PCMDI

Program for Climate Model Diagnosis and Intercomparison

UCRL-JC-126621

PCMDI Report No. 40

COMPARING GCM SIMULATIONS, ENSEMBLES AND MODEL REVISIONS USING COMMON PRINCIPAL COMPONENTS

by

Sailes Sengupta and James S. Boyle

May 1997

PROGRAM FOR CLIMATE MODEL DIAGNOSIS AND INTERCOMPARISON
UNIVERSITY OF CALIFORNIA, LAWRENCE LIVERMORE NATIONAL LABORATORY
LIVERMORE, CA 94550

Abstract

The technique of common principal components (Flury, 1988) is applied to compare the results of a number of GCM simulations. The data used are from 30 AMIP simulations and an ensemble of five AMIP integrations from a single GCM. As suggested by the name, the common principal components analysis of these data identifies the common elements of the integrations. The components are determined in the spatial domain. The results are for the seasonal cycle of precipitation over the US and for the seasonal and interannual variations of global 200 hPa divergence.

The precipitation analysis shows that the models share only the broadest aspects of the seasonal cycle in precipitation and the common vector loadings vary substantially between the models. The analysis of the ensemble integrations for the same variable indicates that the model variations are generally quite a bit larger than the variations amongst the members of the ensemble. The ensemble results give some confidence that the differences among the models are robust and that they would not change greatly for other realizations of the model integrations. In itself, the ensemble analysis displays the usefulness of the CPC technique in succinctly combining the output of any number of realizations from a given model. Their common character is a robust signal and is the answer often sought from ensemble experiments.

The CPC analysis of the five ensemble simulations shows that the interannual variations in the divergence are dominated by the ENSO events for the AMIP decade, 1979-1988 The analysis indicates that the ENSO global signal is robust across all the simulations such that one simulation is all that is necessary to characterize the global response. The intrinsic variability of the model begins to dominate the components higher than the first.

The difference between the 200 hPa divergence of four closely related AMIP models and the NCEP /NCAR reanalysis was analyzed using the CPC technique. The models all share the ECMWF dynamical core. When compared to the analogous analysis of the ensembles, it is seen that the models form two distinct pairs although some distinctive characteristics are retained. The CPC is shown to have value in documenting the impact of modifying parameterizations on the simulations.

The CPC analyses tend to support the observation that the models often have more in common with each other than with the observations. The CPC approach has utility in answering many useful questions posed in the arena of model comparison when used in conjunction with other techniques.

1. Introduction

The Atmospheric Model Intercomparison Project (AMIP) of the World Climate Research Programme provides an infrastructure for the comparison of atmospheric general circulation models (AGCMs) and their response to the specified SST variations. The participants in AMIP simulate the global atmosphere for the decade 1979 to 1988 using a common solar constant and CO_2 concentration, and a common sequence of monthly averaged SST and sea ice data set. An overview of AMIP is provided by Gates (1992). In this work a statistical tool is presented to address the task of model comparison and verification.

The AMIP was intended to document the state of AGCM modeling and to facilitate diagnosis of the causes of any differences that showed themselves. The models do display a number of large differences, but it is not a trivial task to figure out the causes of these. For example, the tendency of the models to have poles which are too cold and a tropical region which is too warm has been documented for some time (Boer et al., 1991), but if there is any fundamental flaw which explains this error it has yet to be unambiguously identified.

There have been numerous useful statistical techniques put forth for the purposes of model verification and analysis. Many of these have as their major thrust a pairwise comparison. The model output is compared to the observations (e.g. 500 hPa geopotential) or the relationship between two variables(e.g. SST and precipitation) is examined between the models and the observations. This approach is well summarized in the comprehensive work of Bretherton et. al (1992). The AMIP analysis presents a slightly different variation on this theme. The large number of models to compare (~30) places a practical restriction on the type of pair-wise analyses that can be carried out. There is also value in ascertaining the systematic errors that cut across all the models. Presumably, such errors would indicate a fundamental gap in understanding or a defect in parameterization implementation that might be corrected if identified.

The desire for a concise analysis of systematic errors gives rise to two potentially conflicting requirements. First, there is a need for a parsimonious representation of the model and observed spatial and temporal evolution to facilitate comparison given the large mass of data. Second, it is important to be able to identify the physical processes that are related to the model errors and so the analysis needs to preserve suf-

ficient spatial and temporal detail so that this can be accomplished. For example, global mean temperature is an efficient reduction of the temperature data, but these data alone will most likely not reveal the processes responsible for deviations from the observations.

An additional complicating factor in model verification comparison is the realization that the models are chaotic in the sense that some aspects of a simulation can change substantially when started from a slightly different set of initial conditions. It is important to be able to isolate the common aspects of an ensemble of model simulations. These aspects are presumably the robust features that are representative of systematic errors and basic behavior of the model.

The purpose of this paper is to put forward the technique of Common Principal Components (CPC, Flury 1988) as a framework for model comparison and as a useful way to incorporate ensemble information. This technique has been applied recently to the comparison of ocean models by Frankignoul et al. (1995). This technique will be seen as a useful complement to the other analysis tools documented in the literature.

In the next section the basic concepts associated with CPCs will be outlined and some comparisons drawn to previously published techniques. In section three the data sets used in the following applications are described while section four presents a number of applications of the CPC technique. These point out the usefulness and interpretation of the CPCs and how they might complement other methods. This section also makes some points about the character of the AMIP integrations. Section four discusses some conclusions and extensions of the techniques, and a summary of a strategy for model comparison, including ensemble integrations.

2. Common Principal Components

Common principal components are most easily described in comparison to the closely related principal component analysis. Principal components(PC), also referred to as Empirical Orthogonal Functions(EOF), have a long history in atmospheric analysis since being introduced by Lorenz (1956). PCs are invaluable tools in that they provide an efficient method of compressing the data in both space and time and present the results in terms of independent modes of variability. The principal vectors are the eigenvectors of the data covariance matrix whose elements are formed from

the differences from some specified means. Each successive eigenvector is orthogonal to the set of previous ones and explains the maximum amount of possible remaining variance in the data. The eigenvectors are usually arranged in order according to the percentage of variance explained. The PCs are derived directly from the data themselves as opposed to some *a priori* set of functions such as in Fourier analysis. Often, but not necessarily, the leading vectors can be associated with some aspects of physical processes.

CPC is a generalization of the PC technique to the case of several groups. Rather than a single covariance matrix, there are now two or more. The basic assumption is that the principal component transformation is identical in all the populations considered, while the variances associated with the components may vary between groups. Thus, similar to PC analysis, the output of the CPC results in a set of spatial patterns (vectors), but unlike PCs there are more than one time series associated with each pattern set. The ordering of the eigenvectors varies by group, and it is by no means necessary that all the groups share the same ordering of the vectors. An advantage in using the CPC model is that one can compare *corresponding* principal components. A formal test of significance for the hypotheses of (partial) commonality of the principal axes of representation of two (or several) fields of data along the line given in Flury (1988) is, however, not possible to implement directly. These tests of significance require that the sample fields (over discrete time instants) be independent. This is generally an incorrect assumption for almost all meteorological fields. This problem itself does not preclude the use of common principal components as a diagnostic tool for understanding the commonality of the fields. The temporal correlation can also be addressed by a sparser temporal sampling. This is impractical for the short span of time represented by the AMIP simulations, but could be implemented for much longer integrations.

In terms of pairwise comparisons there exist powerful techniques, summarized in Priesendorfer(1988) and Bretherton et al. (1992). If there is a fiducial field to compare against, say the observed fields, then a commonly used method is to determine how the principal modes of the other groups project onto those of the basefield. The CPC analysis does not replace this technique, but rather complements it. The most effective means to illustrate how the CPC analysis can complement other methods is to provide some examples. As with any diagnostic technique there are situations where its use is inappropriate and other times when it can provide useful informa-

tion. The next two sections will present examples on data sets where the CPC technique has proved useful.

The algorithm used for determining the common eigenstructure was that of Flury and Gautschi (1986). The code was tested against the IMSL (1991) routine KPRIN and the results were identical. The IMSL routine was not used since it was desired to have access to some intermediate results and the IMSL routines were unable to permit this. The covariance matrices were computed using the IMSL routine CORV. The principal component analysis was carried out using the IMSL routine PRIN.

4. Data sets

The data used consisted of observationally based analyses and the corresponding fields produced by AMIP simulations. The data were all monthly means for the 120 months from January 1979 to December 1988.

Precipitation over the U.S.

The data used in this part of the study consists of precipitation observations gridded to a 4 degree latitude by 5 degree longitude grid. The observations are from surface stations over land from Schemm et al. (1992), and satellite Microwave Sounding Unit (MSU) estimates from Spencer (1993) over the oceans. The bulk of the analysis grid used here is over the United States, where the observational network provides reliable precipitation fields. The 120 month mean was subtracted from each gridpoint to form the deviations. The seasonal cycle was retained since it was of interest to compare how well the GCMs simulated this cycle. The data comprised a matrix of 120 time points at 95 space points. Figure 1 shows the spatial coverage of the data. The model data were interpolated to the observational grid using an area weighting scheme which preserved the spatial mean.

Global velocity potential

The input data for these calculations was the velocity potential computed from the 200 hPa winds. The observational data were from the NCEP/NCAR reanalyses. The model data were from two sources: the first was the from the AMIP integrations, and the second was a small ensemble of AMIP integrations using the ECMWF AMIP model. This ensemble had five members, each integration differing only in the initial conditions used. The initial conditions for the first run were the observed data for 1

Jan 1979, while the initial conditions for the subsequent runs were taken from the ending state of the previous run.

All the data were transformed to the orthogonal spherical harmonics and the spherical harmonic series was truncated at T10. This limits the results to large scale features but allows a fit in the spatial domain since there are 110 spatial coordinates (110 coefficients of the spherical harmonics decomposition) and 120 time points. This truncation is also well within the horizontal resolution of all the AMIP models. From the basic monthly data two sets of deviations were computed for the covariance matrix to be used for intercomparison. The first described the interannual variations, in which the seasonal cycle was removed by subtracting from each month the 10 year mean of that month. The second set retained the seasonal cycle but computed the differences between the models and the NCEP/NCAR reanalysis fields for the 120 months of the AMIP decade. Insofar as the reanalysis depicts reality these data could be considered error fields. From these sets of spherical harmonic data the covariance matrices were formed for input into the PC and CPC algorithm.

3. Applications of CPCs to Data

U.S. precipitation

The leading two principal vectors of the observational precipitation data set are shown in Fig. 1. The time series of the leading three principal vectors are in Fig. 2, while the percent variance explained by the leading four PCs is shown in Table 1. The leading PC can be interpreted in terms of the seasonal cycle of precipitation described by Hsu and Wallace (1976) and Horn and Bryson (1960). There is a winter maximum of rainfall on the west coast and off the Gulf and east coast, and a summer maximum in the central US. It is useful to point out the high level of interannual variability displayed by the principal components in Fig. 2. The PVs were also computed from the 30 individual precipitation fields of the AMIP models. Just considering the first two vectors, this meant comparing 60 figures. From a first examination of the PVs. the models appeared to be very poor in their simulation but it was difficult to make any general categorization. A CPC analyses was then performed on the 31 data sets (30 models + observations). The percent variance explained by the leading three CPCs for the observations is shown in Fig. 3.

Figure 3 can be used to illustrate a few points about this analysis. The first is

that the percent variance explained can be ordered differently for each data set. In Fig. 3 we have arbitrarily chosen to use the ordering of the observed data set, but for some models these are clearly not the leading sequence. The second point is that the common mode dominant among the models is not the first mode for the observations, but the second mode. This information was quite useful in making sense of the PV charts of the individual models. Going back to the individual plots of the PVs and comparing the second PV of the observations to the leading PV of the models indicated a systematic problem with many of the models in depicting the precipitation over the eastern and central US. This correspondence was not at all obvious when trying to look through all the charts due to the many slight variations displayed. The dominant seasonal cycle of the models resembles Fig. 1b, rather than 1a. The models tend to place the precipitation characteristic of the central US too far eastward. Figure 3 lends credence to the common wisdom that the models tend to look more like each other than the real atmosphere.

In Fig. 3 it can also be noted that the models generally have a larger percent variance for a single CPC mode than do the observations. Those that do not in the three vectors displayed inevitably have such a component elsewhere. This is also true for the individual PC computations. Note that the percent variance of the observational data set has dropped substantially between Table 1 and Fig. 3. In Fig. 3, the fit is to a common set of vectors, and the models have such a large variation compared to the observations that the common fit cannot be expected to be as good as in the PC case. The corresponding CPC vectors for the observed set (not shown) do, however, bear a close qualitative resemblance to Fig. 1. Table 2 gives the percent variance explained by the leading principal components of the UCLA AMIP model, whose data are fairly typical of the model behavior. The steeper spectrum is an indication that the models occupy a simpler world than that depicted by the observations in Fig. 2 and Table 1. Figure 4 shows the time series of the leading three PCs of the UCLA model; every year is much like the next.

Beyond common characteristics of the models, Fig. 3 illustrates the substantial differences between the models. The models with low values in Fig. 3 do not have a large variance explained by any of the three leading CPCs of the observations. These outliers have less in common with the observations or wiht other models whose peaks in variance explained occur at other components. The models and observations do share a common variation in the west coast precipitation. This is likely a conse-

quence of the fact that the models all use the same set of observed SSTs which play an important, and evidently dominant, role in determining the variation of the rainfall in the western US.

This example illustrates that the CPC analyses can illuminate some key differences and commonality of the models and observations. These facts can then be taken into account as other tools are brought to bear on the problem. It can guide the choice of pair-wise comparisons, or indicate models that are so far off that they might not be profitable to examine further.

200 hPa velocity potential

Ensemble comparison

Outside the arena of multiple model comparison, the analysis of ensembles of simulations presents the potentially most useful aspect of CPCs for the analysis of GCM output.

Figure 5 presents the time series of the leading CPC for the 200 hPa velocity potential for the five simulations of the ECMWF model ordered with respect to the AMIP simulation. These data represent the interannual variations, the monthly means being removed. The ensemble members all share the same first four CPCs, although they vary slightly at higher components. The leading four explain more than 80% of the variance and Table 3 gives the percent variance explained by the leading three. Two things are evident from Fig. 5. First, the simulations all follow a similar time evolution which clearly reflects the pattern of the ENSO activity for the decade. This is clear from comparing Fig. 5 to Fig.6 which is a plot of the Southern Oscillation Index (SOI) from the Climate Analysis Center. This index is the difference in atmospheric pressure between Darwin, Australia and Tahiti. It is tightly linked to the cycle of the equatorial Pacific SST and the atmospheric response to the SST. The two distinct dips in 1982/83 and 1986/87 represent two strong ENSO events, the 82/83 event being the strongest on record. Figure 5b makes use of the same data as Fig. 5a except that it shows the differences in each simulation from the mean of all the simulations at each time point. These difference curves are in a sense a measure of the non-deterministic component of the flow. The mean time series has an ENSO signal that clearly rises above the noise level during the larger excursions of the SOI. During the interim periods one cannot distinguish the influence of the SST variations from

the model's intrinsic noise.

Figure 7 is a geographical plot of the divergence pattern of the leading CPV. It should be noted that most of the amplitude of the signal is over the Tropical Pacific and the pattern is broadly consistent with that expected from observed precipitation anomalies associated with ENSO events. There is enhanced upper level divergence in the eastern equatorial Pacific. The CPC data compression retains sufficient information to be able to make some physical interpretation of the modes identified.

The point to be made here is not the discovery of new relationships but a measure of the impact of the SSTs on the simulations with varying initial conditions. The ECMWF model appears to have a robust, reasonable simulation of the ENSO response for the global 200hPa divergence.

Ensemble integrations are now commonplace among the major weather forecasting centers, Toth and Kalnay (1993). The CPC methodology provides a framework for combining ensembles into a single field. This combining is necessary since the number of members of the ensemble is often greater than twenty. This would provide more information than can be easily assimilated by a forecaster and the CPC technique provides a consensus summary which is usually the type of information needed. There are some indications that the SST are predictable a month or a season in advance, and if the atmospheric models are then driven by these SSTs, a climate prediction can be made. A CPC analysis of an ensemble of such atmospheric predictions would be an efficient way of producing a robust climate forecast.

Figure 8 is the same as Fig. 5b, except for CPC 2. In these curves there is also an influence of the ENSO variations. The mean curve is less above the "noise" level than in Fig. 5b. By the third CPC (not shown) the mean curve is almost completely engulfed by the ensemble variations, which restricts the number of conclusions that can be drawn using a single run or even five runs from this perspective.Figure 5 indicates that beyond the ENSO maxima the model and the observations do not have a great deal in common for this mode. Each has a different response given a common SST forcing for the decade. This is not unexpected since on the global scale a great many more variables influence the interannual variability of the upper level divergence field besides the equatorial Pacific SSTs.

In the foregoing we have presented some aspects of the relationships of the members of the ensemble to each other. A logical next step is to ascertain what relation the five ensemble members have to the observed data. An easy path is to just include the

observations as another data set with the five ensembles in performing the CPC analysis. The percent variance explained by the leading three PCs for this analysis which now is over six data sets (five ensemble members + NCEP reanalysis) is shown in Fig. 9. Figure 10 is the time series for the leading two CPCs for this analysis with the NCEP reanalyses shown in the solid line. Figure 10a shows that the leading mode is associated with the ENSO variations. Comparing Figs. 10a and10 b indicates that the model does a fair job in tracking the ENSO variations for the period, however, beyond the first CPC the correspondence almost vanishes. In Fig 10a the ensembles show an agreement with each other and the observations, while in Fig. 10b the ensembles agree with each other but are at odds with observations. The percent variance explained (Fig. 9) indicates that the leading vector is more dominant in the observations. It is clear that this mode in the observations lies outside those of the members of the ensemble. A single run would be adequate to capture this aspect of the ECMWF AMIP integration.

To formalize this conclusion, a non-parametric randomization test (Noreen, 1989), has been performed for the percent variance explained by the CPCs to test the hypothesis that the observations could come from the population indicated by the model simulation. The randomization test was applied on the percentage of variance explained by each of the first three common principal components as follows.

For each of the three CPCs two sets of differences of explained percentages of variance are computed. The first set consists of the five differences in absolute value obtained by subtracting the observed PVE from those of the five simulations. The second set consists of the ten absolute differences in PVE derived from the ten possible pairs selected from within the group of the five CPCs of the five simulations. Thus the two sets of differences represent respectively the model/observation difference and the within-model differences for each CPV. Table 4 summarizes the procedure and findings. The randomization test permutes the combined set of 15 differences in all possible manners computing the simulated distribution of the difference of the mean of the first 5 with that of the rest. It then estimates the probability of the differences being as high as or higher than the actual difference. Let ϕ denote the fraction of time that under random permutation of the fifteen, the mean of the first five exceeds the mean of the remaining ten by more than the actual value. The estimated values of ϕ in the three CPCs were 0.002, 0.0, and 0.749 respectively. From a sampling distribution of the estimated ϕ the probabilities are computed leading to the significance lev-

els in Table 5. The values in Table 5 indicate that ϕ is almost certainly less than 0.01 for the first two CPCs and almost certainly greater than 0.1 for the third.

Model development and evolution

Divergence error fields

In the work presented next, the difference fields computed by subtracting the 200 hPa divergence of the NCEP/NCAR reanalysis from the AMIP models are compared. In this case the common behavior would reflect some common error in the models. In these data the seasonal cycle was retained, and the covariance matrices were formed from the deviations from the 120 month mean.

The most brute force approach is to take all the 30 models for which we had data, and to calculate the CPCs for the difference fields between the models and the reanalysis. The results for this analysis are not shown since they did not reveal much. The leading three common vectors, which were shared by all the models, closely resembled the PC analysis of the NCEP/NCAR reanalysis. This could be interpreted as meaning that the models are all in error, and that what they share most in common is a difference with the reanalysis with no particular common pattern. Going into the analysis, it was hoped that the common error patterns might indicate specific regions or times when the models had particular problems in simulating the upper level divergence. It might be thought that there would emerge specific locations where the convective parameterizations would evince a common breakdown, especially in the Tropics. However, for these data there is not a common localized systematic error as was evident in the US precipitation data. The models' common feature is that they are different from this observational data set, but evidently they are different in a host of ways.

The CPC analysis was then applied to a selected subset of four models, which a priori were expected to have some common type of error patterns. The models chosen for this analysis were the ECMWF, UGAMP and MPI (ECHAM-3) AMIP models and the MPI (ECHAM-4) model run using the AMIP boundary conditions. These models were chosen since they all share the same basic formulation of the ECMWF model. Indeed, they all started from the same code. The UGAMP differs from the ECMWF only in the penetrative convective scheme used. UGAMP uses the Betts-Miller con-

vective adjustment while the ECMWF uses the Tiedtke mass flux scheme, as do the MPI models. Although sharing the same convective parameterization, the MPI and ECMWF models differ in many ways as documented by Phillips (1994). The chief differences are the treatment of the land surface processes and the radiation, particularly the interaction with clouds. Figure 11 presents the percent variance explained for the leading three vectors. It is evident that the ECMWF and UGAMP models and the two MPI models form two pairs. It might be felt that for this field, upper level divergence, the convective parameterization might play an overwhelming role in determining the model characteristics. However, in this case the models sharing the same convective scheme are different, while the two models alike except for this parameterization are similar. Figure 12 are the time series of the two leading CPCs for this group of models.

The MPI changes were intended to improve the performance of the ECMWF forecast model in climate simulations. The MPI models show a clear reduction in amplitude of the leading CPC, which is dominated by the seasonal cycle. The two MPI simulations appear to differ mostly in amplitude, except during the ENSO events of 82/83 and 86/87 when they also get out of phase. Table 6 provides the mean absolute values of the time series, which in the case are actually errors with respect to the NCEP/NCAR reanalysis. There is actually a slight increase in the difference for the second vector for the MPI models with respect to the other two. Figure 13 shows two leading CPVs. There is considerable amplitude in the Tropics, which might be expected but also there is significant contribution in the midlatitudes. The maximum centered just west of Central America is a characteristic error for this suite of models. The variations of the monsoon over eastern Asia are evident in the figure.

The ensemble of five ECMWF integrations were analyzed for the velocity potential difference field in order to be able to judge if the variations between the single runs of the four models lie outside what might be expected from the intrinsic variability. It would be better to run such an ensemble analysis on each model, but the ECMWF model is most probably a fair proxy since they all share the ECMWF dynamical core and many parameterizations. The results are shown in Figs. 14 and 15. What is interesting is that the UGAMP and ECMWF models are actually indistinguishable given the variability represented by the ensemble by the tests outlined above. It would appear that it would take a larger number of integrations to establish if the UGAMP and ECMWF are truly unique from this particular perspective.

4. Discussion and conclusions

CPCs have been shown to be useful in the comparison of the output of a large number of GCMs. This type of analysis attempts to characterize the systematic model errors, as here is value in ascertaining the systematic errors that cut across all the models. Presumably, such errors would indicate a fundamental gap in understanding or a defect in parameterization implementation that might be corrected if identified. CPCs have also demonstrated utility in as a natural way to summarize the robust results of an ensemble of simulations from a single model. In the future, this technique will be applied to time series of longer simulations. One idea is to compare the common components between decades of a multidecadal coupled climate system model.

It is important to indicate what the CPCs are not. The technique is not a replacement for describing the coupled patterns of, say SST and 500 mb geopotential. The analysis of coupled patterns is well described by Bretherton et al. (1992) using SVD, CCA, CAA, PC. Nor is the CPC a replacement for the simple PC. The CPC are a compromise among the covariance matrices provided to the algorithm and do not preserve the powerful concise description inherent in PCs. The CPC does not permit the rotation of individual members of the group being analyzed. In this work the CPC analyses has been used to complement the PC information.

For the purpose of comparing two fields the technique of projecting the model filed onto the observational PC, as described by Priesendorfer (1988), provides information that the CPC cannot. However, there are fields for which the observational quality is dubious. In any case the CPC provides a kind of consensus viewpoint of corresponding components which has been shown to be useful. The CPC technique is well worth adding to the tools of data analysis. The concise description of ensemble data probably holds the greatest promise for general modeling beyond the AMIP type intercomparisons.

A formalism for the space-time comparison of geophysical data can be built by first identifying the commonality in the spatial structure by means of a CPC analysis, followed by a comparison of the corresponding CPCs as time series. 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. This notion of predictability of one series in terms of the other can be extended one

more step by regarding the two series in a bivariate context. More precisely, one may consider one series as a (linear)filtered version of another and estimate in an 'optimal' way the filter coefficients (Newton 1988). One may expect to do a little better if one allows for nonlinearity in the filter.

One can also look at two time series of a specified CPC pair resulting from two model outputs (or a model output and observations) and find a predictive function of one in terms of the other. 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 leading CPCs. In the context of model intercomparison, two data sets (model/model or model/observation) would be considered as 'similar' with increasing degree of similarity in the order indicated below, if :

(i) The significant common principal components within each pair under comparison explain a 'large' portion of the variations in the fields under comparison and

(ii) a high degree of predictive skill is demonstrated when one of the series is used in the prediction of another.

Acknowledgments. The cooperation of the ECMWF in making their forecast model available and in providing expert technical advice for this research is gratefully acknowledged. This work was performed under the auspices of the Department of Energy Environmental Sciences Division by the Lawrence Livermore National Laboratory under contract W-7405-ENG-48.

5. References

Boer, G. J., K. Arpe, M. Blackburn, M. Deque, W. L. Gates, T. L. Hart, H. le Treut, E. Roeckner, D. A. Sheinin, I, Simmonds, R. N. B. Smith, T. Tokioka, R. T. Wetherald, and D. Williamson, 1991: An Intercomparison of the Climates Simulated by 14 Atmospheric General Circulation Models., WRCP - 58, WMO/TD - No. 425, World Meteorological Organization, Geneva.

Bretherton, C.S., C. Smith and J. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Clim.*, **5**, 541-560.

Flury, B.,and Gautschi, W., 1986: An algorithm for simultaneous orthogonal transformation of several positive definite symmetric matrices to nearly diagonal form. *Siam.Jour. Sci. and Statist. Comp.*, **7**, 169-184.

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

Frankigoul, C., S. Fevrier, N. Sennechael, J. Verbeek, and P. Braconnot, 1995: An intercomparison between four tropical ocean models: Thermocline variability. *Tellus*, **47A,** 351-364.

Gates, W. L., 1992: AMIP: The atmospheric model intercomparison project. *Bull. Amer. Meteor. Soc.*, 73, 1962-1970.

Horn, L. H., and R. A. Bryson, 1960: Harmonic analysis of the annual march of precipitation over the United States. *Ann. Assoc. Amer. 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**, 1093-1101.

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

Lorenz, E. N., 1956: *Empirical Orthogonal Functions and Statistical Weather Prediction*, MIT, Dept. of Meteorology, Science Report 1, 49pp.

Phillips, T. J., 1994: A summary documentation of the AMIP models. PCMDI Report No. 18, Program for Climate Model Diagnosis and Intercomparison, University of California, Lawrence Livermore National Laboratory, Livermore, CA, 343pp.

Preisendorfer, R. W., 1988: 'Principal Component Analysis in Meteorology and Oceanography,' Developments in Atmospheric Science 17, Elsevier, Amsterdam, 425pp.

Noreen, E. W., 1989: *Computer Intensive Methods for Testing Hypothesis: An Introduction*, John Wiley and Sons, Inc., New York, 238pp.

Schemm, J. S., S. Schubert, J. Terry, and S. Bloom, 1992: Estimates of monthly mean

soil moisture for 1979-1989. NASA Technical Memorandum 104571, 254pp

Spencer, R. W., 1993: Global oceanic precipitation from the MSU during 1979-91 and comparisons to other climatologies. *J. Clim.*, **6**, 1301-1326.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317-2330.

PC 1	PC 2	PC 3	PC 4
25	13	12	6

Table 1: Percent variance explained for principal components of observed precipitation

PC 1	PC 2	PC 3	PC 4
38	14	4	4

Table 2: Percent variance explained for principal components of precipitation simulated by the UCLA model

Model run	CPC 1	CPC 2	CPC 3
1	37	31	7
2	39	28	9
3	41	25	9
4	38	28	7
5	43	24	8

Table 3: Percent variance explained for common principal components of 200 hPa velocity potential of ECMWF AMIP model

	Model/obs	within model	difference of means	estimated ϕ
CPC - 1	8,6,4,7,4	2, 4, 1, 4, 2, 1, 2, 0, 3, 3	3.6	0.002
CPC -2	11, 9, 5, 8, 7	2, 6, 3, 4, 4, 1, 2, 3, 2, 1	5.2	0.0
CPC - 3	2, 0, 0, 2, 1	2, 2, 0, 1, 0, 2, 1, 2, 1, 1	-0.2	0.749

Table 4: Model / observation and within-model absolute difference in PVE for the leading three CPCs

	P(ϕ<0.01)	P(ϕ<0.05)	P(ϕ<0.1)
CPC - 1	1.0	1.0	1.0
CPC - 2	1.0	1.0	1.0
CPC - 3	0.0	0.0	0.0

Table 5: Probabilities associated with levels of significance of the randomization tests

Model	CPC 1	CPC 2
ECMWF	2.7	1.8
UGAMP	3.4	1.9
MPI	2.2	2.1
MPI- 2	2.2	2.3

Table 6: Mean absolute error of the CPCs for the four models shown. The error is with respect to the reanalysis data

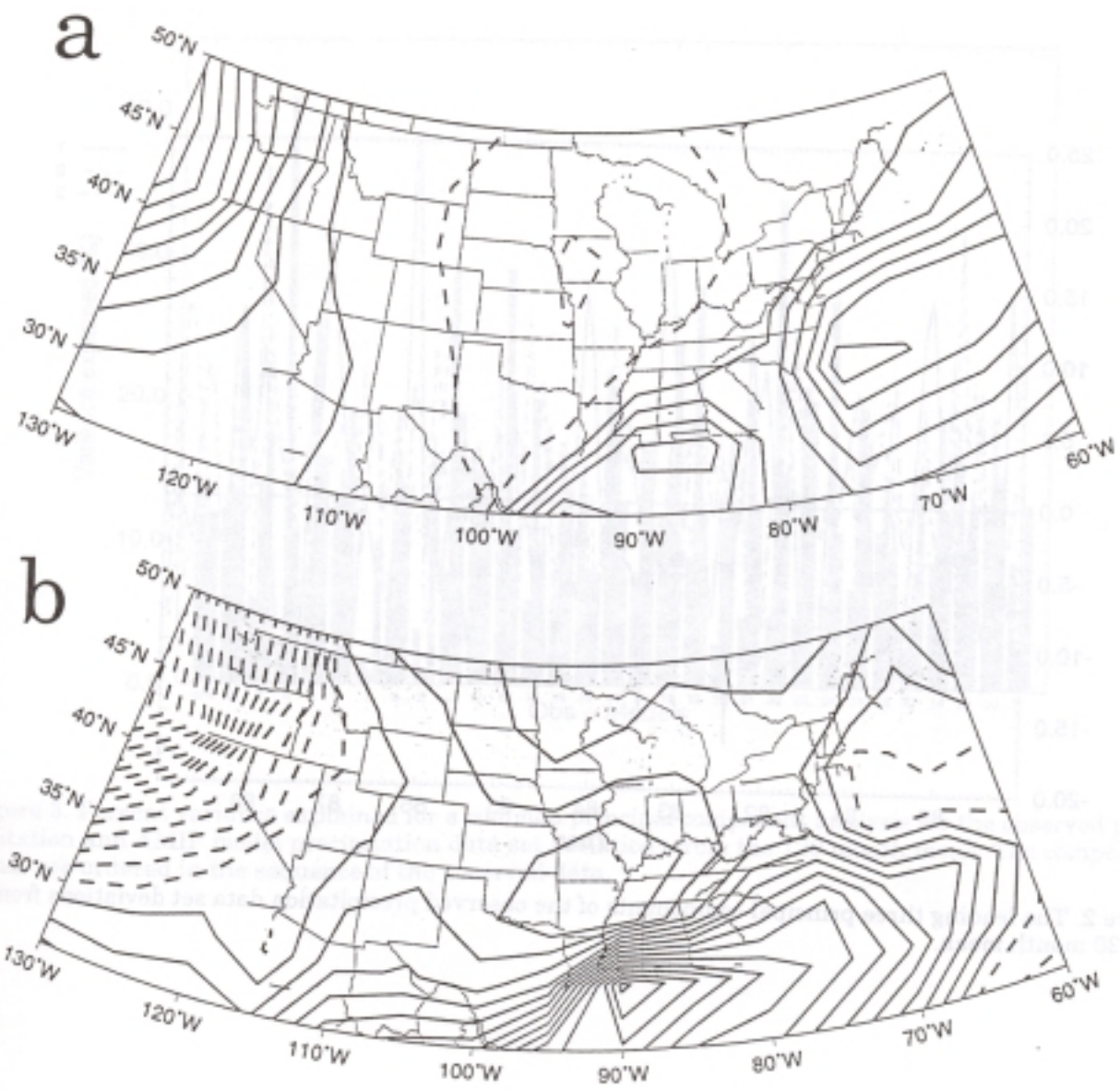

Figure 1. (a) The leading principal vector of the observed precipitation data set deviations from the 120 month mean. This mode explains 25% of the variance. Contour interval is 0.01. Dashed lines indicate negative values. Solid lines indicate zero and positive values.
(b) As in (a) except for the second principal vector. This explains 14% of the variance.

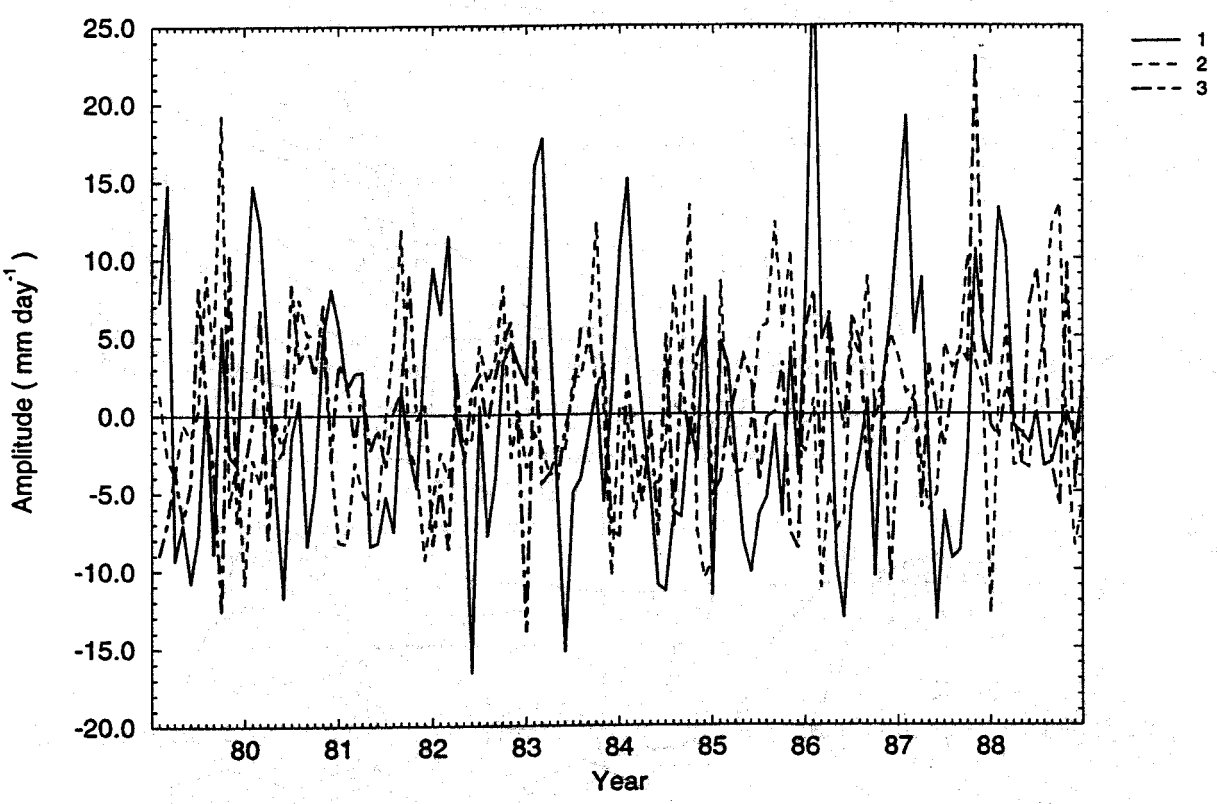

Figure 2. The leading three principal components of the observed precipitation data set deviations from the 120 month mean.

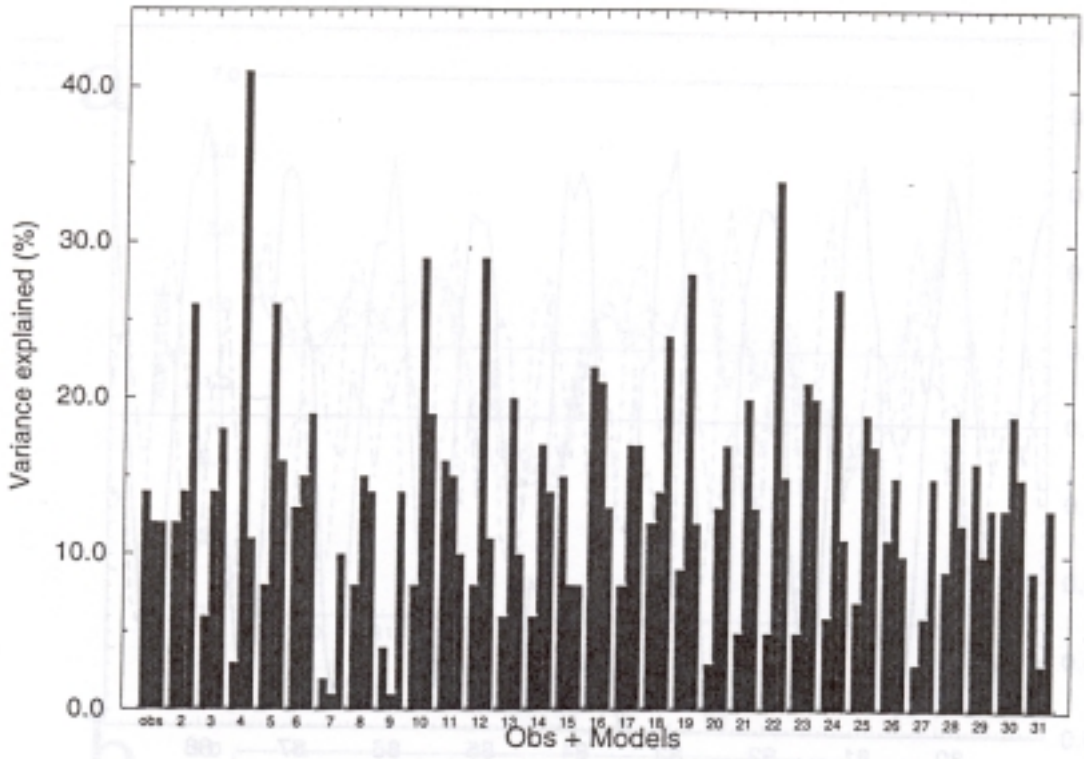

Figure 3. Percent variance explained for a common principal component analysis for the observed precipitation and AMIP model precipitation data set deviations from the 120 month mean. The components are ordered in the sequence of the observed data.

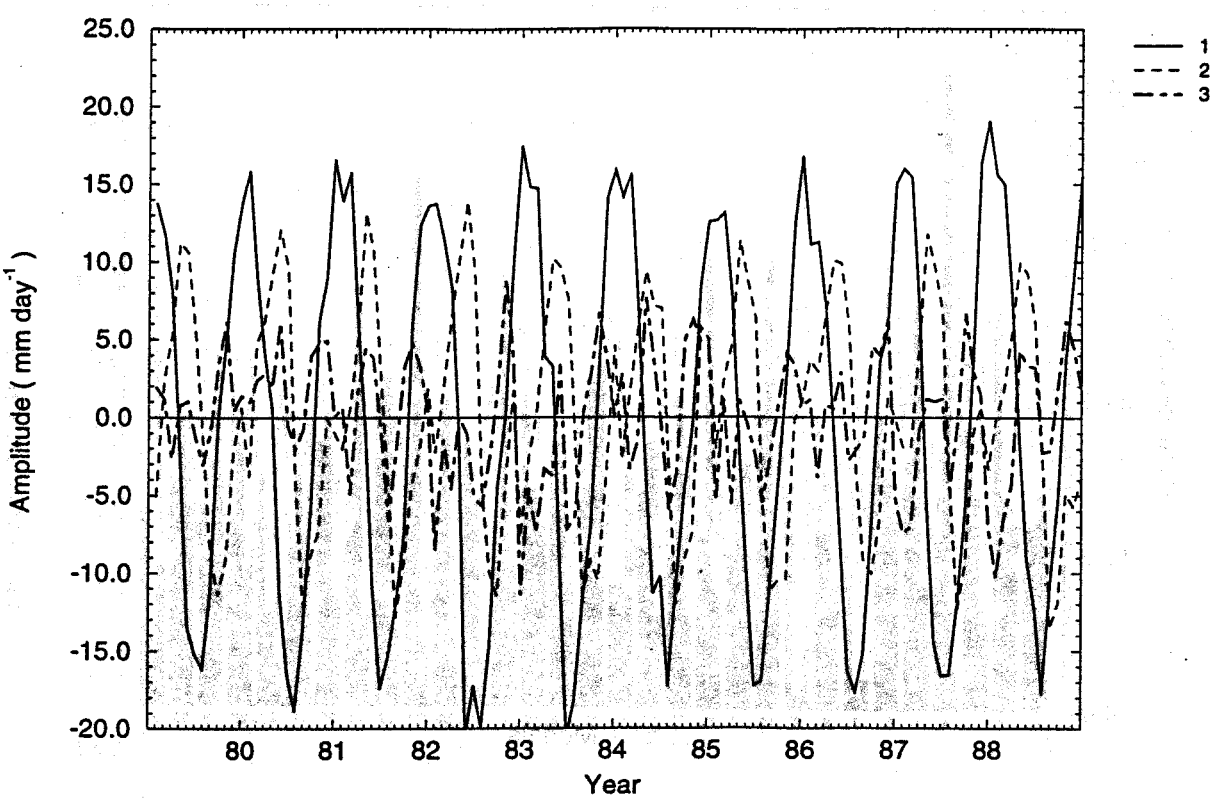

Figure 4. The leading three principal components of the UCLA AMIP precipitation data set deviations from the 120 month mean.

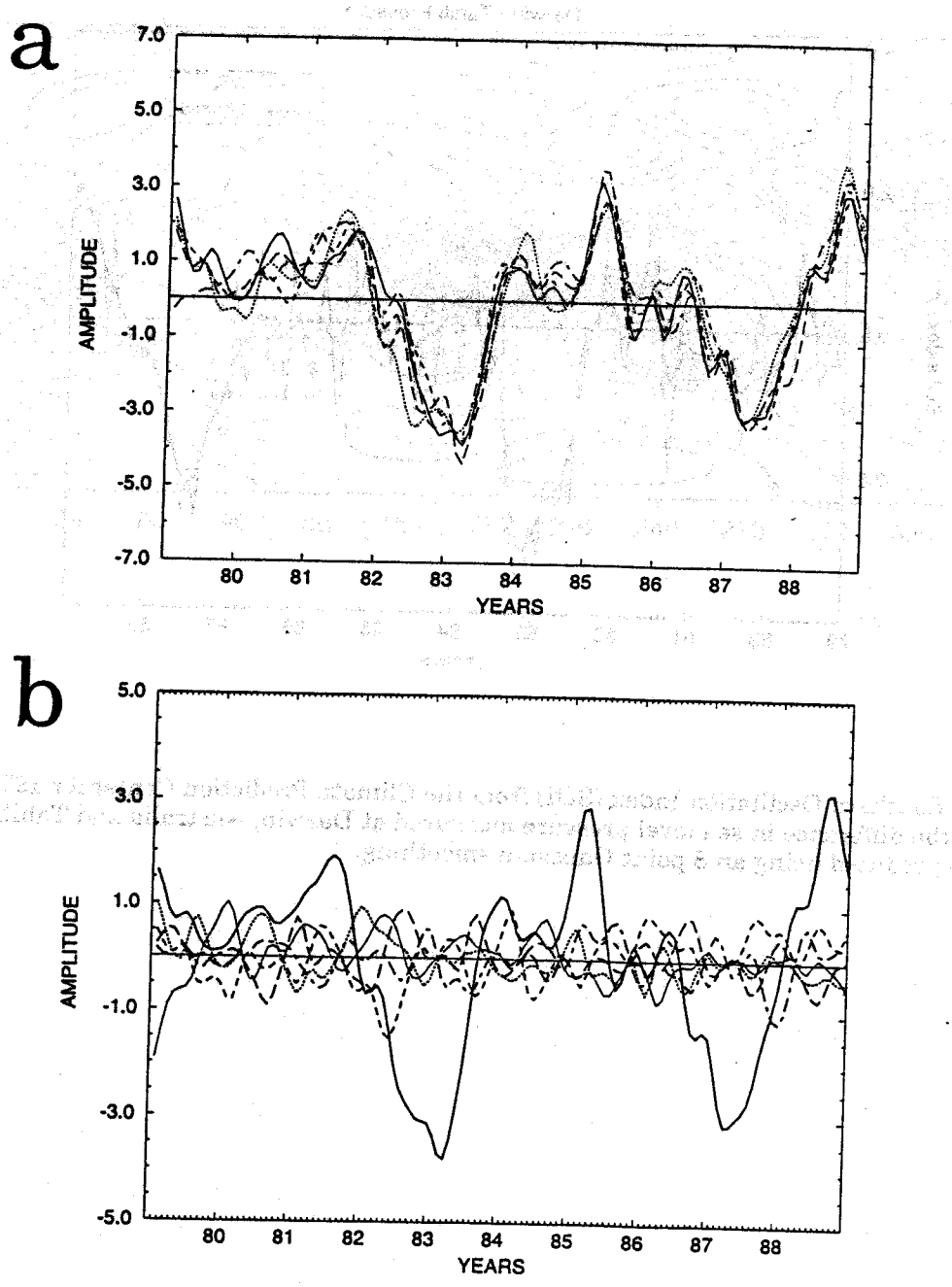

Figure 5. (a) The leading common principal component for each of the five members of the ECMWF AMIP ensemble of simulations for the 200 hPa interannual variations of the velocity potential. (b) The same data as in (b) except the mean (thick solid line) of the five members is shown along with the deviations of each simulation from this mean.

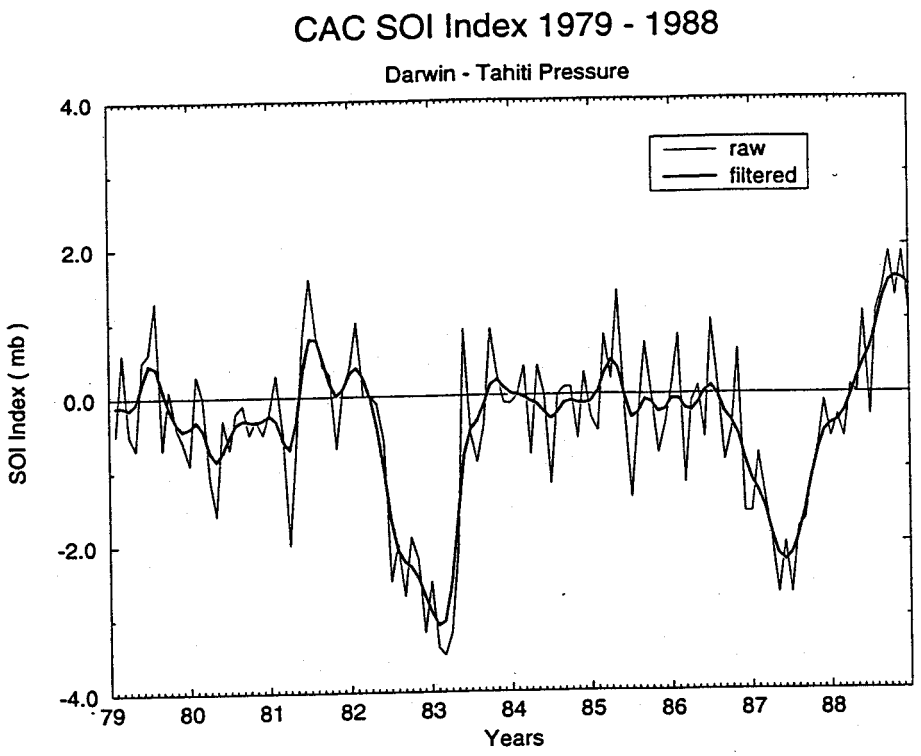

Figure 6. The Southern Oscillation Index (SOI) from the Climate Prediction Center for 1979 to 1988. This index is the difference in sea level pressure measured at Darwin, Australia and Tahiti. The filtered curve is produced using an 8 point Gaussian smoothing.

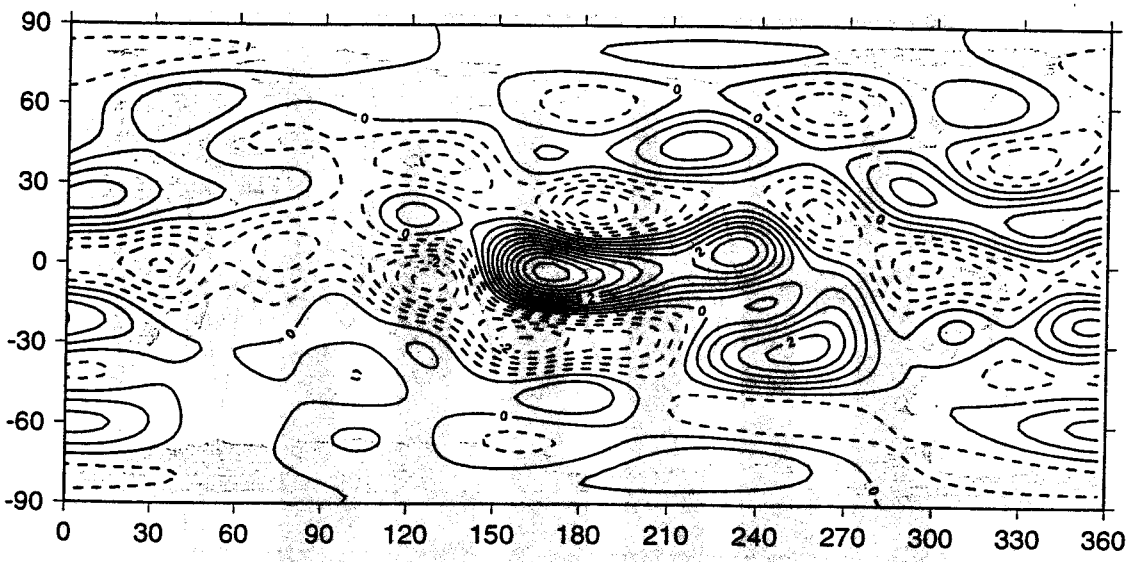

Figure. 7. The leading common principal vector of the five ECMWF AMIP ensembles for the 200 hPa interannual variations of velocity potential. Contour interval is 0.01. The dashed lines indicate negative values. The solid lines are positive and zero.

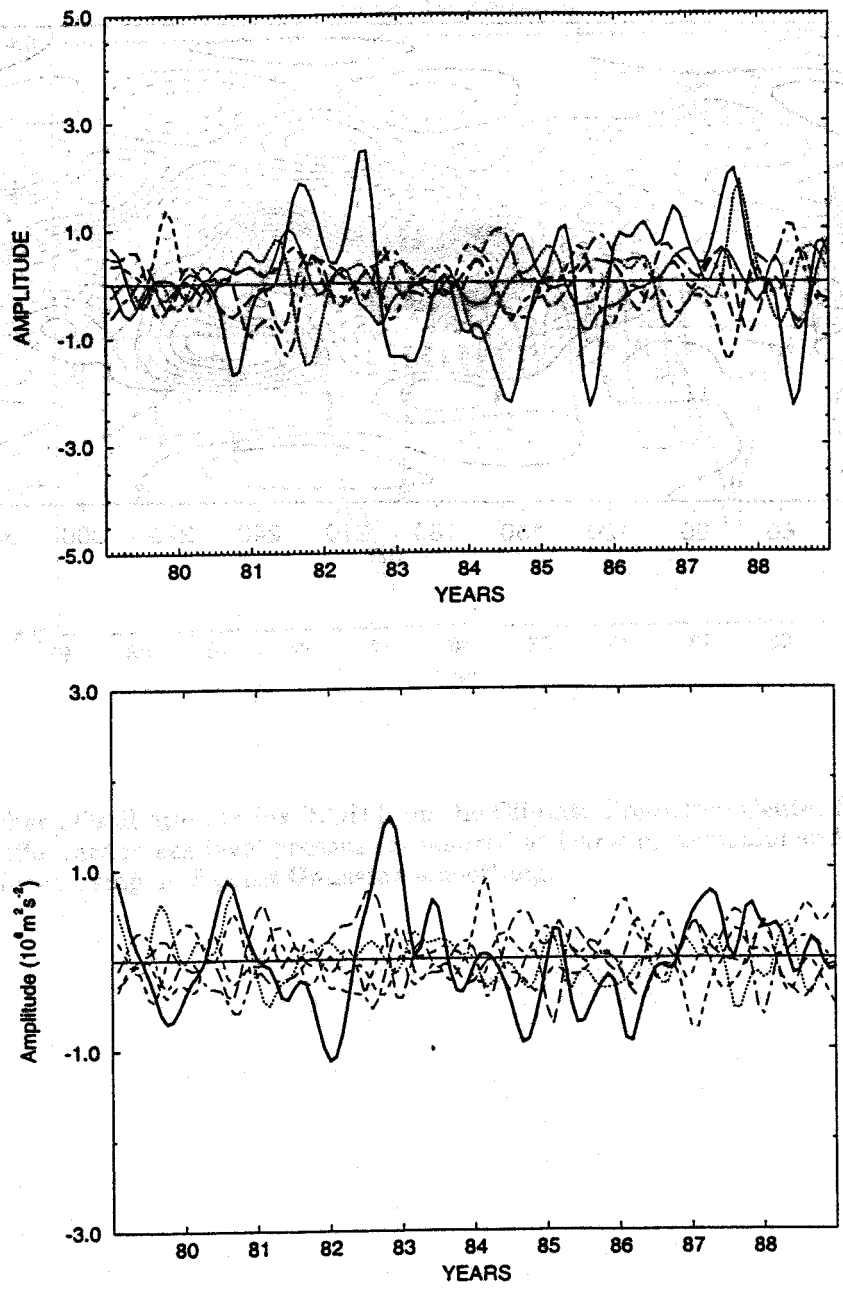

Figure 8. (a) As in 5b except for the second component.
(b) As in 5b except for the third component. Note the scale change on the ordinate.

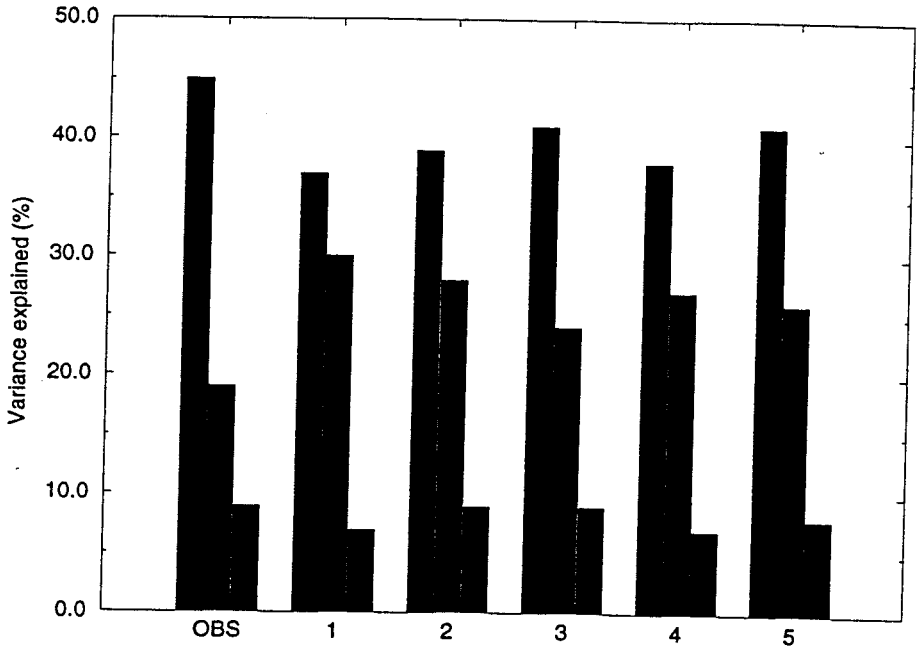

Figure 9. Percent variance explained for the leading three common principal components of the observed (NCEP reanalysis) and five ECMWF AMIP simulations for the 200 hPa velocity potential.

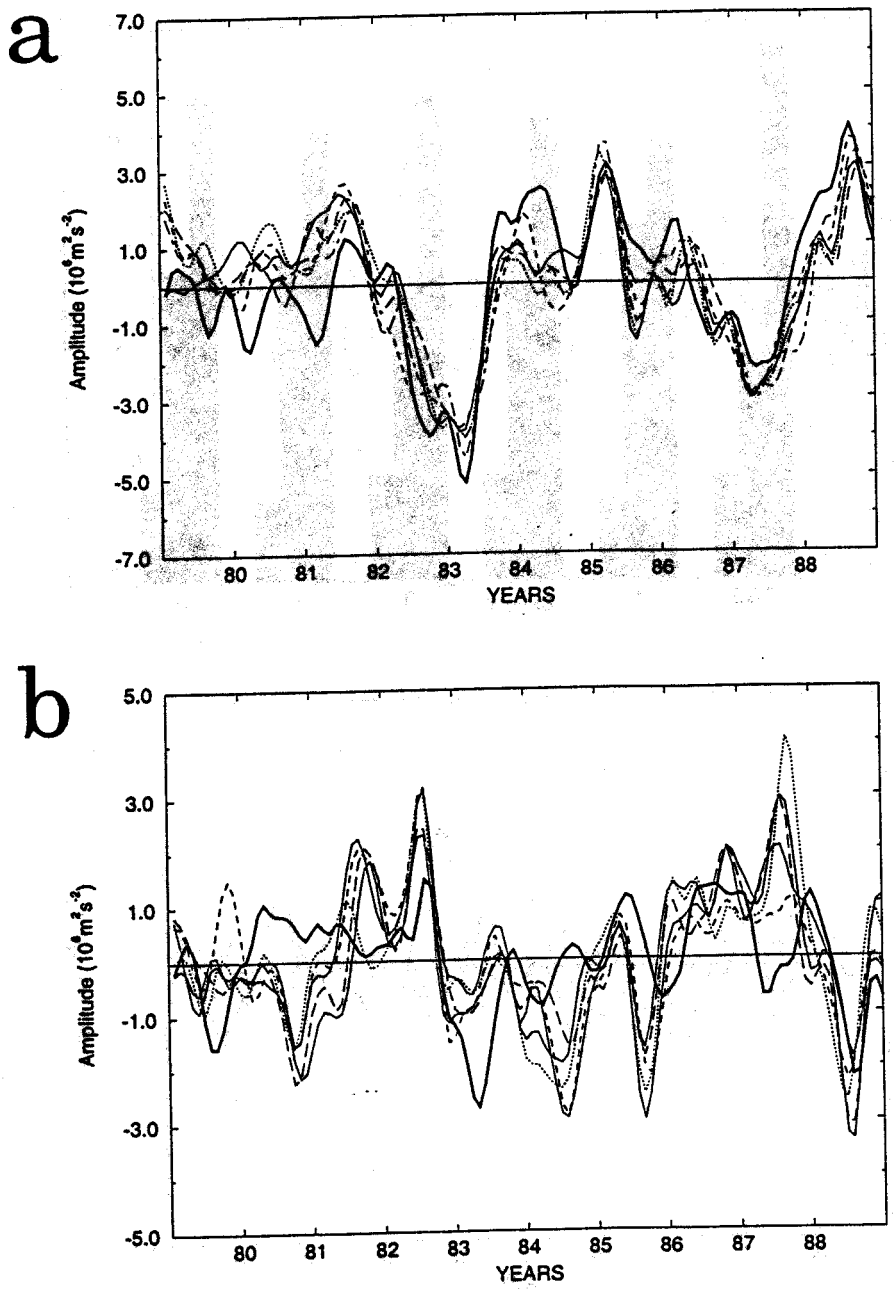

Figure 10. (a) The leading common principal component of the analysis with the NCEP reanalysis and the five members of the ECMWF AMIP ensemble of simulations for the 200 hPa interannual variations of the velocity potential. The thick solid line is the reanalysis data.
(b) As in (a) except for the second component.

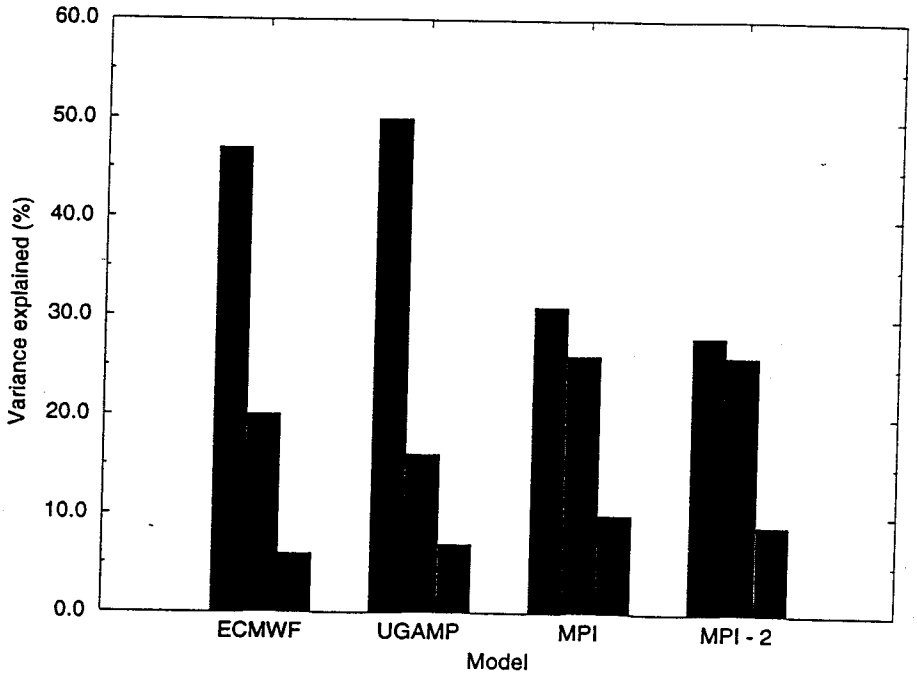

Figure 11. Percent variance explained for a common principal component analysis for the difference in the global 200 hPa velocity potential of four AMIP models and the NCEP/NCAR reanalyses. The components are ordered in the sequence of the ECMWF model.

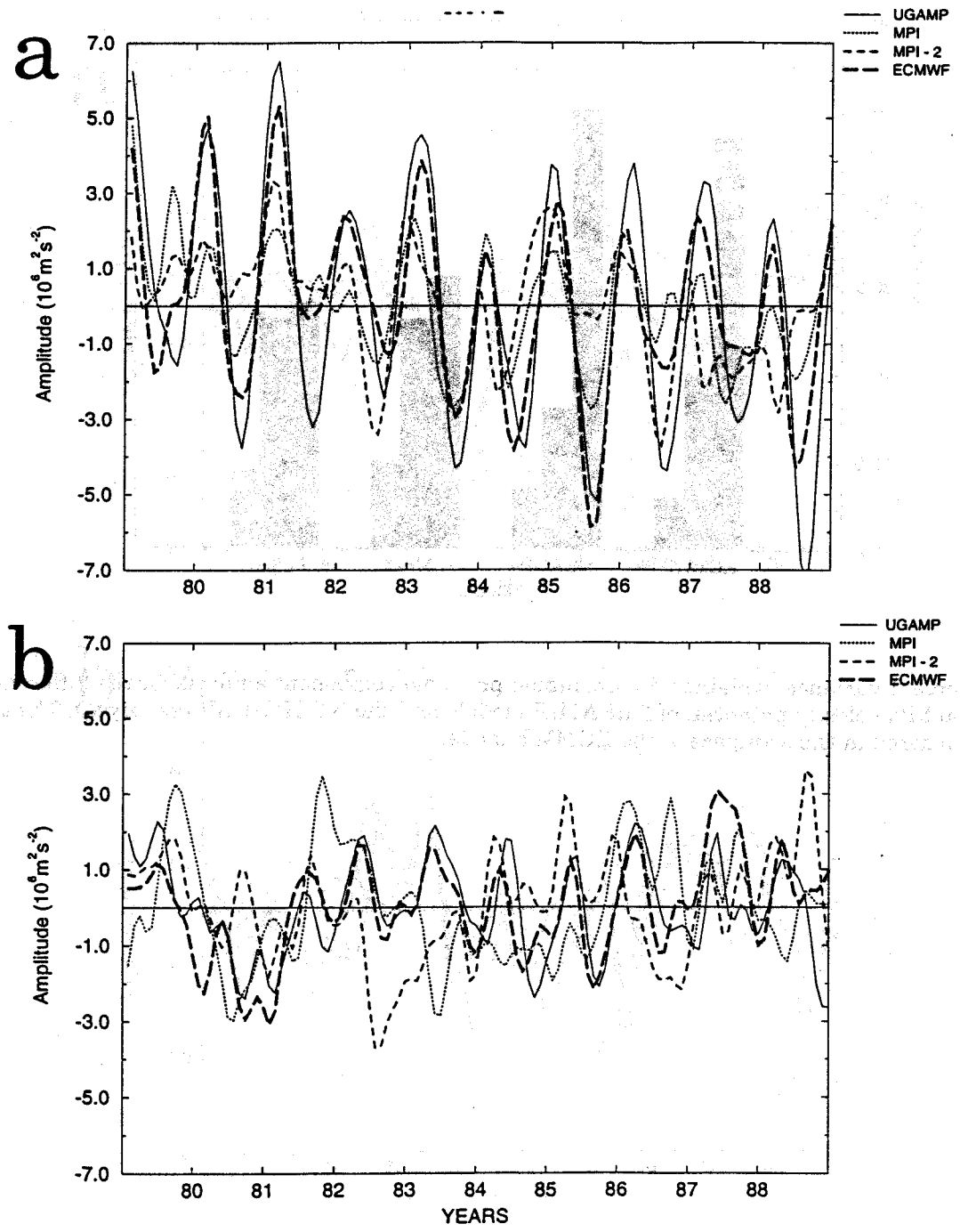

Figure 12. (a) The leading common principal component for each of the four models for the difference in the global 200 hPa velocity potential of each model from the NCEP/NCAR reanalyses.
(b) As in (a) except for the second CPC component.

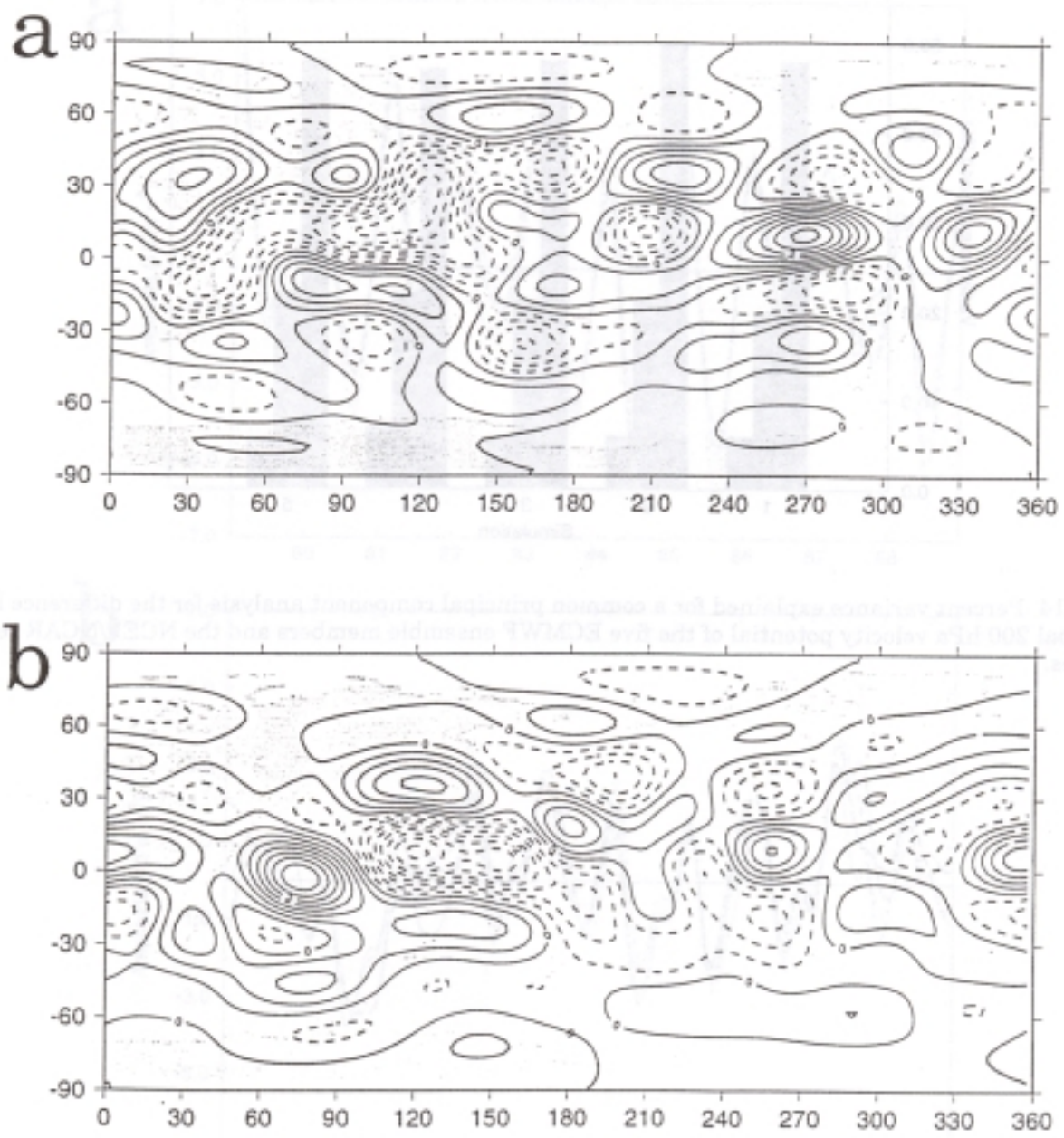

Figure 13 (a)The leading common principal vector of the four models for the difference in the global 200 hPa velocity potential of each model from the NCEP/NCAR reanalyses 200 hPa interannual variations of velocity potential. Contour interval is 0.01. The dashed lines indicate negative values. The solid lines are positive and zero.
(b) As in (a) except for the second CPC vector.

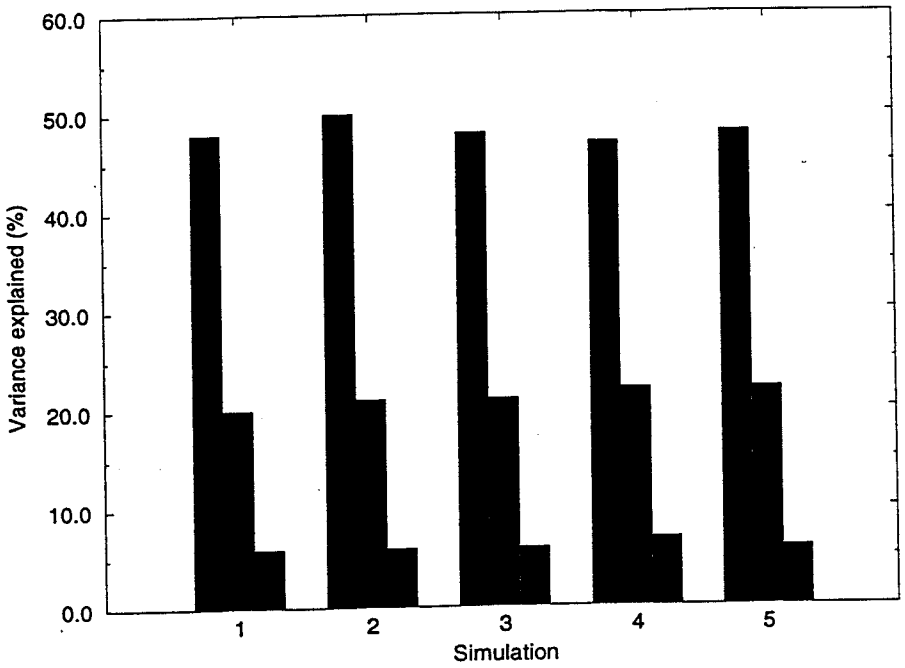

Figure 14. Percent variance explained for a common principal component analysis for the difference in the global 200 hPa velocity potential of the five ECMWF ensemble members and the NCEP/NCAR re-analyses.

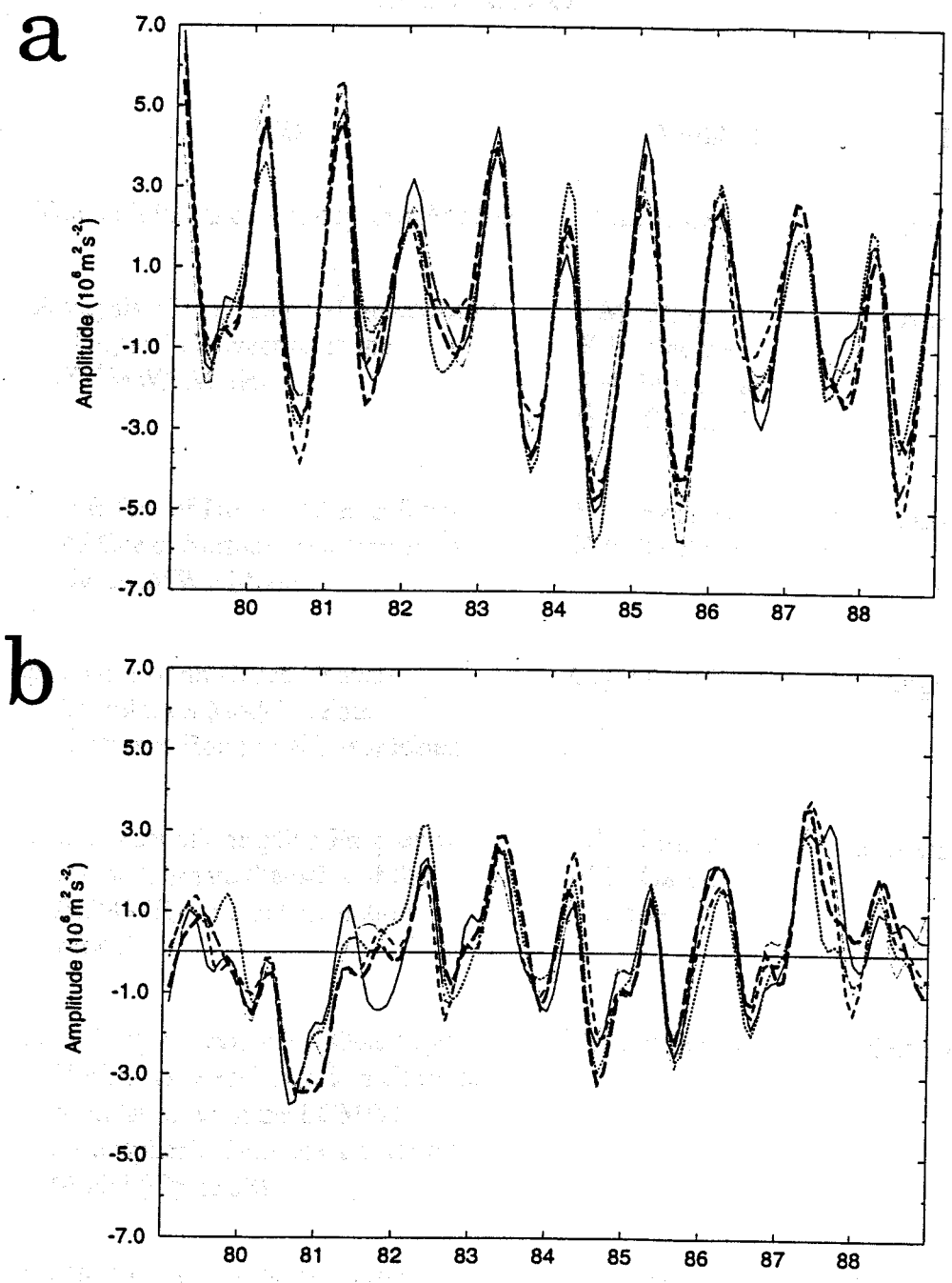

Figure 15 (a) The leading common principal component for each of the five ECMWF ensemble members for the difference in the global 200 hPa velocity potential of each model from the NCEP/NCAR reanalyses.
(b) As in (a) except for the second CPC component.