# MULTIVARIATE RECURRENCE ANALYSIS

Francis W. Zwiers
Canadian Climate Centre
Downsview, Ontario. Canada

Hans von Storch
Max Planck Institute for Meteorology
Hamburg, F. R. Germany

## 1. Introduction

The concept of recurrence analysis was recently introduced to the meteorological literature by Storch and Zwiers (1987), hereafter referred to as SZ. In that paper, the concept was dealt with in a univariate manner; in the present paper some of the ideas discussed by SZ are extended to the multivariate domain.

Recurrence analysis is a tool which can be employed In the analysis of climate experiments with General Circulation Models (GCMs), and in model-model and model-reality comparisons. The aim is to characterize aspects of the differences between two climates which are recurrent. For example, In an El-Niño sensitivity experiment, such as the one discussed in SZ, the aim is to discover aspects of the GCM's response to a given sea surface temperature (SST) anomaly which are likely to recur each time a new, independent realization of the perturbed climate Is simulated. Geometrically the concept is very simple: the difference between two climates is *recurrent* if there is very little overlap between the corresponding clusters of realizations of climate states. A precise definition of the concept in the case of a univariate state 'vector' is given In SZ. An appropriate definition for the multivariate case is given below.

## 2. Multiple Discriminant Analysis and Recurrence Analysis

SZ described several univariate statistical tests which can be used to test for a *recurrent* response. These tests can be characterized as one of two types. Either the minimum degree of recurrence which is of interest is specified *apriori* and a test is conducted to determine whether this level of recurrence has been obtained; or the degree of recurrence is estimated from the control and experimental samples, and a test is then conducted using this estimate. In either case it is possible to devise parametric and non-parametric tests.

Except for the language which has been employed, the methodology is essentially a univariate version of that used by statisticians in multiple discriminant analysis (MDA). In MDA it is presumed that observations can only come from one to two (or several) populations. Samples taken from the two (or more) populations in question are used to estimate a rule with which to classify future observations of unknown origin as belonging to one of the parent populations. Having estimated such a discrimination rule, probabilities of misclassification can be estimated, and the "significance" of the rule can be tested. The probabilities of misclassification are directed related to the degree of recurrence of the climate response: If the degree of recurrence is small, the probability of misclassification will be large and vice-versa. Anderson (1984) contains a good description of MDA in a rather general setting.

While the statistical literature on MDA is relatively rich, it is primarily concerned with the parametric problem in which it may be assumed that both samples come from Gaussian populations. Linear or quadratic discriminant functions arise depending upon whether one is willing to assume equality of variance-covariance matrices in the two sampled populations. In the absence of any information about the costs of misclassification or the relative likelihood that the next observation will come from the control or experimental population, the discriminant function is constructed by forming a likelihood ratio

$$L(z)=p_x(z)/p_y(z).$$

The decision Z *belongs to climate* X Is made if $L(z) > 1$. In the case of equal variance-covariance matrices $\Sigma$ the function reduces to

$$W = Z^t\Sigma^{-1}(\mu_x - \mu_y) - 0.5(\mu_x - \mu_y)^t\Sigma^{-1}(\mu_x - \mu_y)$$

and the decision Z *belongs to climate* X is made if $W > 0$. Because samples in climate comparison problems are always small (and particularly the experimental samples in climate experiments), it is a matter of necessity that we assume equality of variance-covariance matrices. The remainder of this paper assumes this to be the case.

Assuming that both populations are Gaussian, it is quickly shown that the discrimination statistic W has a Gaussian distribution with variance $\nabla^2$ and mean $\nabla^2/2$. The sign 01' the mean depends upon whether the observations comes from the control or experimental population. The quantity $\nabla^2$ is the Mahalanobis distance which is given by

$$\nabla^2 = (\mu_x - \mu_y)^t\Sigma^{-1}(\mu_x - \mu_y).$$

It then follows immedlately that the probability of misclassification is given by

$$1 - p = \phi(\Delta/2) \tag{2.1}$$

where φ is used to designate the Gaussian cumulative distribution function, This observation in turn leads to a definition ofp-recurrence In the multivariate case with Gaussian populations and equal variance-covariance matrices:

*The reponse, or difference between two climates, is sald to be* p-recurrent *lf the probabil1ty of misclassification using the linear discriminant function* W *is* 1-p.

In actual fact the variance-covariance matrix and difference of means vectors must be estimated before an estimate of the misclassification probability can be made. The simplest approach is to estimate the unknown parameters from the data and plug them into expression (2.1). However, this approach does not take sampling variability into account. The literature on MDA contains the derivation of several expressions similar to (2.1) which do attempt to take sampling variation into account, and these are reviewed by Page (1985). An alternate and more robust approach is to use the bootstrap (Efron, 1983) or crossvalidation (Snapinn and Knoke, 1984) to estimate the probability of misclassification. We have used the latter because of its simplicity and the fact that that it provides relatively good estimates of the probability of misclassification.

Unfortunately. the MDA literature is not very rich in non-parametric approaches to the classification problem. However. one approach, which has been suggested by Conover und Iman (1980), is to make an approximate transformation to the Gaussian distribution by replacing the observed vectors with vectors of ranks. The i'th component in a vector would be replaced by the rank of its value amongst the values of all other i'th components in both samples, Having completed the rank replacement, one proceeds as described above. Our experience is that this works very well, and that even with relatively small samples, very little is lost in sensitivity.

## 3. Constructing Tests for Recurrence Analysis

As noted above, two approaches may be taken to testing for recurrence. One method entails estimating the probability of misclassification by cross-validation, the bootstrap, or the "plug-in" method or a variant, and then making a decision on the basis of this information. We prefer this approach because it provides diagnostic information about the degree of recurrence of the experimental response. The second method requires that a statement is made *apriori* about the minimum level of recurrence which is of interest. A test is then conducted which looks for evidence of even greater levels of recurrence.

In the latter case we proceed as follows. Suppose that we wish to test the null hypothesis that the response is at least p-recurrent. We compute the Hotelling T2 statistic, given by

$$T^2 = \frac{n \bullet m}{n + m}(\bar{x} - \bar{y})^t S^{-2}(\bar{x} - \bar{y})$$

where S is the estimated variance-covariance matrix and n and m are the sizes of the control and experimental samples respectively. Under the null hypothesis

$$F = \frac{n + m - p - 1}{p \bullet (n + m - 2)} T^2$$

is distributed as a non-central F random variable with p and n+m-p-l degrees of freedom and with non-centrality parameter

$$\delta^2 = \frac{n \bullet m}{n + m} \Delta^2$$

From (2.1) we see that the non-centrality parameter of the F-distribution may be specified apriori as

$$\delta^2 = \frac{nm}{n + m} Z_{1-p}^2$$

where $Z_{1-p}$ is the 1-p'th quantile of the standard Gaussian distribution. Having derived the non-centrality parameter from the null-hypothesis it is then a relatively simple matter to determine the appropriate critical value for the test.

## 4. An Application

As an example, we re-analyzed the results of the El-Niño sensitivity experiments conducted with the Canadian Climate Centre (CCC) GCM which are described in SZ. In these experiments, the control sample consists of 30 December, January, February (DJF) means which were extracted from a 20-year and a 10-year control integration. The experimental sample consists of five DJF means extracted from simulations with anomalous prescribed SSTs. The S5T anomalies used were twice the Rasmussen and Carpenter (1982) anomaly (2RC), minus twice this anomaly (-2RC), and the observed DJF anomaly during the 1982/83 El-Niño (82/83). The experiments are described in more detail In 5Z.

The method of analysis consists of the following steps:

1) DJF means of 500 mb height and temperature fields were smoothed by averaging over 10° latitude by 20° longitude boxes.

2) the first five empirical orthogonal functions (EOFs) were computed from the control sample for data In the 30°N-30°5 band for each field. The computed EOFs were then used to reduce the dimension of the observed DJF means to five for both the control and the experimental climates.

3) a multiple recurrence analysis was conducted using the reduced data for each field and each experiment 80th the parametric and rank replacement techniques described above were employed. In each case the degree of recurrence was estimated by cross-validation.

The results for the parametric and rank-replacement techniques were found to be very consistent, and only those for the parametric technique are displayed in Table 1.

Table 1

| Var | Exp | $T^2$ | Sig. | Rec. |
|---|---|---|---|---|
| $T_{500}$ | 2RC | 110.8 | 0.00 | 1.00 |
| | -2RC | 83.2 | 0.00 | 1.00 |
| | 82/83 | 150.4 | 0.00 | 1.00 |
| $Z_{500}$ | 2RC | 58.0 | 0.00 | 1.00 |
| | -2RC | 145,5 | 0.00 | 1.00 |
| | 82/83 | 141.6 | 0.00 | 1.00 |

Table 1: Results of the multiple recurrence analysis in the 30°/N - 30°S latitude band. The analysis compares 30 control DJF means with five experimental DJF means derived from three El-Niño sensitivity experiments. The column headed "Sig." contains the p-value, or significance, of the computed $T^2$ statistic. The column labelled "Rec." contains the estimated degree of recurrence of the experimental result.

The results illustrate that the CCC GCM simulates significant and recurrent large scale responses to El-Niño type SST anomalies even if the local response is significant only near the equator (SZ).

The analysis described above was actually carried out for a hierarchy of latitude bands beginning with 10°N -10°S and extending to the 60° N - 60°S. At the beginning of this hierarchy, the EOF's represent possible response patterns restricted to the tropics. As the North-South extent of the band increases, the patterns encompass more extratropical variation, and eventually describe primarily extratropical variation. The results of the analysis using the largest band is displayed in Table 2.

Table 2

| Var | Exp | $T^2$ | Sig. | Rec. |
|---|---|---|---|---|
| $T_{500}$ | 2RC | 29.9 | 0.00 | 0.89 |
| | -2RC | 13.5 | 0.06 | 0.74 |
| | 82/83 | 56.3 | 0.00 | 1.00 |
| $Z_{500}$ | 2RC | 3.2 | 0.73 | 0.54 |
| | -2RC | 21.9 | 0.01 | 0.83 |
| | 82/83 | 12.0 | 0.09 | 0.69 |

Table 2: As Table 1, rxpect 60°N - 60°S.

The results illustrate that as we look for a response in larger spatial structures which describe primarily extratropical variation, the degree of recurrence weakens considerably. We see that the extratropical 500 mb temperature response is strongest in the experiment using observed SSTs (82/83) and that the response in the extratropical 500 mb flow is strongest in negative event experiment (-2RC). Also observe for 500 mb height in the -2RC experiment, that the T2 statistic Is significantly large, but the response is not particularly strong. The estimated level of recurrence is only 0.83, indicating that a realization of DJF 500 mb height from the -2RC climate would be misclassified as belonging to the control climate almost 20-percent of the time indicating that there is considerable overlap of the populations of seasonal means of these climates.

## 5. Conclusions

We have described an extension of the concept of recurrence analysis to multivariate problems by using the techniques of multiple discriminant analysis. Four statistical procedures were proposed; two of which are parametric and require the Gaussian assumption, and two of which employ a rank replacement prior to analysis. As an example, a parametric multiple recurrence analysis was conducted with the results of several El-Niño experiments conducted with the CCC GCH.

## 6. References


Anderson, T.W., 1884: **An Introduction to Multivariate Statistical Analysis**, Wiley, New York.

Conover, W.J. and R.L. Iman, 1980: The rank transformation as a method of discrimination with some examples. **Commun. Statist.-Theor. Meth., A9(5)**, 465-487.

Efron, B., 1983: Estimating the error rate of a prediction rule: Improvement on cross-validation. **J. Amer, Stat. Assoc., 78**, 316-331.

Page, J. T., 1985: Error-rate estimation in Discriminant Analysis. **Technometrics, 27**, 189-198.

Rasmuesen, E.M. and Carpenter, T.H.,1982: Variations In tropical sea surface temperature and surface wind fields associated with the Southern Oscillation / El-Niño. **Mon. Weath. Rev., 110**, 354-384.

Snapinn, S. M., and J. D. Knoke, 1984: Classification error rate estimators evaluated by unconditional mean squared error. **Technometrics, 26**, 371-378.

Storch, H. v., and F.W. Zwiers, 1987: **Recurrence Analysis of Climate Sensitivity Experiments.** Max Planck Instltute for 11eteorology, Report No. 3, Max Planck Institute for Meteorology, Bundesstrasse 55, D 2000 Hamburg 13, F. H. Germany.