such a level of generality our basic problem of model intercomparison becomes hopelessly intractable. Simplifications are possible by introducing the notion of separable processes that was first introduced by Quenouille (1952) and later developed by Martin (1990) and Basawa et al. (1990).

Separable Processes

The three dimensional parameter estimation problem can be simplified to the problem of estimating three one dimensional parameters by introducing the assumption of separability. We define the notion of separability as follows.

Let $\theta = (x,y,t)$, $x \in X$, $y \in Y$ and $t \in T$ denote a 3-d index where X, Y, and T are finite one-dimensional lattices.

DEF. The stochastic process X is "Weakly stationary" if $E(X^2_\theta) < \infty$ and both $E(X_\theta X_{\theta +H})$ and $E(X_\theta)$ are independent of θ, for all (vector) lag $H = (x,y,t)$ in the product space X x Y x T.

DEF. The "Covariance" Γ at lag H is given by

$$\Gamma(H) = Cov(X_{\theta +H}, X_\theta)$$

and the "Correlation" τ at lag H as

$$\tau(H) = \Gamma(H) / \Gamma(O)$$

where, $O = (0,0,0)$ refers to the origin in the index space.

DEF. The stationary process X_θ is "separable" if $\tau(H)$ is factorizable:

$$\tau(H) = \tau_1(x) \, \tau_2(y) \, \tau_3(t)$$

where τ_1, τ_2, τ_3 are the lag correlations in the factor spaces X, Y and T respectively.

Before introducing the factor representation of a weakly stationary 3-d separable process let us recall that the ARMA(p,q) representation of a 1-d process (using the time lag operator L_t) is given by:

$$F(L_t) \, f_t = \theta(L_t) \, Z_t$$

where Z_t is a White Noise (WN) process. F and θ are polynomials of degrees p and q respectively, representing the AR and the MA components of the model. A similar

representation in a factorized form is possible for a 3-d weakly stationary separable process as:

$$F_x(L_x)F_y(L_y)F_t(L_t)f(x,y,t) = \theta_x(L_x)\theta_y(L_y)\theta_t(L_t)Z(x,y,t) \qquad (5)$$

where F_x, F_y, F_t, θ_x, θ_y, θ_t are polynomials of specified degrees and $Z(x,y,t)$ is a white noise process of three dimensions. The separability assumption allows us to represent the process by (5) and proceed with the maximum likelihood estimation of the param-eters involved (Basawa et al., 1990).

An application of a simplified version of the space-time ARMA model has been considered by Niu and Stein (1990) for analyzing the monthly averages of global ozone data observed during 1979–88. The model is simplified to include only one spatial dimension. The processed satellite ozone data is provided in the form of a global latitude, longitude lattice system. For such data, observations in a given latitude band at different longitudes are highly interrelated. For observations in different latitude bands this inter relatedness is much less pronounced, suggesting that the ozone data should be modeled for each latitude band separately as bidirectional autoregressive moving average (BARMA) processes with perhaps a trend and a seasonal component in each. Thus for a fixed latitude band, they modeled the observed ozone concentration $f(t)$ at longitude j and time t as

$$f(t) = TREND(j,t) + SEASON(j,t) + BARMA(j,t)$$

where:

$$BARMA(j,t) =$$

$$\sum_{k=-p1}^{p1} \alpha_k BARMA\,(j+k,t)\ +$$

$$\sum_{k=1}^{p2} \beta_k BARMA\,(j,t-k)\ +$$

$$\sum_{k=0}^{q} \delta_k \varepsilon_{j-k}(t)$$

Here $\varepsilon_j(t) \sim NID(0, \sigma^2(t))$, with $\sigma^2(t)$ allowed to be time-dependent (seasonal). The authors then successfully fit a BARMA(1,1) model to their data. Kim and North (1992) present another interesting application of a space-time ARMA model to climatic data. It should be noted that there is considerable flexibility in choosing the nature of the trend and the seasonal component as well as in the choice of the orders p1, p2 and q. Fitting an appropriate BARMA model to the model/observation data makes the tasks of model validation and change detection simpler.

4. Reduced multiple time series

In section 2 we pointed out that a multiple time series based on a full global grid is difficult to analyze due to its computational complexity. In addition, a problem inherent in intercomparison of models is the problem of multiplicity (Tukey 1977, Hasselman 1979). Generally speaking, the problem of multiplicity arises when statistical significance tests are performed simultaneously. With each individual test having a nominal significance level, the probability of error of the first kind increases geometrically with the number of tests performed which, in this context, is the number of grid points investigated. The problem of multiplicity has been handled in various manners (Tukey 1977, Zwiers 1987, Katz 1992). No matter how the complexity is handled some form of data reduction is essential. In search of parsimony one may proceed along three possible directions.

Summary statistics
Temporal evolution of selected spatial features derived from the gridded data (e.g., first and second order spatial statistics) can be studied for change detection or model validation.

Using selected grid points
Representative grid points can be selected based on some statistical feature selection procedure, analogous to those used in pattern recognition systems.

Feature extraction procedure based on principal vectors
One may select a few significant principal vectors and study their temporal evolution.

In the following, we briefly discuss each of these procedures.

a.   Summary Statistics

Various summary statistics can be computed for a given field and their time evolution can be compared for the model(s) and observations. Various time series characteristics of these summary statistics for the fields under comparison can be estimated and compared. We cite below a few of the spatial statistics that have been found useful in the field of pattern recognition, especially in remote sensing applications. (See for example Haralick et al. (1979), Sengupta et al. (1989) and Rabindra et al. (1992)). These statistics are based on the notion of 'texture' in the image processing (IP) literature, which in turn depends on the joint distribution of two grid point values separated by a fixed vector distance. Such a joint distribution has been called a co-occurrence matrix in the IP literature. The (i,j)th entry in this matrix represents the relative frequency of occurrence of the value pair (i,j) occurring in the field at any pair of grid points separated by a fixed vector distance. Among the second order statistics derived from the co-occurrence matrix that have been found useful in the IP literature are: moments, run length, entropy, angular second moment, correlation, local homogeneity and contrast (Welch et al., 1988, 1989, 1991). The characteristic spatial properties of a field are summarized in these statistics and can be used as the components in an MTS for the purpose of model intercomparison/validation.

b.   Selected Grid Points

Let X be a p-variate random vector representing the model output or observations at p grid points. Let Y be a linear transform of X with $q(\leq p)$ components so that

$$Y = A^t X \tag{6}$$

where A is a p x q matrix with $q \leq p$ and $A^t A = I_q$, the identity matrix of order q. Further, let Σ_x, Σ_y denote the covariance matrices for the random vectors X and Y respectively, so that,

$$\Sigma_y = A^t \Sigma_x A$$

Let the nonnegative characteristic roots of Σ_x be arranged in decreasing order of magnitude be

$$\mu_1 \geq \mu_2 \geq \mu_3 \ldots \geq \mu_q \geq 0$$

McCabe (1984) suggests the following further restriction on the structure of the matrix A. The q columns of A are obtained by an ordered sampling without replacement from the columns of I_p. The effect of applying the transformation (6) then is tantamount to the selection of a subset of size q of the variables. The selection process can be based on a suitable optimality criterion. Although these criteria can start from very different objectives, most of them are equivalent to one of the following:

Maximize the total variance criterion: $\qquad \sum_{i=1}^{q} \mu_i \qquad\qquad (7)$

Maximize the generalized variance criterion: $\qquad \prod_{i=1}^{q} \mu_i \qquad\qquad (8)$

McCabe has termed these q variables 'principal variable', a term akin to 'principal components' in multivariate statistics. If q is sufficiently small and yet adequate for a parsimonious description of the data then a multiple comparison based on Bonferoni inequality (Katz 1992) can be employed effectively.

c.   Principal vectors / components

If A is subject only to the requirement of column orthogonality (6) then the result is known as the empirical orthogonal functions (EOF) or what is the same as principal vectors in classical multivariate statistics. As indicated earlier, a substantial volume of literature exists on the uses (and sometimes abuses) of the method of EOF in atmospheric/meteorological/oceanographic studies. A strategy for the model(s)/observation intercomparison will now be outlined. First note that we need to consider three types of comparison (Priesendorfer and Mobley, 1982):

• Model/model intercomparison where outputs from two or more GCMs are compared;

- Model/observation comparison where a GCM output is compared with observation;

- Model/analysis comparison where a GCM output is compared with analysis based on an incomplete set of observations (ECMWF analysis for example).

d. Comparison strategy

- Compute the first few 'significant' principal components for the data fields under consideration and compare the resulting sets of time series in the frequency as well as time domain.

- Contour the principal vectors corresponding to the 'significant' eigenvalues and compare the structure of these fields.

There are various strategies for retaining the 'significant' principal vectors in an eigenstructure analysis. However, in comparing two or more fields regarding their eigenstructure there may not exist any natural ordering of the principal vectors where one can legitimately compare the k th principal vector of one field with the same for the other fields under comparison for all k. This question of correspondence of principal vectors has apparently been ignored in the literature. We shall address this problem in a different manner in the next section by using a new tool called the Common Principal Components first introduced by Flury (1984).

5. Common principal components: A tool for studying common covariance Structure

Let there be k fields under comparison with p components each. Let S_i, $i = 1, 2, .$ k, be their respective covariances. One of the most important questions in the intercomparison of these fields is: Are the covariance matrices 'similar' in any meaningful way? Flury (1988) has provided the following levels of hierarchy of similarity of covariance matrices:

$$(\Sigma_i), i = 1, 2, 3, ...k.$$

I. Σ_i's are all equal.

II. Σ_i's are proportional, that is $\Sigma_i = c_i \Sigma$ for some constants c_i and some fixed covariance matrix Σ.

III. The common principal component model.

One of the main objectives in traditional principal component analysis is to find a coordinate system in which the representation of the p components of a multivariate vector are uncorrelated. In the search for a **common** covariance structure it is then natural to ask if it is possible to find a coordinate system in which the p variables are uncorrelated, not only in one field, but in two (or several) fields simultaneously. This leads to the following inquiry. Does there exist a single orthogonal matrix B of order p such that

$$\Sigma_i = B\, M_i\, B, i=1,2, ... , k ?$$

IV. The partial common principal components model (PCPC)

In this model, the full orthogonality restriction is relaxed so that only a subset consisting of q ($< p$) of the p principal axes are required to be in common to all k fields under comparison. The remaining (p-q) of the principal axes are allowed to be field specific for each of the k fields. The precise mathematical formulation is relegated to appendix A.

We shall concentrate on the levels III(CPC) and IV(PCPC) in the rest of this article.

One major advantage in using the CPC/PCPC model is that one can compare **corresponding** principal components. In addition, one can test the hypotheses of (partial) commonality of the principal axes of representation of two (or several) fields of data. The statistical tests of significance are given in appendix A. It should be noted here that the application of the tests of significance requires that the sample fields (over discrete time instants) be independent. This is, in most cases not a valid assumption since fields over successive time instants are in fact correlated. This

problem itself does not preclude the use of common principal components as a diagnostic tool for understanding the commonality of the fields without a formal test of significance (It should be pointed out that the current use of EOF for intercomparison does not go beyond a visual comparison of the fields). In order to apply the tests of significance however, the following modifications in the procedure are suggested.

First, based on the auto-correlation structure of the time samples one may subsample from them so that the effect of auto-correlation is statistically insignificant on successive samples. Of course this will typically reduce the effective sample size by a factor of 3 to 4 for monthly data.

Second, one may consider the principal component analysis in the 'sample space setting' (Preisendorfer and Mobley, 1988) where each time series on a single grid point is regarded as a sample, providing as many samples as there are grid points. Spatial instead of temporal subsampling should now be used to ensure approximately the independence of the samples should one decide on a significance test.

6. Time series model identification and intercomparison

A crucial step in the model/model (observation) intercomparison as described in section 5 is the correct ARMA identification of the time series obtained from the common principal vectors. One may employ the Box-Jenkins methodology (Box and Jenkins, 1970) for this. However, the identification using this methodology is often not unique. In order to get around this problem a method based on PR principles is (Lee and Park 1988) outlined now. This method for time series identification is based on the technique of supervised learning. N classes of time series are labelled by the N pairs (p_i, q_i), i=1,2,..., N based on the identification of N time series ARMA models under consideration. Simulated ARMA series with these identification characteristics are used to 'train' a traditional classifier (linear, piece-wise linear or quadratic). The training essentially consists of building the coefficients of the discriminant functions which in turn are used to form the class boundaries for classification decisions. Once the classifier is trained, a new time series is classified based on the values of the computed discriminant functions generated by the classifier during the training phase. Traditional classifiers of the types mentioned above are often suboptimal in accuracy because of the parametric assumptions and fixed boundary types inherent in the model. A relatively new type of classifiers based on artificial neural networks (ANN)

overcomes these problems. Neural networks can be used to build classifiers that can create arbitrarily complex class boundaries for decision surfaces for optimal performance. Furthermore, no distributional assumptions are needed in this classifier. In fact, Lee and Park (1991) have recently extended their work by replacing the traditional linear discriminant-based classifier with a classifier based on artificial neural network which has improved their classification accuracy.

In addition to the ability to perform accurate classification, ANN also has the capability of accurate and robust function approximation. This last capability makes ANN a strong competitor of other traditional methods that are based on regression, interpolation or spline functions. In this area, ANN enjoys the advantage of fully utilizing the parallel distributed processing capabilities that are inherent in such networks by construction. Furthermore, unlike the traditional function approximation methods, ANN does not require the specific nature of the approximating function to be provided. This implicit function approximation capability has recently been explored in a paper by Elsner and Tsonis (1992) in the context of a meteorological prediction.

a. Time series Prediction and intercomparison

We have indicated in section 1 that one way of comparing two (or more) time series resulting as the common principal components is to check how well the identification of parameters in one can be used to predict the other. More specifically, one would identify the ARMA model in one series by estimating the orders p, q and the associated (p+q) ARMA coefficients and then use these to predict the other series. In a nonparametric setting one may use an ANN directly for the prediction bypassing the need to identify the model first as indicated in the following paragraph.

b. ANN in time series prediction and intercomparison

The fundamental problem in a one-step prediction can simply be stated as the estimation of a mapping f:

$$X(t + n) = f(X(t), X(t+1), \ldots, X(t+n-1)) \tag{9}$$

where X(T), T = t, t+1, ..., t+n-1 denotes the value of the time series at time T. In the context of the problem of model intercomparison, one can look at two time series X(T), Y(T) of a specified CPC pair resulting from two model outputs (or a model output and

observations) and find a predictive function of the form (9) for the X(T) series. This function is encapsulated in the form of weights of the ANN trained by 'examples' selected as time segments of fixed length (n+1) from the series X(T). These weights are analogous to the regression coefficients in a regression model. The input (X(t), X(t+1),..., X(t+n-1)) is a segment of length n, and the output is X(t+n) both taken from the time series segment (X(t),..., X(t+n)) for t=1, 2,..., N-n, where N is the length of the series. This estimated function can then be used to predict Y(t+n) based on an input segment (Y(t), Y(t+1),..., Y(t+n-1)) for different values of t. The corresponding ANN output is the predicted value of the Y(t) series based on the X(t) series. A widely used measure of skill of a predictor is the correlation coefficient R between actual and predicted values (Anthes, 1984). This or other measures of predictive skill can then be used to validate the similarity of the two models (or the model and the observations). The process can of course be repeated for the comparison of all significant CPC. Work is in progress towards this and other model validation/intercomparison (Sengupta and Boyle, 1992).

7.   A strategy for space-time data field intercomparison

In section 5 we introduced the notion of common principal vectors as a means of representation of data from multiple fields in a common frame of reference. This led to groups of time series of principal components where the time series within each group need to be intercompared. Section 6 dealt with the methods of intercomparison by traditional methods as well as methods based on ANN. In this section we combine the two steps to outline an overall strategy of space-time field intercomparison.

1. For a given meteorological/atmospheric/oceanographic variable over k space-time fields, determine if the mean fields under comparison are very much alike. If they are, one may consider the comparison of anomaly fields. For example, one might consider monthly temperature anomaly patterns over a certain period as given by observation, GCM output, or analysis derived from an incomplete set of observations.

2. Test the hypotheses of q common principal vectors in succession for q=1,2,..., until the hypothesis of commonality of q principal vectors can be rejected for some q(<p), say, q=Q+1. This indicates that the maximal number of common principal vectors is Q.

3. For each CPV, project each of the k space-time fields on the CPV by taking the scalar product of the CPV and the sample field at each time instant to get k univariate time series.

4. Using the methods indicated in section 6, the k univariate time series can now be intercompared with regard to their evolutionary patterns. The methods in section 6 are all based on time domain analysis. However, frequency domain characteristics can also be intercompared.

Thus, in search for a coupled set of patterns one first looks at the similarity in the spatial distribution through the orthogonal common principal vectors and then in the temporal characteristics of the derived principal components. The latter task can be accomplished by traditional model identification procedures or by using pattern recognition methods as indicated in section 6.

In the context of model intercomparison, two fields would be considered to be 'similar' with increasing degrees of similarity in order indicated below, if:

• The significant common principal components explain a 'large' portion of the variations in the fields under comparison. In most practical applications of principal components, when dealing with 'explained' variations, a cumulative proportion of 70–95 percent is considered 'large.'

• The ARMA(p,q) identification of the CPC time series are identical.

• The estimates of the AR and MA coefficients using these parameters in each case when used to predict the other series shows a high degree of predictive skill. (A widely used measure of skill of a predictor is the correlation coefficient R between actual and predicted values (Anthes, 1984)).

8. Application

The data sets used for the example application of the CPC methodology are vertically integrated temperatures from three sources. The first source is a ten year climate simulation of the decade 1979 to 1988, the second is the operational analyses from the European Centre for Medium Range Forecasts (ECMWF), and the third is

the observations from the Microwave Sounding Unit (MSU) carried on the NOAA polar orbiting satellites. The MSU data (specifically channel 2) represent a weighted vertically averaged temperature. These data are described in detail by Spencer and Christy (1990). Hurrel and Trenberth (1993) describe a method for computing the equivalent of the MSU temperature given a vertical array of temperature values. Their method was applied to the model output and the operational analyses to yield three data sets for comparison.

The model data which were used for this experiment is from the AMIP integration of the ECMWF model, cycle 36. The model has 19 levels in the vertical and for this experiment was set to a horizontal resolution of T42. The model is in almost all respects the same as that described by Miller et al. (1992). The sea surface temperatures are specified in accordance with the AMIP guidelines. The surface land temperatures are allowed to vary in accordance to the surface parameterizations employed. The integration started with the ECMWF analyses for 1 January 1979 and continued for 10 years.

The operational analyses are those of the ECMWF. A rather complete discussion of the nature of these data for the period 1979 to 1988 is given by Hoskins et al. (1989).

The data for the comparison are all available on global grids. It was decided to perform the analysis on a limited grid covering 15 S–15 N and 120 E–255 E of the equatorial Pacific region. The spacing between the data points is 15 degrees of latitude and longitude. This region was chosen since it is the area within which the major atmospheric and oceanic anomalies occur during ENSO events. In this area the SST have strong variations that affect the atmospheric temperature. Since the only connection of the model to actual conditions is the SST patterns, it was felt that this region was a good testbed to demonstrate the accuracy of the simulation. The data were available on latitude longitude grids of approximately 2.5x2.5 degrees, although only data every 15 degrees was used in order to reduce the amount of spatial correlation between data points.

Although the data were available as monthly means, the ECMWF analyses were available only from 1980 to 1988; the computations were therefore performed for 9 years yielding 108 time samples. As shown by Newell and Wu (1992) the data possess a strong temporal correlation. In their paper they provide estimates of the effective sample size using the observed MSU temperatures for the same time period. They used these estimates to determine significance levels for correlation coefficients.

Another reason for restricting the region of computations to the equatorial Pacific is that the time correlation is relatively homogeneous in this region, which allows for a less ambiguous estimate of the effective sample size than a region which comprises a spectrum of time relations should one choose to construct a statistical test based on estimates.

Finally, anomalies were computed for all the data sets by subtracting the 9 year mean for each month from each month's data. This procedure effectively removes the seasonal cycle which would dominate the more interesting non_seasonal signals. As a result of these decisions the analysis was performed using a 3 x 10 spatial grid with 108 time samples. This small amount of data allowed the analyses to be run quite easily on a Sun 1+ workstation. The IMSL (1991) subroutine KPRIN based on the FG algorithm for common eigenstructure (Flury and Gautschi, 1986) was used in this work.

Results

The first two common principal vectors for each of the three pairs are shown in Figs. 1a–f for each case and for each field in the pair. For purposes of comparison, Figs. 2 a–f display the first two principal vectors of the individual data sets. Also plotted, in
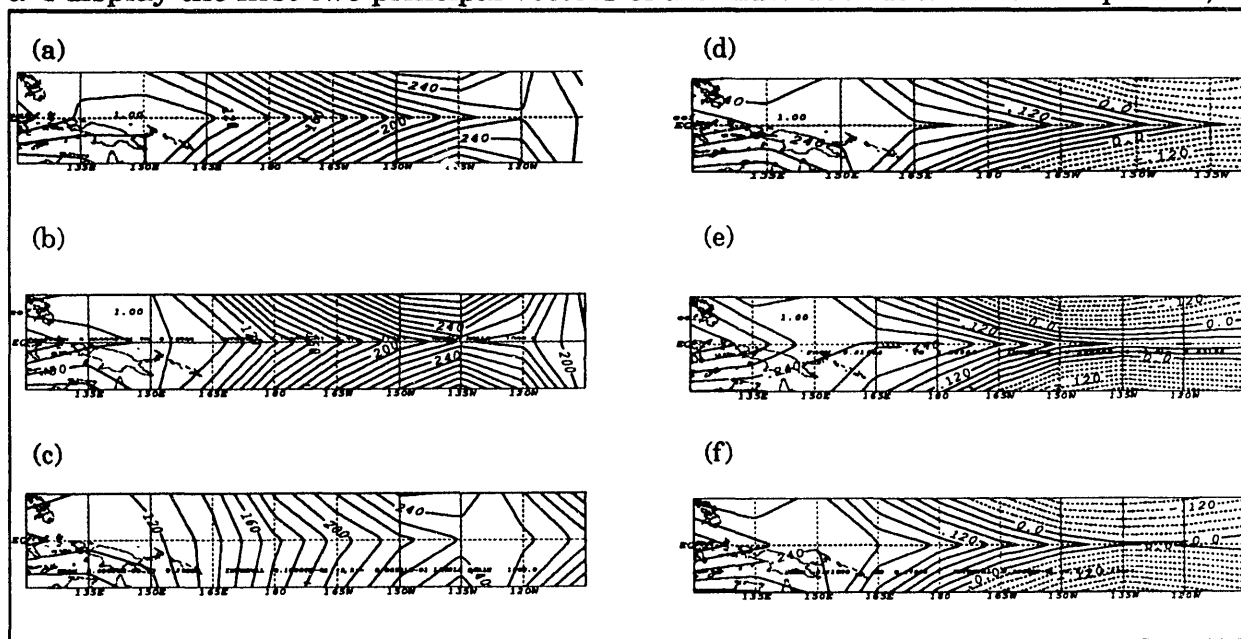

Fig. 1. (a) The first common principal component vector using the model and observation data sets of vertically integrated temperature. (b) As in (a) except for the analyses and observation data sets. (c) As in (a) except for the model and analyses data sets. (d) As in (a) except for the second CPC. (e) As in (b) except for the second CPC. (f) As in (c) except for the second CPC.
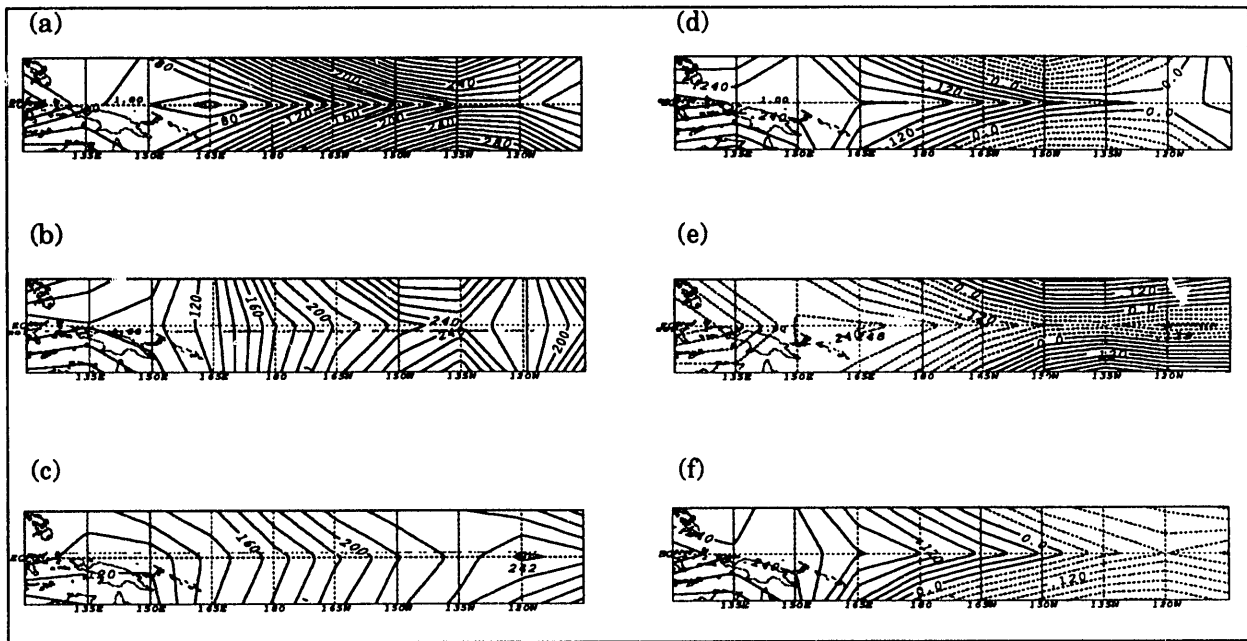
Fig. 2. (a) The first principal component vector for the observations. (b) As in (a) except for the model data set. (c) As in (b) except for the analyses data set. (d) As in (a) except for the second principal component. (e) As in (b) except for the second principal component. (f) As in (c) except for the second principal component.
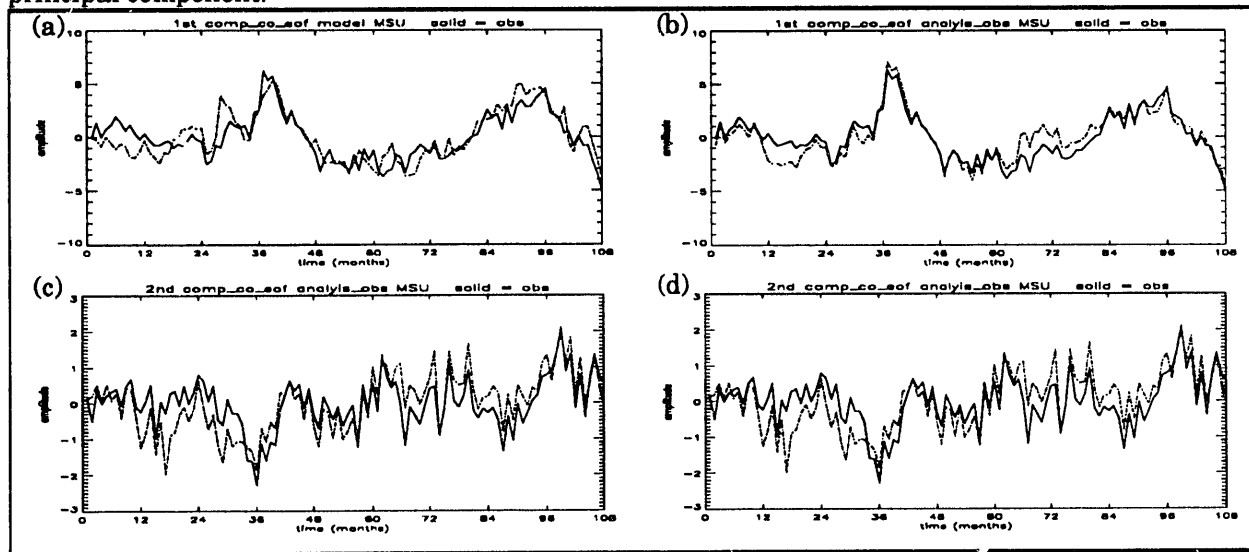

Fig. 3. (a) Time series of the projection of the model and observed data onto the first CPC of the model and observations. The solid line is the observed data. (b) Time series of the projection of the analysis and observed data onto the first CPC of the analysis and observations. The solid line is the observed data. (c) Time series of the projection of the model and observed data onto second CPC of the model and observations. The solid line is the observed data. (d) Time series of the projection of the analysis and observed data onto the second CPC of the analysis and observations. The solid line is the observed data.

-20-

Table 1: CPC 1 in ANAL/OBS Intercomparison EQPAC region

	p	q	AR coeffs	MA coeffs	Residual Variance
ANAL.	1	24	a1=-0.825	b1=-0.2970 b24=-0.278	1.300
Observation	1	24	a1=-0.847	b12=-0.258 b24=-0.308	1.119

Table 2: CPC 2 in AMIP/OBS Intercomparison EQPAC region

	p	q	AR coeffs	MA coeffs	Residual Variance
AMIP	1	0	a1=-0.721 a2=0.103		0.321
Observation	2	0	a1=-0.846 a2=-0.324		0.323

Table 3: CPC 2 in ANAL/OBS Intercomparison EQPAC region

	p	q	AR coeffs	MA coeffs	Residual Variance
ANAL.	24	1	a1=-0.919 a18=-0.204 a24=0.230	-0.696	0.337
Observation	2	0	a1=-0.362 a2=-0.321	b12=-0.258 b24=-0.308	0.324

Table 4: CPC 1 in AMIP/OBS Intercomparison EQPAC region

	p	q	AR coeffs	MA coeffs	Residual Variance
AMIP	11	3	a1=-0.0948 a2=0.103	b1=-0.260	1.088
Observation	1	24	a1=-0.846	b12=-0.246 b24=-0.293	1.148

Figs. 3a–d, are the projections of the data sets representing the time-varying fields along these principal vectors resulting in the principal components of these time-varying fields. These should be interpreted as the principal modes of time evolution of the spatial fields measured along the common principal vectors. In addition the correlograms of the time series of the first two CPC's (Fig. 3) are shown in Figs. 4 a–d. The results of an ARMA analysis for the first two principal components are shown in Tables 1 to 4.

The first common principal vectors for all three cases, Figs. 1a-c, look similar. This implies that the model, analysis and observations are in good agreement with respect to the basic structure of the anomaly fields of integrated temperature in this region. The percentage of explained variance by the first CPC is at least 70% for all the cases. The analyses and observations have percent explained variance by the lowest CPC of roughly 70% while the corresponding model percent is 80%. This might imply that the model patterns are less complex. The pattern seen in Figs. 1a–c is also typical for the first EOF in the SST fields for the region (Weare et al., 1976, 1981, and 1983), Hsiung and Newell, 1983).

The PCA fields, Figs. 2a–c, for the first mode show that the observations evince a much larger gradients although the pattern in all three fields is similar. The analysis appears to be an intermediate version of the observations and the model. This is perhaps not surprising in that the ECMWF model was used as the data assimilation platform for the analysis. This results in the analysis having characteristics of the model and observations. The patterns in Figs. 1a–c can also be compared to the PCA components in Figs. 2a–c. It is apparent that the CPC algorithm also produces blending of the characteristics of the input fields. Comparing the PCA analysis, Figs. 2a–c,

a definitely indicates that the first mode of the three data sets are different, but they do not provide a common basis for comparision. The CPC analysis puts the (pairwise) fields in a common framework allowing for a direct comparison and more detailed analysis of their time evolution.

In the time series of Figs. 3a–b the peaks occurring during the El Niño years of 1982–83 and 1986–87 clearly standout. It is the El Niño pattern that is captured by the CPC's of Figs. 2a–c. A low frequency component of about 4.5 to 5 years appears to be present in all cases. It cannot be determined from this limited time series whether this is a real phenomenon or an artifact of the short time sample. In general the agreement between the model and the observations is not quite as good as between the analysis and observations but in both cases the ENSO events are clearly captured. The agreement for the second CPC, Figs. 3c–d, is not as good, but the amplitude of this component is reduced from the first (note the scale change on the ordinate between Figs. 3a–b and 3c–d). The correlograms for the model and analysis, Figs. 4a–b,
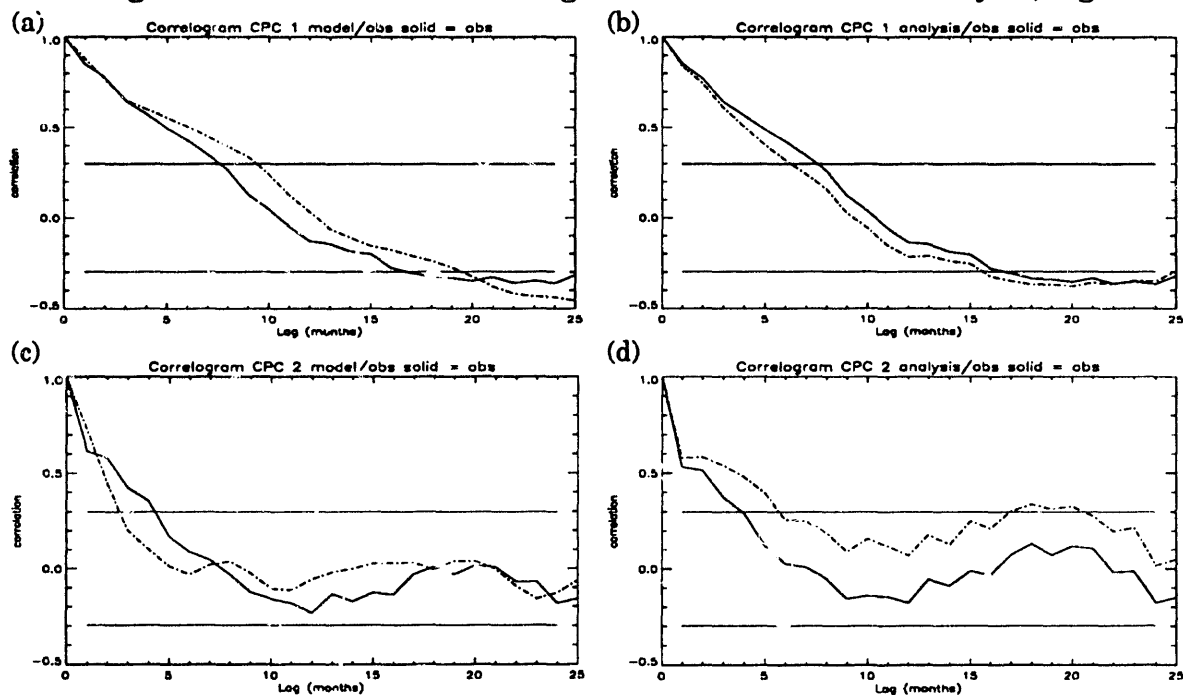

Fig. 4. (a) Correlogram of the time series of the projection of the model and observed data onto the first CPC of the model and observations. The solid line is the observed data. The 95% simultaneous confidence bands for white noise autocorrelations are also displayed on the figure. (b) Correlogram of the time series of the projection of the analysis and observed data onto the first CPC of the analysis and observations. The solid line is the observed data. The 95% simultaneous confidence bands are also displayed on the figure. (c) As in (a) except for the second CPC. (d) As in (b) except for the second CPC.

show opposite tendencies, in that the models have greater autocorrelations, Fig. 4a, and the analysis has smaller autocorrelations. Thus, for their respective first CPC modes, the model has more persistence, and the analyses less than observed. This is consistent with the inference drawn before that the model has less complex structure in the spatial domain. This relationship is reversed for the second CPC mode, Figs. 4c–d, at least for the autocorrelations above the 95% confidence interval. In general, the second modes, Figs. 4c–d, exhibit somewhat less persistence than the first, Figs. 4a–b.

The results of the model identification/intercomparison procedure for the time series corresponding to the first two common principal components of the model/observation and analysis/observation pairs are shown in Tables 1–2. The analysis/observation intercomparison for the first CPC shows the same (p,q) estimates (p=1, q=24). The estimated coefficients also are close in value. This indicates that the first principal components in the analysis and the observations for the region under consideration are nearly identical. However, the ARMA order structures are different for the principal components corresponding to the model/observation pair. In CPC 2, the model/observation pair shows the same MA order (q=0) but slightly different autoregressive order (p=1 for the model compared to p=2 for observation), whereas, the analysis/observation pair show little resemblance in their respective orders. It should be noted however that the ARMA estimates of the orders and coefficients are not unique and hence a more appropriate intercomparison should be based on the predictability of one series by the evolutionary characteristics of the other as indicated in section 6. This approach although not shown here will be used in our future work.

9. Conclusions and Further Research

An appropriate representation of the spatio-temporal data derived from the observations or as model outputs is a necessary first step in model validation/intercomparison. Several possible representations have been considered in this report. Of these, two are particularly attractive for GCM models on a global scale. The first is the bilateral auto-regressive model restricted to a fixed latitude band, with perhaps a time-varying trend and a seasonal component added to it (Niu and Stein, 1990). The second starts with a reduced data set in the form of a MTS consisting of a few orthogonal common principal components for the fields under comparison. The use of CPC's

are not limited to model intercomparisons only. In fact, it is a powerful tool in detecting coupled patterns involving several simultaneous spatio-temporal fields and meteorological variables. This last feature makes it a potentially valuable tool in understanding the physical processes associated with these fields. Some of our future effort will be directed towards this aspect of the applications of CPC's.

Time series model identification methodology applied to the CPC time series has been indicated as the next step in the model intercomparison process. This identification, which is not necessarily unique, need not however be carried out explicitly. We have indicated briefly a method based on ANN that is capable of being used for model intercomparison purposes, bypassing the need for explicit ARMA model identification. Developing this method further will be a part of our future work.

Appendix A

We provide here the mathematical formulation of the statistical hypotheses referred to in the levels III and IV of the hierarchy introduced in section 5. These are respectively the CPC and PCPC hypotheses regarding the covariance structures of the space-time fields. We use the same notations introduced in section 5.

$$\Sigma_i = B\, M_i\, B', \quad i = 1, 2, \ldots k.$$

where B_{pxp} is an orthogonal matrix of order p,

$M_i = \text{diag}(\mu_{i1}, \mu_{i2} \ldots, \mu_{ip})$ is a diagonal matrix for $i = 1, 2, \ldots k$. In other words, the covariance matrices are simultaneously diagonalizable. The common principal vectors then provide us with a common frame of reference for the fields under comparison. The model so derived is thus called the Common Principal Components (CPC) model.

The Partial Common Principal Components Model (PCPC)

In this model, we assume that there are q common principal vectors, $\beta_1, \beta_2, \ldots \beta_q$, for all covariances Σ_i and (p-q) field-specific principal vectors $\beta_j^{(i)}$, $j = q+1, \ldots, p$; $i = 1, \ldots, k$.

$$\Sigma_i = \sum_{j=1}^{q} \mu_{ij} \beta_j \beta'_j + \sum_{j=q+1}^{p} \mu_{ij}^{(i)} \beta_j^{(i)} (\beta_j^{(i)})^t$$
$$i = 1, 2, \ldots k; \quad q < p\text{-}1.$$

Statistical Hypotheses

In III, the log-likelihood statistic for testing the null hypothesis of common principal vectors (H_{CPC}) against the hypothesis of arbitrary Σ_i s is asymptotically chi-square distributed with $(k\text{-}1)p(p\text{-}1)/2$ degrees of freedom and is formally given by

$$\chi^2 = \sum_{1=1}^{k} n_i \frac{\det \hat{\Sigma}_i}{\det S_i}$$

where S_i is the is the ith sample covariance matrix and $\hat{B}$, $\hat{\mu}$, and $\hat{\Sigma}_i$ are respectively the maximum likelihood estimates of B, M_i, and Σ_i under the restrictions of the hypothesis H_{CPC} of common principal components. Note that $\hat{\Sigma}_i = \hat{B}\hat{\mu}_i\hat{B}'$.

Numerical algorithms for estimating B and m_i are available (Flury and Gautschi, 1986) and FORTRAN executable codes can be found in the IMSL statistical package (IMSL, 1991).

The hypothesis corresponding to the level IV (PCPC of order q) of the hierarchy can be written in matrix notation as

$$H_{IV}(q): \Sigma_i = B_i M_i B_i^t, \quad i=1,2,...k$$

where Σ_i's are positive definite symmetric matrices of order p, M_i = diag $(m_{i1}, m_{i2}...,m_{ip})$ and $B_i = (B_c, B_i)$. Each B_i is a (pxp) orthogonal matrix sharing a common part B_c with p rows and q columns independent of the group i and a group-specific part B_i with p rows and (p-q) columns. The test statistic for testing $H_{IV}(q)$ against arbitrary Σ_i can again be based on the likelihood ratio principle and its asymptotic distribution turns out to be chi-square. Clearly there are more parameters estimated in this model under this null hypothesis than under the null hypothesis of level III. This makes for fewer degrees of freedom of the chi-square statistic viz. (k-1)q(2p-q-1)/2. Furthermore, the numerical algorithm to solve the system of equations for the MLE is highly involved. In fact, the proof of convergence of this algorithm is nonexistent. There is, however, an approximate procedure to find the ML estimates of the parameters involved which in turn can be used in a conservative manner to test the null hypothesis. It is conservative in the sense that the approximate test can be used only for the purpose of accepting the null hypothesis when the computed chi-square is less than the corresponding critical value. One cannot however reject the null hypothesis in the opposite case. We have not used either of these versions in this work. Instead we have used the full CPC version of common principal components extraction in IMSL(1991)

subroutine KPRIN in which the columns of B (common principal vectors) are arranged in increasing order of the sums of the corresponding eigenvalues over the k groups under consideration. Depending on the relative magnitudes of these sums, only a few of these common principal vectors are retained for the comparison study. In our case we retained only those principal vectors where contribution from each group in a sum exceeded 5% of the total of the eigenvalues for that group.

Appendix B
List of Symbols

Roman

a	K-vector
A	Rectangular column-orthogonal matrix
A_i	Kx K matrix
B	orthogonal matrix
H	vector lag: (x,y,t)
K	number of grid points
M	number of horizontal grid points
N	number of vertical grid points
p	auto-regressive order in an MTS
q	moving average order in an MTS
R^K	K-dimensional euclidean space
t	time instant
T	finite index set of a time series
w	a single realization of a time series
x	a vector representing the grid location
X	(x,t)
Y_t	discreet time MTS
Z_t	white noise

Greek

α_k	spatial AR parameters
β_k	temporal AR parameters
δ_k	temporal MA parameters
σ	error standard deviation
ε_t	error vector at time t in a MTS
θ	(x,y,t), the three dimensional index parameter in a STSP
Γ	covariance matrix
τ	correlation matrix

Operators

L_x	Lag operator along x - direction
L_y	Lag operator along y - direction
L_t	Lag operator along time axis
Φ	Polynomial
Σ	Covariance matrix
μ_i	eigenvalue

Acronyms

AMIP	Atmospheric model intercomparison project
ANN	Artificial neural network
AR	Auto-regressive
ARMA	Auto-regressive moving average
BARMA	Bi-directional auto-regressive moving average
CPC	Common principal components
CPV	Common principal vector
ECMWF	European center for medium range weather forecasts
ENSO	El Niño southern oscillation
EOF	Empirical orthogonal function
GCM	General circulation model
IP	Image processing
MA	Moving average
MSU	Microwave sounding Unit
MTS	Multiple time series
NID	Normal and independently distributed
PCA	Principal components analysis
PCPC	Partial common principal components
PR	Pattern recognition
RMTS	Reduced multiple time series
SST	Sea surface temperature
STSP	Space - time stochastic process
TSM	Time series model
VAR	Vector auto-regressive

VARMA	Vector auto-regressive moving average
WGNE	Working group on numerical experimentation
WN	White noise

REFERENCES

Akaike, H., 1974: A new Look at the Statistical Model Identification. *IEEE trans. Auto. Contrl.*, **AC-19**, 716–723.

Anthes, R.A., 1984: Predictability of Mesoscale Meteorological Phenomena, Predictability of Fluid Motions, G. Holloway and B. J. West Eds., *Amer. Inst. Phys.*, 247–270.

Bartlett, M.S., 1971: Physical Nearest Neighbor Models and Non-linear Time Series. *J. Appl. Prob.*, **8**, 222–232.

Basawa, I.V., P.J. Brockwell, and V. Mandrekar, 1990: Statistical Analysis of Separable ARMA Processes on a Lattice. *Tech. Reprt.*, The University of Georgia.

Besag, J., 1974: Spatial analysis and the statistical analysis of lattice systems. *J. Statis. Soc.*, **B 36**, 192–236.

Box, G. E. P., and G. M. Jenkins, 1970: *Time series analysis, forecasting and control.* Holden Day, San Francisco, CA, 324 pp.

Bretherton, C.S., Catherine Smith, and John Wallace, 1992: An Intercomparison of Methods for Finding Coupled Patterns in Climate Data. *J. Clim.*, **5**, 541–560.

Cox, D.R., 1974: Discussion of Mr. Besag's Paper. *J. Roy. Statist. Soc.*, *B36*, No. 2, 226–236.

Elsner, J.B. and A.A. Tsonis, 1992: Nonlinear Prediction, Chaos and Noise. *Bull. Amer. Meteor. Soc.*, **73**, 49–60.

Flury, B., 1984: Common Principal Components in k Groups. *J. Amer. Statist. Assoc.*, **79**, 892–898.

Flury, B. and Gautschi, W., 1986: An Algorithm for Simultaneous Orthogonal Transformation of Several Positive Definite Symmetric Matrices to Nearly Diagonal Form. *Siam. J. Sci. and Statist. Comp.*, **7**, 169–184.

Flury, B., 1988: *Common Principal Components and Related Multivariate Models.* John Wiley, New York, 258 pp.

Gates, W. L., 1992: AMIP: The atmospheric model intercomparison project. *Bull. Amer. Meteor. Soc.*, **73**, 1962–1992.

Hannan, E.J. and Quinn, B.G., 1979: The Determination of the Order of an Autoregression. *J. Roy. Stat. Soc.*, **B41**, 190–195.

Hannan, E.J. and Rissanen, J., 1982: Recursive Estimation of Mixed Autoregressive Moving Average Order. *Biometrika*, **69**, 81–94.

Haralick, R.M.,K.S. Shammugan, and I. Dinstein, 1979: Textural Features for Image Classification. *IEEE Trans. Syst. Man. Cybern*, **SMC-3**, 610–621.

Hasselman, K. 1979: On the signal to noise problem in atmospheric response studies. In, *Meteorology over the Tropical Oceans*, D. B. Shaw (ed.), *Royal Meteorological Society*, London, 251–259.

Hoskins, B. J., H. H. Hsu, I. N. James, M. Masutani, P. D. Sardeshmukh and G. H. White, 1989: *Diagnostics of the global atmospheric circulation–Based on ECMWF analyses 1979–1989*. WRCP–27, WMO/TD–No. 326, Geneva, Switzerland, 217 pp.

Hsiung, J., and R. E, Newell, 1983: The principal non-seasonal modes of variation of global sea surface temperature. *J. Phys. Oceanogr.*, **13**, 1957–1969.

Hurrell, J. W. and K. E. Trenberth, 1992: An evaluation of monthly mean MSU and ECMWF global atmospheric temperatures for monitoring climate. (submitted to *J. Climate*).

IMSL Stat/Library, 1991: IMSL Inc., Houston, 1578 pp.

Katz, R. W., 1992: Role of statistics in the validation of general circulation models. *Climate Research*, **2**, 35–45.

Kim, K. Y. and G. R. North, 1992: Seasonal cycle and Second Moment Statistics of Simple Coupled climate systems. (submitted to *J. Geophys. Res*).

Lee, K.-C, and S.-J. Park, 1988: Decision support in time series modeling by pattern recognition. *Decision Support Systems*, **4**, 159–207.

Lee, K.-C., J.-S. Yang and S.-J. Park, 1991: Neural Network based time series modeling: ARMA model identification via ESACF approach. *Proc. International Joint Conference on Neural Networks*, Singapore.

Livezy, R. E., 1985: Statistical Analysis of General Circulation Model Climate Simulation: Sensitivity and Prediction Experiments, *J. Atmos. Sci.*, **42**, 1139–1149.

Lutkepohl, H., 1991: *Introduction to multiple time series analysis*. Springer Verlag, New York, 545 pp.

Lorenz, E.N.,1956: Empirical Orthogonal Functions and Statistical Weather Prediction. *Sci. Rept. No. 1*, MIT, Dept. of Meteorology, Cambridge, Mass., 49 pp.

Martin, R.J., 1990: The Use of Time Series Models and Methods in the Analysis of Agricultural Field Trials, *Commu. Statist. Theory Meth.*, **19(1)**, 55–81.

McCabe,G.P., 1984: Principal Variables, *Technometrics C*, **26**, 137–144.

Miller, M. J., A. C. M. Beljaars and T. N. Palmer, 1992: The sensitivity of the ECMWF model to the parameterization of evaporation from the Tropical oceans. *J. Climate*, **5**, 418–434.

Newell, R. E., and Z-X. Wu 1992: The interrelationship between temperature changes in free atmosphere and sea surface temperature changes. *J. Geophys. Res.*, **97(D4)**, 3693–3709.

Niu, X. and M. Stein, 1990: Space-Time ARMA Models for Satellite Ozone Data, Interface '90, *Proceedings of the 22nd. Symposium on the Interface*, May 16–19, 225–234.

Oshumi, A., 1988: Linear Estimation of Random Fields with Second Order Increments and its Application. *Automatica*, **24**, 203–215.

Preisendorfer, R. W., 1988: *Principal Component Analysis in Meteorology and Oceanography*, C.D. Mobley (ed.), Elsevier, New York, 425 pp.

Preisendorfer, R.W. and C.D. Mobley, 1982 a: *Data Intercomparison Theory I*, NOAA Tech. Memo **ERL PMEL-38**, 48 pp.

_____, 1982 b: *Data Intercomparison Theory II*, NOAA Tech. Memo **ERL-PMEL-39**, 94 pp.

_____, 1982 c: *Data Intercomparison Theory III*, NOAA Tech. Memo **ERL-PMEL-39**, 123 pp.

Quenouille, M. H., 1952: *Associated Measurements*. Butterworths, 241 pp.

Rabindra, P., S.K. Sengupta, and R.M. Welch, 1992: An Interactive Hybrid Expert System for Polar Cloud and Surface Classification, *Environmetrics*, 3(2) pp.

Sengupta, S.K., R.M. Welch and M. Navar, 1989: Performance of Two Texture Based Classifiers of Cloud Fields Using Spatially Averaged LANDSAT Data. *Proc. IGARSS-89/12th Canadian Symposium on Remote Sensing*, Vancouver, B.C.

Sengupta, S.K. and J. S. Boyle, 1992: Artificial Neural Network in Climate Model Diagnostics and Intercomparison Methodology, *Artif. Intell. Res. Envir. Sci. Workshop*, Monterey, Calif., Oct. 27–31.

Spencer, R. W., and J. R. Christy, 1990: Precise monitoring of global temperature trends from satellites., *Science*, **247**, 1558–1562.

Spencer, R. W., J. R. Christy, and N. C. Grody, 1990: Global atmospheric temperature monitoring with satellite microwave measurements: Method and results 1979–84. *J. Climate*, **3**, 1111–1129.

Tjøstheim, D., 1978: Statistical Spatial Series Modelling, *Adv. Appl. Prob.*, **10**, 130–154.

Tjøstheim, D., 1981: Autoregressive modelling and spectral analysis of array data in the plane. *IEEE Trans. Geosci. Remo. Sens.*, **GE-19**, 15–23.

Tjøstheim, D., 1983: Statistical spatial series modelling II: Some further results on unilateral lattice processes., *Adv. Appl. Prob.*, **15**, 562–584.

Tukey, J. W. 1977; Some thoughts on clinical trials, especially problems of multiplicity, *Science*, **198**, 679–684.

Weare, B. C., A. R. Navato, and R. E. Newell, 1976: Empirical orthogonal analysis of Pacific Sea surface temperature. *J. Phys. Oceanogr.*, **6**, 671–678.

_____, 1981: Autoregressive Modelling and Spectral Analysis of Array Data in the Plane. *IEEE Trans. Rem. Sens.*, **GE-19**, 15–23.

_____, 1983: Statistical Spatial Series Modelling II: Some Further Results on Unilateral Lattice Processes. *Adv. Appl. Prob.*, **15**, 562–584.

Welch, R.M., S.K. Sengupta, and D.W. Chen, 1988: Cloud Field Classification Based upon High Spatial Resolution Textural Features, Part I: Grey Level Co-Occurrence Matrix Approach. *Jour. Geophys. Res.*, **93**, **No. D12**, 12663–12681.

Welch, R. M., M. S. Navar, and S. K. Sengupta, 1989: The Effect of Spatial Resolution upon Texture-Based Classification. *Jour. Geophys. Res.*, **94, No. D12**, 14767–14781.

Welch, R.M., S.K. Sengupta, A.K. Goroch, P. Rabindra, N. Rangaraj and M.S. Navar, 1991: Polar Cloud and Surface Classification Using AVHRR Imagery: An Intercomparison of Methods. *Jour. Appl. Meteor.*, **31, No. 5**, 405–420.

Whittle, P., 1954: On Stationary Processes in the Plane. *Biometrika*, **41**, 434–449.

Wigley, T.M. and B.D. Santer, 1990: Statistical Comparison of Spatial Fields in Model Validation, Perturbation and Predictability Experiments. *J. Geophys. Res.*, **95**, 851–865.

Willmott, C.J., S.G. Ackleson, R.E. Davis, J.J. Feddema, K.M. Klink, D.R. Legates, J. O'Donnell, and C.M. Rowe, 1985: Statistics for the Evaluation and Comparison of Models. *J. Geophys. Res.*, **90**, 8995–9005.

Zwiers, F.W., 1987: Statistical Considerations for Climate Experiments, Part II: Multivariate Tests. *J. Clim. Appl. Meteor.*, **26**, 477–487.

Zwiers, F.W. and H.J. Thiebaux, 1987: Statistical Considerations for Climate Experiments, Part I: Scalar Tests. *J. Clim. Appl. Meteor.*, **26**, 464–476.

END

DATE
FILMED
9 / 10 / 93