MULTIPLE RECURRENCE ANALYSIS

Francis W. Zwiers
Canadian Climate Centre, Downsview, Ontario, Canada M3H 5T4

Hans von Storch
Max Planck Institut für Meteorology, Hamburg, F.R. Germany

1. Introduction

Ordinary statistical procedures for the comparison of two climates are generally based on point and interval estimates of the difference of means and corresponding tests of significance. While such objective procedures are a useful part of any analysis comparing a simulated climate with observations or another simulated climate, they do not paint a complete picture of the differences and may in fact obscure important information about the strength of the difference between two climates.

In particular, estimates of the difference between two mean states do not reveal, in an objective way, whether the observed pattern of differences can be recognized as a characteristic of individual realizations of the "experimental" climate. (Our discussion will be in terms of a climate experiment in which there is a "control" simulation and an "experimental" simulation which, for example, may have been forced by anomalous sea surface temperatures (SSTs) or may have an atmosphere of a different composition as in a CO_2 experiment). Tests of the significance of the difference of means also fail to address this question. Such tests are particularly difficult to interpret because the size of difference which can be perceived to be significant by testing procedures depends upon sample size. Even small, physically insignificant differences will be declared to be significant given large enough samples.

The correct interpretation of the word "significant" (from a statistical point of view) is that the observed difference is "unusual"; (i) relative to the null hypothesis that the means are equal, <u>and</u> (ii) relative to the statistical model which is implicit in the testing procedure. The fact the a difference is determined to be "significant" is only a flag that the difference is measurable in some statistical sense, and that it may bear further investigation. We propose that such further investigation should include statistical procedures which are addressed to the question "Are there characteristics of the experimental climate which can be recognized in every realization of the experimental climate?". In many experiments, such as El-Niño sensitivity experiments, it is such differences which provide the important clues to our understanding of the response mechanism.

We have developed procedures for estimating how "recurrent" an estimated response is, and for testing hypotheses about the degree of recurrence. We say that a response is <u>P-recurrent</u> if the probability that an unknown climate realization either either the control climate or the experimental climate is correctly classified is P. A complete description of the techniques and their derivation may be found in Zwiers and Storch (1988; hereafter referred to as ZS)

The recurrence estimators which we have developed are based on Multiple Discriminant Analysis (MDA) methodology. The MDA literature is replete with descriptions of many "misclassification error rate estimators" which can be exploited as recurrence estimators. We give a very brief overview of some of the relevant ideas. In the Gaussian setting, a difference between two climates is P-recurrent if the control and experimental climate means, say μ_c and μe are separated by Mahalanobis distance Δ^2= $(\mu_e - \mu_c)\,^t\sum^{-1}(\mu_e - \mu_c)$ where $\sum$ is the common variance-covariance matrix and where Δ is related to P through the relation $P = \Phi(\Delta/2)$. Here $\Phi(\cdot)$ represents the Gaussian cumulative distribution function. Thus P can be estimated by estimating Δ. A very efficient, and nearly unbiased estimate of Δ^2 can be had by replacing μ_e, μ_c and $\sum^{-1}$ with their respective unbiased estimators. The result is the so-called DS estimate of Δ^2 which is given by $DS = [(n_e+n_c-k-1)/(n_e+n_c-2)]\cdot(\bar{x}_e-\bar{x}_c)\,^tS^{-1}(\bar{x}_e-\bar{x}_c)$ where n_e and n_c are the control and experimental climate sample sizes respectively, $\bar{x}_e$ and $\bar{x}_c$ are the corresponding sample means, k is the dimension of the observed climate state vectors and S is the usual pooled estimate of the variance-covariance matrix of the observed state vectors. Other "parametric" estimators of Δ^2 which we have studied include the "D-method" (which is optimistically biased) and the "M-and OS-methods" which are both based on asymptotic expansions. The latter two are computationally more complex than DS but have about the same statistical properties.

"Non-parametric" estimators can also be constructed through the application of the "linear discrimination function" (see ZS). The so called "R-method" estimator is constructed by estimating a discrimination rule and using it to classify the observations in both samples. As one might expect, this estimator is optimistically biased. This problem is effectively cured with the "U-method" estimator. Here a single observation is deleted from one of the two samples. Then all remaining observations are used to specify a discrimination rule which is then used to classify the withheld observation. This process is repeated for all observations in the two samples, and the proportion correctly classified is used as the estimate of P. The literature also contains descriptions of several estimators which are based on the bootstrap procedure. The nonparametric estimators are generally more robust than their parametric counterparts, but they are somewhat less efficient and their estimates of P are somewhat more variable.

Parametric and non-parametric tests of a priori specified values of P can also be constructed. The parametric approach is a modification of the usual Hotelling T^2 test for equality of means, the difference being that the observed value of the test statistic is compared to the critical values of a non-central (rather than central) F distribution. The non-centrality parameter, and consequently the critical values, increase with the value of P specified in the null hypothesis and the sample sizes. The result is that much stronger evidence is required to conclude that an experimental response is significantly greater than P-recurrent than is needed to conclude that the response is significant in an ensemble mean sense.

3. An Application

In ZS we describe an application of recurrence analysis to a sequence of El-Niño sensitivity experiments conducted with the Canadian Climate Centre General Circulation Model. Details of the results of the experiment may be found in Boer (1985) and Storch and Zwiers (1988). The experiment consists of three sets of five climate simulations with anomalous SST distributions. Specifically, the observed 1982/83 December/January/February (DJF) SST distribution and two simulated SST distributions with anomalies corresponding to +2 and -2 times the standard Rasmuessen and Carpenter (RC) SST anomaly were imposed. Samples of five DJF 500 mb height fields were obtained from each experiment and compared with a sample of 76 DJF 500 mb height fields obtained from 3 extended climate simulations with prescribed annually varying climatological SSTs.

Comparisons between control and experimental climates were made in two latitude bands (30°S - 30°N and 60°S - 60°N) using two EOF truncations (EOFS 1-5 and EOFS 1-10 of the 76 control climate realizations). The severe reduction of dimensionality is required because the techniques described above are "multivariate". An additional side benefit of a recurrence analysis in this case is that discrimination rules are produced which may be used to test the fidelity of the model's response to the experimental SST conditions. We tested these rules by applying them to a set of observed DJF mean NH 500mb height anomalies. The general findings of the analysis are:

- in the 30°S - 30°N band the response is significant, and also significantly greater than 84%-recurrent. In fact, estimates of P ranged from 98% to almost 100%. "Sun dial" plots of the response indicate that the +2xRC and 82/83 SST anomaly responses have somewhat different structure (the signal is not projected onto the same set of EOFs) although they are of comparable magnitude as measured by P or the T^2 statistic. Such plots also show that the broad structure of the response to the cold SST anomaly is opposite to that of the response to the 82/83 anomalies. In all cases it is possible to characterize the response in such a way that individual realizations can be easily identified.

- in the 60°S - 60°N band the response is also significant but not significantly greater than 84% recurrent. Estimates of P ranged between 83% and 93%. In this large region, the first 10 EOFs characterize primarily temperate, northern hemisphere features of the control climate. It is thus not surprising that the response is not found to be strongly recurrent.

- discrimination rules based on the model's response in a 20°N -30°N latitude band were remarkably successful at distinguishing between warm and non-warm years (using the +2xRC experiment) and cold and non-cold years (using the -2xRC experiment) when the rules were used to classify observed DJF NH 500mb height anomalies derived from (mostly) NMC analyses made between 1955 and 1984. Classifications were somewhat less successful with the discrimination rules derived from the 82/83 experiment. This is perhaps due to the difference in structure between the model's responses to the +2xRC and 82/83 SST anomalies. The success of these rules is a strong indication of the fidelity of the model's response to anomalous SST's.

4. Summary

We have shown that recurrence analysis is a powerful tool with which to extract more information from climate comparison studies. Not only do we learn about the significance of a response, but we also obtain information about the relevance to the response to individual climate realizations. In some climate experiments it may also be possible to use discrimination rules that arise from a recurrence analysis to test the fidelity of the model's response to anomalous conditions by applying the rules to archives of climate data. The use of recurrence analysis has been illustrated by applying it to a set of El-Niño sensitivity experiments conducted with the Canadian Climate Centre General Circulation Model.

5. References

Boer, G.J., 1985: Modeling the Atmospheric Response to the 1982/83 El-Niño. Coupled Ocean-Atmosphere Models, J. C. J. Nihoul, editor, Elsevier, Amsterdam, 7-17.

Storch, H. von and F.W. Zwiers, 1988: Recurrence Analysis of climate sensitivity experiments. J. Climate, 1, 151-171.

Zwiers, F.W., and H. von Storch, 1988: Multiple Recurrence Analysis. Report No. 17, Max Planck Institut für Meteorologie, Bundesstraße 55, D 2000 Hamburg 13, F.R. Germany. 49pp.