A Statistical Comparison with Observations of Control and El Niño Simulations Using the NCAR CCM*

Hans von Storch

Max Planck Institut für Meteorologie, Bundesstraße 55, 2000 Hamburg 13, FR Germany

(Manuscript received 26.01.1987, in revised form 29.04.1987)

Abstract:

Extended simulations, done with the NCAR Climate Community Model (CCM) in the perpetual January mode, are studied in terms of Northern hemisphere 500 mb height with respect to two questions:

(a) To what extent does the NCAR CCM reproduce the observed characteristics of the January mean flow?

(b) Is the extratropical response of the NCAR CCM to the doubled RASMUSSON/CARPENTER standard El Niño SST anomaly similar to any of the 500 mb height anomalies observed during the last few El Niño events, or similar to the response of the ECMWF T21 GCM to the same SST anomaly?

These problems are considered by means of a statistical procedure proposed by STORCH et al. (1985), consisting of the steps "reduction of degrees of freedom", "performance of a multivariate statistical test", and "a-posterioi univariate analysis". This basic version is used to compare objectively the model's climatology with that of the observed one: The intensity and phase of stationary and transient eddies is fairly realistically simulated, but the midlatitudinal January zonal mean gradient of 500 mb height is significantly too steep.

In the modified version, the reduction of degrees of freedom is achieved by projecting the GCM data onto so-called "guess patterns" which were derived a priori from theoretical or empirical considerations. Advantageously, the results of this procedure demonstrate not only the existence of a nonzero signal but also its similarity with the guess patterns. Considering the NCAR CCM response to the El Niño SST anomaly, we selected the Northern hemisphere 500 mb height anomalies observed during Januaries 1973, 1977 and 1983 as guess patterns and also the ECMWF T21 GCM response to the same anomalous SST. The result of the analyses is that the NCAR CCM response is significantly nonzero, parallel to the January 1977 anomaly, antiparallel to the 1973 anomaly and perpendicular to the 1983 anomaly. It is negatively correlated with the ECMWF T21 GCM response.

Zusammenfassung: Statistische Analyse eines Kontroll- und eines El Niño Response-Experimentes mit dem NCAR CCM

Januar-Simulationsdaten des NCAR Climate Community Model (CCM) werden mit statistischen Methoden daraufhin untersucht, ob die Modell-Klimatologie sich in statistisch signifikanter Weise von der beobachteten Januar-Klimatologie unterscheidet. Ferner wird untersucht, ob die extratropische Antwort des NCAR CCM auf eine El Niño SST Anomalie

(a) statistisch signifikant von Null verschieden ist,

(b) in einem objektiven Sinne ähnlich ist den beobachteten extratropischen Anomalien während der letzten El Niño Ereignisse, 1972/73, 1976/77 und 1982/83, und

(c) in einem objektiven Sinne ähnlich ist dem statistisch signifikanten Reponse des ECMWF T21 GCM auf gleiche SST Anomalie.

* Dedicated to Prof. Dr. H. Flohn on his 75th birthday

0005-8173/87/04 0464–14 $ 03.00/0

Betrachtet wird das nordhemisphärische, extratropische Feld der 500 mb Topographie.

Das dabei eingesetzte statistische Verfahren ist im wesentlichen der u.a. von STORCH et al. (1985) vorgeschlagene Algorithmus „Reduktion der Freiheitsgrade", „Anwendung eines statistischen, multivariaten Tests", „a-posteriori univariate Analyse". Die Grundform des Verfahrens wird eingesetzt, um die Modell- und die beobachteten Klimatologien objektiv zu vergleichen: Demnach ist das NCAR Zirkulationsmodell recht erfolgreich in der Reproduktion der Intensität und der Lage der stationären und transienten Eddies, aber das zonal-zeitliche Mittel des 500 mb Geopotentials hat einen signifikant zu steilen meridionalen Gradienten in mittleren Breiten.

In der modifizierten Form des Verfahrens werden „guess patterns" zur Reduktion der Freiheitsgrade eingesetzt: Wenn Muster zur Verfügung stehen, denen der Modell Reponse aufgrund einer a priori Überlegung ähneln sollte, so werden die Modelldaten auf diese „guess patterns" projeziert. In dem so entstehenden niederdimensionalen Raum wird dann der statistische Test durchgeführt. Vorteil dieses Verfahrens ist, daß man nicht nur eine Aussage über die Existenz eines von Null verschiedenen Signals bekommt sondern auch eine objektive Aussage über die Ähnlichkeit des untersuchten und eines a-priori ausgewählten Musters. Um den extratropischen NCAR CCM Reponse auf die doppelte Rasmusson-Carpenter Standard El Niño SST Anomalie zu untersuchen, werden als „guess pattern" die beobachteten Anomalien der Januare 1973, 1977 und 1983 sowie der ECMWF T21 GCM Response verwendet. Als Resultat ergibt sich zunächst die statistische Signifikanz des simulierten Musters sowie eine überzufällige Ähnlichkeit, d.h. eine signifikant positive Projektion, mit dem Januar 1977-Muster. Das NCAR CCM-Signal ist aber antiparallel zum Januar 1973-Muster und steht senkrecht auf dem Januar 1983-Muster. Das ECMWF T21 GCM Response auf die gleiche El Niño Anomalie ist mit dem des NCAR CCM global negativ korreliert.

Résumé: Analyse statistique d'expériences de contrôle ou de simulation du phénomène «El Niño» par le CCM du NCAR

On étudie en termes de la hauteur de la surface de 500 mb de l'hémisphère Nord des simulations réalisées avec le Climate Community Model (CCM) du NCAR pour un mois de janvier perpétuel. Deux problèmes sont envisagés:

a) jusqu'à quel degré le modèle reproduit-il les caractéristiques observées de l'écoulement moyen de janvier?

b) La réponse du modèle, aux latitudes extratropicales, au doublement de l'anomalie standard de RASMUSSON/CARPENTER de la température de la mer caractérisant le phénomène «El Niño» est elle semblable à l'une quelconque des anomalies de la hauteur de la surface de 500 mb observées lors des dernières manifestiations du «Niño» ou semblable à la réponse du modèle de circulation généale T21 du ECMWF?

On étudie ces problèmes à l'aide de procédés statistiques proposés par STORCH et al. (1985) et qui comportent les étapes suivantes: «réduction du degré de liberté», «performance d'un test statistique multivarié» et «analyse univariée a posteriori». La version de base est utilisée pour comparer objectivement la climatologie du modèle avec celle observée. L'intensité et la phase des perturbations stationnaires et mobiles sont simulées de façon tout à fait réaliste mais le gradient zonal moyen de janvier de l'altitude de la surface de 500 mb aux latitudes moyennes est trop fort.

Dans le version modifiée, la réduction des degrés de liberté se réalise en projetant les données du modèle de circulation générale sur des «configurations estimées» déduites a priori de considérations théoriques ou empiriques. Ce procédé montre, de manière avantageuse, non seulement l'existence d'un signal non nul mais aussi sa similitude avec les configurations estimées. En considérant la réponse du modèle du NCAR aux anomalies de température de surface du «Niño», nous avons sélectionné pour les configurations estimées, les anomalies d'altitude de la surface de 500 mb de l'hémisphère Nord observées durant les mois de janier de 1973, 1977 et 1983 ainsi que la réponse du modèle T21 du ECMWF aux mêmes anomalies de température. L'analyse montre que la réponse du modèle du NCAR est significativement non nulle, parallèle à l'anomalie de janvier 1977, antiprallèle à celle de 1973 et perpendiculair à celle de 1983. Elle est corrélée négativement à la réponse du modèle du ECMWF.

1 Introduction

The significance of multicomponent signals is today often assessed initially by a series of univariate tests in the grid point space and subsequently by the performance of a statistical test which makes use of the rate of univariate rejections of the "local null hypothesis of no signal" (STORCH, 1983; LIVEZEY and CHEN, 1983).

Another procedure consists of an a-priori reduction of degrees of freedom of the multicomponent signal, a multivariate test and eventually of a final univariate analysis. The reduction of degrees of freedom is achieved by projecting the original data onto a subspace spanned by a few, a-priori selected "guess patterns". This concept was proposed by HASSELMANN (1979) and shown to be useful mainly by the work of HANNOSCHÖCK and myself (for an overview, see LIVEZEY, 1985). It is described in some detail by STORCH and KRUSE (1985) and will be summarized in Section 2 of this paper.

In a **basic version** of the procedure, the guess patterns do **not** depend on the problem under consideration (Section 2). How this basic version may be applied successfully to midlatitude 500 mb height topography data simulated by the NCAR Climate Community Model ("CCM"; Section 3) is shown in Sections 4 and 5: First, some circulation statistics derived from an extended range "control run" are compared with respective circulation statistics based on DWD analyses of the Januaries 1967 to 1984 in order to assess whether the simulated flow is systematically different from the real flow. Secondly, the 500 mb height topography as simulated in the above mentioned control experiment is compared with data derived from an extended range El Niño type SST anomaly experiment with the NCAR CCM.

In Section 6, a **particular version** of the general concept is discussed. It makes use of **one** problem-dependent guess pattern derived from similar but independent observations of simulations. This approach seems rather robust and simple: The projection into a one-dimensional subspace allows any univariate test to be applied. It permits not only a conclusion concerning the significance, i.e. the stability of the response, but also an assessment as to whether the experimental data contain a signal similar to the prescribed guess pattern. That part of the signal which is perpendicular to the guess pattern may be subjected to the general technique outlined above using problem-independent guesses.

This particular version is used in order to study whether the NCAR CCM's response to El Niño SST conditions is similar to the 500 mb height topography anomaly observed in the El Niño Januaries 1973, 1977 and 1983 or to the response pattern simulated by the ECMWF T21 GCM (Section 7).

2 The Basic Version of the Significance Analysis Strategy

The basic version of the significance analysis used in the study was presented by STORCH and ROECKNER (1983) and STORCH et al. (1985): In its first step, the data are projected onto some a priori "guess patterns", which do not depend on the considered problem. The guess patterns are orthonormal functions (to be precise: discrete functions = vectors) allowing for a fast-converging series expansion. Since only the first few of the expansion coefficients are connected with significant variance, only the first few need to be retained.

The particular form of the orthonormal functions is not relevant in this context, but only their ability to explain efficiently a major part of the variance of the considered multicomponent random variable. For that reason, these function are not shown or discussed. In this study, EOF's are used.

After the dimension of the problem has been reduced, a multivariate test is done to assess whether the null hypothesis of a zero signal conflicts with the data. If so, the test leads to the conclusion that the probability of a zero signal is less than, say, 5 %. This result provides statistical proof of a nonzero signal, but it is no more than evidence of its **existence.** Nothing is said about its attributes, about its distribution or its strength in point space.

The following consideration shows that one should not interpret each detail of a statistically significant mean signal as "significant", i.e. as sample-independent property of the true nonzero signal: Consider a two-dimensional random vector (A, B) with $E(A) \neq 0$ and $E(B) = 0$ (E denotes expectation). If the "signal" E(A) is sufficiently strong, the significance analysis applied to the vector (A, B) will assess $E((A, B)) \neq 0$. Due to sampling, the sample mean $\overline{B}$ is nonzero from which $E(B) \neq 0$ may not be inferred.

Thus, we suggest that the multivariate tests should be followed by an a posteriori analysis where (in the grid point space or in some spectral space) the experimental samples are to exhibit patterns systematically different from the control samples. Those areas have to be identified in which all or nearly all experimental states have larger (or smaller) values than the mean of the control states, or – even better – than all or almost all of the control states. This approach has been formalized successfully to the "recurrence analysis" concept (STORCH and ZWIERS, 1987), evaluating the probability of such events.

This analysis may be done by estimating roughly a 95 % band of the control ensemble; that is, for each grid point an interval containing about, say, 95 % of all control states. Together with the 95 % band all individual experimental states are plotted. In this manner, one may easily demonstrate where the experimental samples reveal structures recurrently different from the control samples. This procedure and the resulting diagrams are used throughout this study.

3 NCAR CCM Experiments and Available Data

With the NCAR Climate Community Model (CCM), two experiments were performed in the perpetual January mode, each integrated over a spin-up period of 200 days and an experimental range of a total of 1200 days. The two experiments differ with respect to the prescribed SST distribution in the equatorial Pacific. The "control experiment" was run with climatologically fixed SST. In the second experiment, a doubled RASMUSSON and CARPENTER (1981) standard January El Niño anomaly centered at the equatorial eastern and central Pacific was superimposed. A thorough description of the experiment and its results is given by BLACKMON et al. (1983).

In the present study, the experimental range of 1200 days is subdivided into a series of 30 Januaries in the following way: First, the time series is broken into adjacent 40 day intervals. Secondly, the first 10 days of each subinterval are disregarded and the remaining 30 days considered to form one "January". In this way, a total of 30 control Januaries and 30 "El Niño" Januaries is obtained.

Because of the serial correlation of the data considered, the 10 day gap between each January is not sufficient to obtain strictly statistically independent monthly means. We do not believe that this is a serious limitation of our analysis even if we cannot rule out completely the possibility that our tests underestimate their risks. However, in all parameters, which appear in the following sections to be statistically significant, we find that the risks are so small that a slight underestimation of this quantity would not harm the results of the analysis.

This study is restricted to the extratropical Northern Hemisphere flow characterized by the 500 mb height topography.

4 Verification of NCAR CCM 500 MB Height Climatology

The basic strategy presented in Section 2 is used to compare several Northern Hemisphere 500 mb height statistics derived from the NCAR CCM control experiment and from DWD analyses of Januaries 1967 to 1984. Results obtained for the ECMWF T21 and T40 GCM and for the Hamburg University Model are given by STORCH et al. (1985).

The first step of the strategy – the reduction of degrees of freedom – is achieved by expanding the data into an EOF series and using only the first 5 EOF coefficients (details: see STORCH and ROECKNER, 1983). For the multivariate test – the second step – we use the generalized Mann/Whitney test (details: see STORCH et al., 1985). All simulated parameters considered in the next two paragraphs turn out to be significantly different from the analyzed data. All test statistics obtained in the various cases were associated with risks less than 1 %. The univariate analysis is carried out as outlined in Section 2 by plotting an estimated 95 % band of the observed state and all individual simulated states.

In Figure 1, the 95 % band of the zonal January mean, $[\bar{z}]$, estimated from the 18 DWD-analysed Januaries and the 30 individual CCM-generated states are shown. The CCM midlatitude gradient of $[\bar{z}]$ is much too steep, which is connected with an overly intense zonal flow in the middle troposphere. The intensity of stationary eddies, $[\bar{z}^{*2}]\cos\phi$, is somewhat too weak and shifted northward (Figure 2). The 30°–60 °N mean of stationary (January mean) eddies, $\{\bar{z}^{*}\}$, exhibits a clearly developed Pacific stationary disturbance, but the Atlantic system is rather weak compared to observation (Figure 3). The difference with respect to the stationary disturbance cannot be regarded as severe; the NCAR CCM, especially if compared with the performance of other GCM's, appears to lie near the upper bound of present-day GCM's ability to reproduce the actual climate (Figure 4).

According to the univariate analysis displayed in Figure 5, the transient eddy variance, $[\overline{z'^{*2}}]\cos\phi$, is well simulated for practical purposes, even if its overall level is simulated slightly too low. This is an example of a difference "simulation minus analysis", which is statistically significant but not physically relevant. A comparison with other GCMs shows that the NCAR CCM simulates the transients favorably (Figure 6).

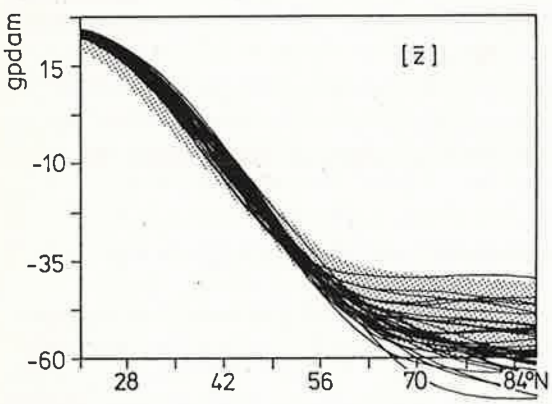


Figure 1 Comparison of Northern Hemisphere NCAR CCM 500 mb January height climatology with data. Univariate analysis of $[\bar{z}]$.

Dotted: 95 % band derived from DWD analyses of Januaries 1967–84.

Lines: Individual NCAR CCM generated states (permanent January).

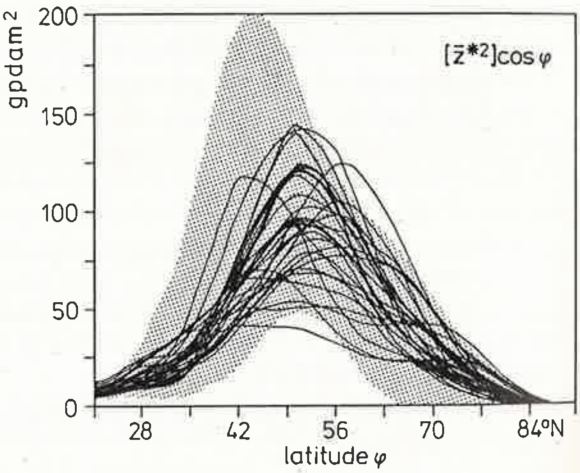
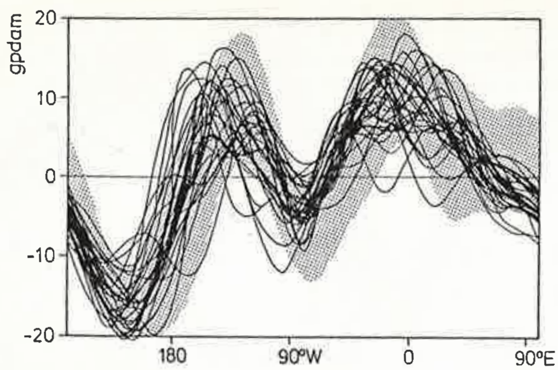


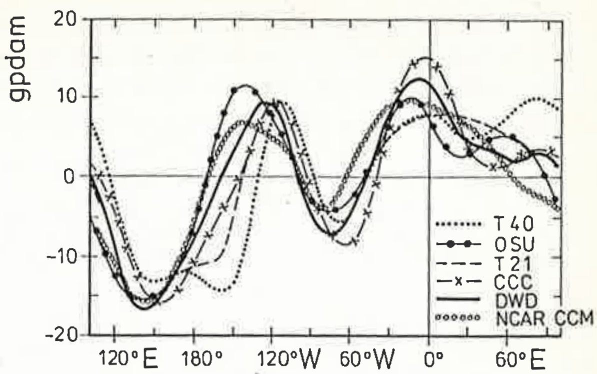
Figure 2 Comparison of Northern Hemisphere NCAR CCM 500 mb height climatology with data. Univariate analysis of zonally averaged variance due to stationary (January mean) eddies, $[\bar{z}^{*2}]\cos\phi$.

Dotted: 95 % band of observed data, derived from DWD analyses of Januaries 1967–84.

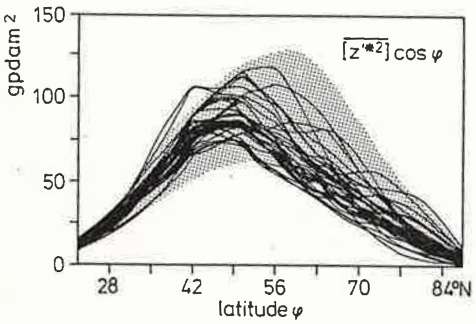
Lines: Individual NCAR CCM generated states (permanent January)



• **Figure 3** Comparison of Northern Hemisphere NCAR CCM 500 mb height climatology with data.
Univariate analysis of the 30° – 60 °N mean of stationary (January mean) disturbances $\{\bar{z}^*\}$.
Dotted: 95 % band derived from DWD analyses of Januaries 1967–84.
Lines: Individual NCAR CCM generated states (permanent January).



• **Figure 4** Comparison of Northern Hemisphere 500 mb January mean height climatology with GCM simulated climatologies.
Quantity: 30°– 60 °N mean of stationary disturbances $\{\bar{z}^*\}$.
DWD = DWD analyses 1967–85; T21 & T40 = ECMWF; CCC = Canadian Climate Centre; OSU = Oregon State University; NCAR = National Centre of Atmospheric Research.



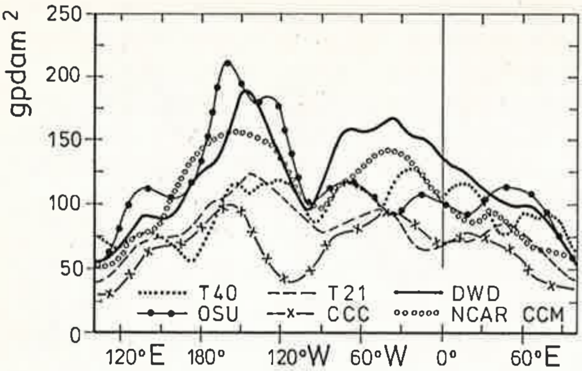
• **Figure 5**
Comparison of Northern Hemisphere NCAR CCM 500 mb height climatology with data. Univariate analysis of the zonally averaged variance due to transitient eddies, $[z'^{*2}]\cos(\phi)$.
Dotted: 95 % band derived from DWD analyses of Januaries 1967–84.
Lines: Individual NCAR CCM generated states.

5 El Niño Experiment Significance Analysis without A-priori Guesses

If no a priori guesses of the El Niño SST anomaly response pattern are available to perform the first step of the significance analysis strategy – the reduction of degrees of freedom – the same procedure as used in Section 4 may be utilized to compare the El Niño SST anomaly GCM experimental data and the GCM control data. All stationary flow and transient eddy circulation statistics considered in Section 4 were tested. Statistically significant signals were detected in the stationary eddy component only.

Figure 7 displays the univariate analysis of the response in terms of the 30°–60 °N mean $\{\bar{z}^*\}$. Apparently, one of the 30 experimental Januaries exhibits a longitudinal distribution totally different from the other 29. Apart from this exception a clear tendency towards a lowered 500 mb topography at the eastern Pacific and eastern Atlantic may be identified as a common property of all experimental samples. Also, the topography is raised over the American continent.

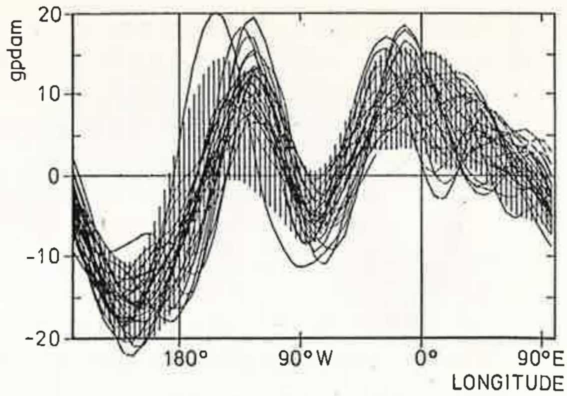
The zonally averaged stationary eddy variance, $[\bar{z}^*]\cos\phi$, is found to be significantly different in the El Niño experiment. The univariate analysis shows, however, that this signal although detectable, is small and therefore unimportant.



• **Figure 6** Comparison of Northern Hemisphere January 500 mb height climatology with GCM-simulated climatologies.

Quantity: 30°– 60 °N mean of variance due to transient eddies: $\{\overline{z'^2}\}$.

DWD = DWD analyses 1967–85; T21 & T40 = ECMWF; CCC = Canadian Climate Centre; OSU = Oregon State University; NCAR = National Centre of Atmospheric Research.



• **Figure 7** Comparison of Northern Hemisphere NCAR CCM 500 mb height data simulatee in the control and in the El Niño anomaly experiment. Univariate analysis of the 30° – 60°N mean of stationary disturbances, $\{\overline{z}^*\}$.

Shaded: 95 % band derived from 30 control run Januaries.

Lines: 30 individual Januaries from El Niño experiment.

6 The Particular Version of the Significance Analysis Strategy

If an a priori hypothesis about the pattern of the GCM response to anomalous SST is available, the basic version of the significance analysis scheme described in Section 2 may be modified. In the modified version, the complete signal S is split up into the component P, parallel to the a-priori fixed guess vector G, and the component Q, perpendicular to G:

$$S = P + Q$$

with P = (S, G) G and (Q, G) = 0. Here

$$(x, y) = \sum_i x_i y_i$$

is defined to be standard dot product. The guess vector G is assumed to be normalized, i.e. (G, G) = 1. The **parallel component** P varies within a one-dimensional subspace spanned by the guess vector G. Thus, the mean and variability of P are completely described by the mean and the variability of the generalized "Fourier-coefficient" f (S): = (S, G). Therefore, any standard univariate test may be used to assess whether the mean difference between the parallel component P simulated in the control and the anomaly experiment is significantly nonzero. In the following, the nonparametric Mann/Whitney test (e.g. CONOVER, 1971) is used. It is based on the relative order of the dot products of the individual control or experimental fields and the guess vector G.

The **perpendicular component** Q is given by

$$Q = S - f(S)G$$

It is situated in the (n − 1)-dimensional subspace perpendicular to G, i.e. (Q, G) = 0. Whether the data conflict with the null hypothesis that the mean perpendicular signal is zero, may be tested with the basic test version described in Section 2 and used in Section 5. If the perpendicular signal is found to be significantly nonzero, this may be taken as an indication that the guess pattern is capable of explaining only part of the complete signal.

7 El Niño Experiment Significance Analysis Using A-priori Guesses

7.1 Available Guesses

Useful guesses may be defined from similar but independent observations or simulations. In this study, three different guesses were considered:

(A) Since the NCAR CCM experiment yields as many as 30 January samples, it is possible to subdivide the total of 30 experimental samples into two subsets each consisting of every second sample. A guess pattern is determined as the mean difference between all control samples and the experimental samples of the first subsample. The second subsample is used to perform the test.

(B) The 500 mb height topography anomalies observed during the Januaries (mature phase; year "+1") of the last few strong El Niño events, i.e. Januaries 1973, 1977 and 1983. The SST anomalies of these El Niño episodes developed quite differently and were associated with different patterns, as was outlined by FU et al. (1986). The height anomalies were derived from the data already used for verifying the NCAR CCM in Section 4, namely from daily DWD analyses.

(C) The 500 mb height topography anomaly simulated in an extended range GCM experiment by CUBASCH (1985) with the T21 GCM of ECMWF. As in the experiment described in Section 3, Cubasch performed a series of "normal" January simulations and of "El Niño disturbed" January simulations. The strength and pattern of the superimposed equatorial Pacific SST anomaly is identical to that used by BLACKMON and his colleagues in the NCAR CCM experiment discussed here. This ECMWF T21 GCM generated response pattern was found to be significantly nonzero and, furthermore, highly coherent with the January 1983 anomaly pattern (STORCH and KRUSE, 1985).

7.2 Results

A NCAR CCM Experiment Guess

As was to be expected a-priori, the guess pattern G derived from every second experimental sample yields a parallel component with maximum significance (risk < 1 %). Figure 8 shows the complete signal and the guess pattern calculated from half of the data. Both maps are quite similar with respect to pattern and magnitude.

As an illustration of the use of the guess pattern as a predictive pattern, the number of control and experimental samples with a positive dot product (S, G) was obtained. The probability of a positive f(S) = (S, G) is only 43 % (close to expectation) in the control run but 97 % in the El Niño SST anomaly experiment.

The signal perpendicular to the guess pattern was not found to be significantly nonzero. According to the univariate analysis (not shown), the acceptance of a zero perpendicular signal appears reasonable: The "anomaly curves" vary irregularly mainly within the 95 % control band. The mean perpendicular signal Q has a magnitude one order less than both the complete and the parallel signal.

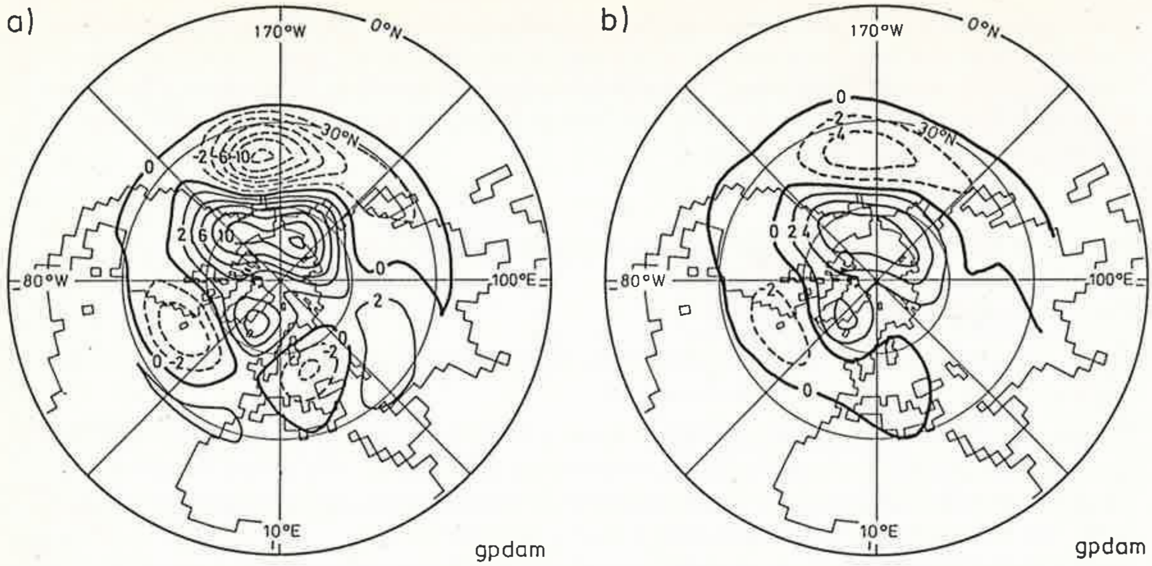


Figure 8 Northern hemisphere extratropical signal "experimental minus control data" of NCAR CCM El Niño experiment in terms of 500 mb height:

a) Complete signal.

b) Projection of the signal derived from half of the experimental data on the guess pattern build up from the other half of the experimental data ("parallel signal").

Spacing: 2 gpdam.

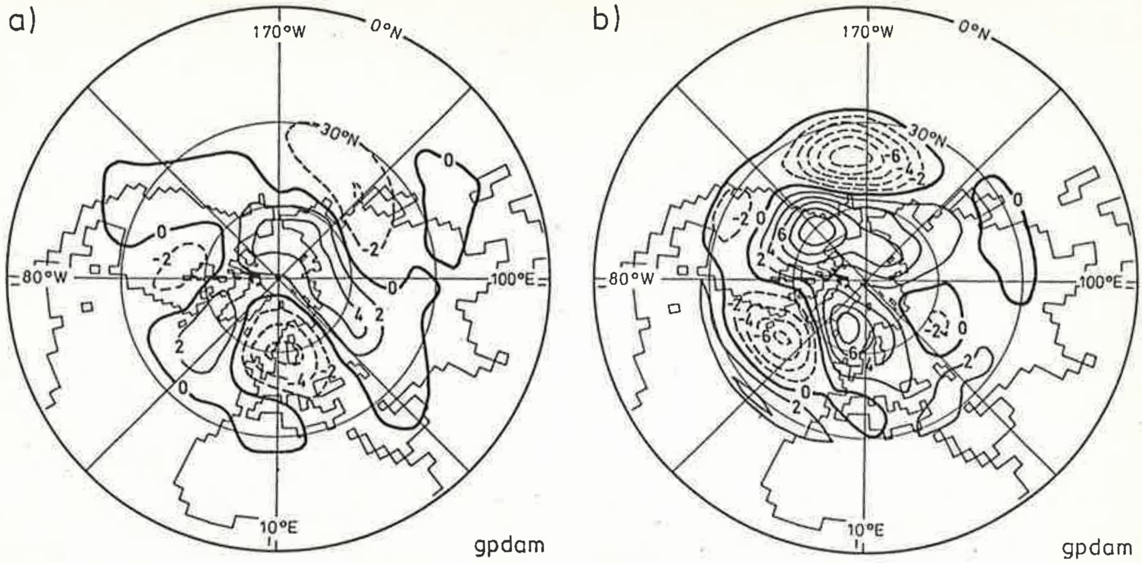
B Observed Guesses

The results obtained with the observed guess patterns January 1973, 1977 and 1983 are interesting and partly unexpected.

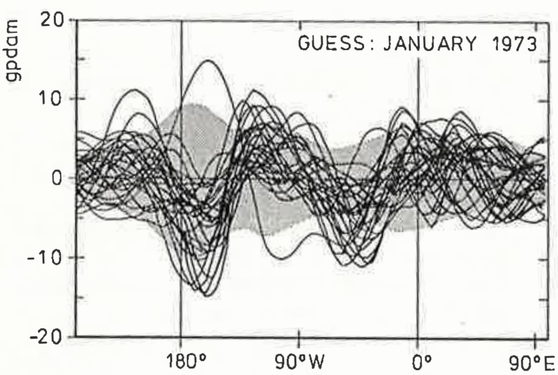
The **January 1973** SST anomaly was quite intense in terms of WRIGHT's (1984) SST index: 1.7 °C. It yields a successful guess, however, with reversed sign: the projection of the GCM signal is significantly **negative** (risk < 1 %). Its strength is about 1/2 of the most powerful guess' projection (A). Figure 9 shows both the parallel and the perpendicular component: The parallel one exhibits the most variance in a sector covering the Atlantic and Eurasia. It is weaker than the perpendicular signal, which resembles the complete signal given in Figure 8a and turns out to be significantly nonzero. According to Figure 10 there is a stable perpendicular signal extending from the dateline downstream to the Atlantic.

The **January 1977** was associated with a weaker SST anomaly-WRIGHT's SST index is 0.9 °C. It yields a successful signal. Its mean projection is 75 % of that found by the most powerful guess (A). The parallel component (Figure 11a) is very similar to the full signal (Figure 8), even with respect to details. The perpendicular component (Figure 11b) is still significantly nonzero, but this significance is – according to Figure 12 – not connected with disturbances common to a clear majority of experimental samples.

The most intense El Niño event in the records is the 1982/83 event. In **January 1983**, WRIGHT's SST index is 2.5 °C. Thus, this event is closest to the considered GCM El Niño experiment with respect to the strength of the equatorial SST anomaly. By means of a similar significance analysis strategy as used in this study, STORCH (1984) showed that the mean Northern Hemisphere circulation in January 1983 was significantly different from those observed in the preceding Januaries.



• **Figure 9** Northern hemisphere extratropical response parallel (a) and perpendicular (b) to guess field "January 1973" derived from DWD analyses of Januaries 1967–84. Spacing: 2 gpdam.



• **Figure 10**

Comparison of NCAR CCM data simulated in the control and in the El Niño anomaly experiment. Component perpendicular to guess "January 1973" derived from DWD analyses.

Univariate analysis of the 30° – 60 °N mean of stationary (January means) disturbances, $\{\overline{z}^*\}$.

Dotted: 95 % band derived from 30 control run Januaries.

Lines: 30 individual Januaries from El Niño experiment.

Therefore, it is quite unexpected that this January 1983 guess pattern fails to achieve the significance of the NCAR CCM response pattern. The signal's component parallel to the guess (Figure 13a) is small and negative. Its length is 7 % of that obtained by (A). The perpendicular signal (Figure 13b) is practically identical to the complete signal.

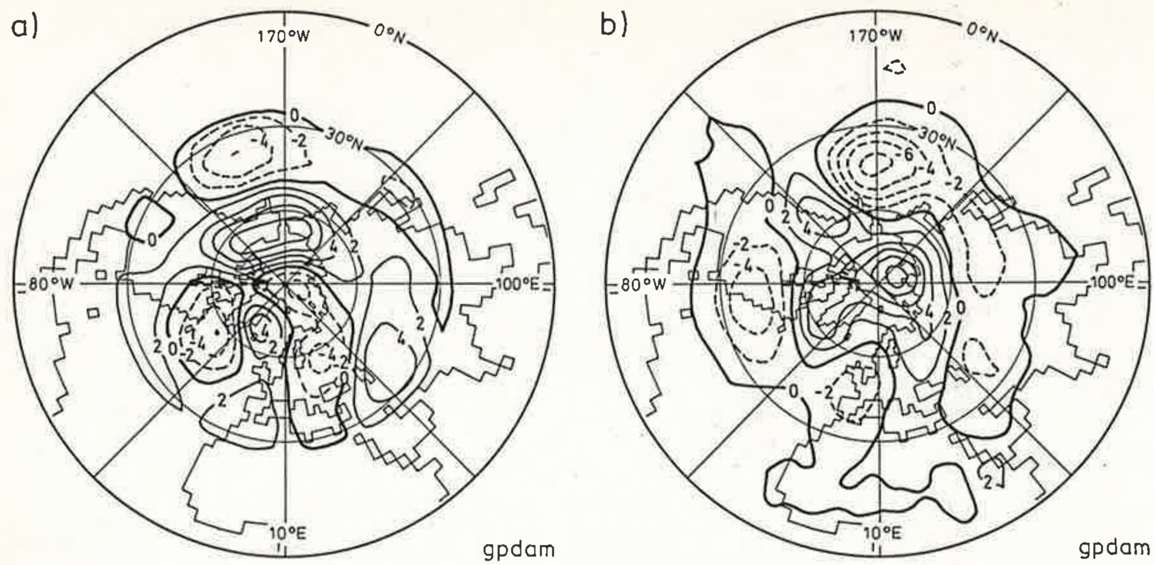


Figure 11 Northern hemisphere extratropical response parallel (a) and perpendicular (b) to the guess pattern "January 1977" derived from DWD analyses Januaries 1967–84. Spacing: 2 gpdam.

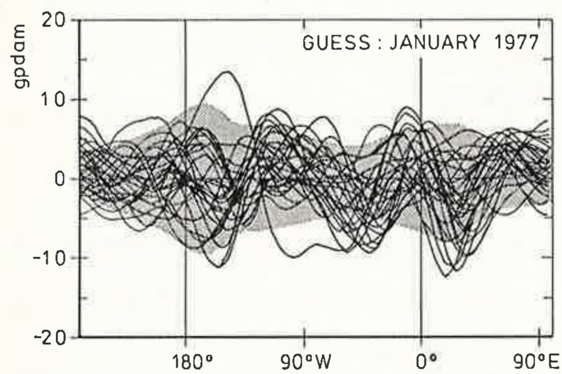


Figure 12

Comparison of NCAR CCM data simulated in the control and in the El Niño anomaly experiment. Component perpendicular to guess "January 1977" derived from DWD analyses.

Univariate analysis of the 30° – 60 °N mean of stationary disturbances, $\{\overline{z}^*\}$.

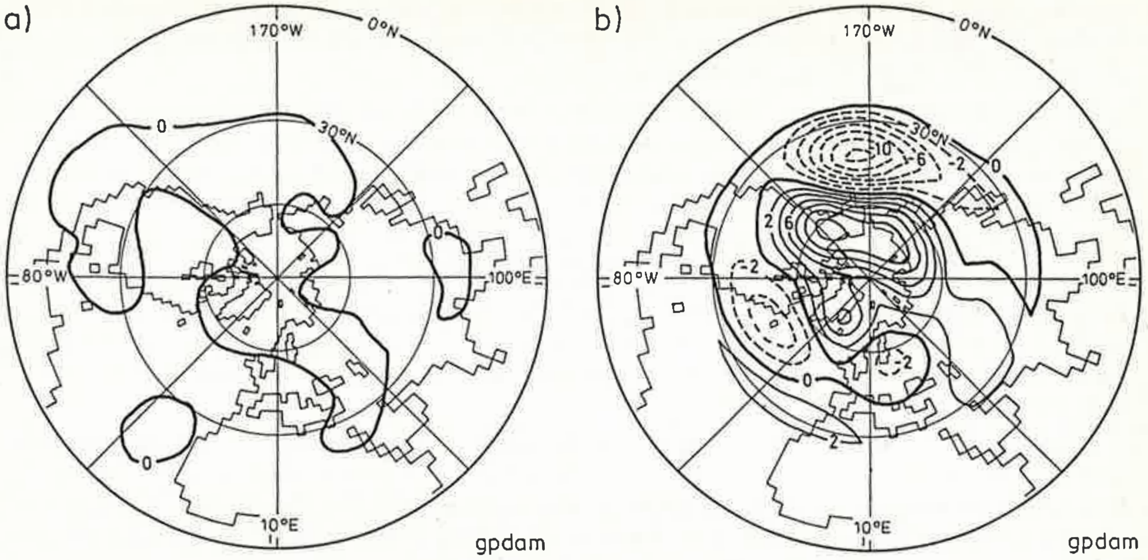
Dotted: 95 % band derived from 30 control run Januaries.

Lines: 30 individual Januaries from El Niño experiment.

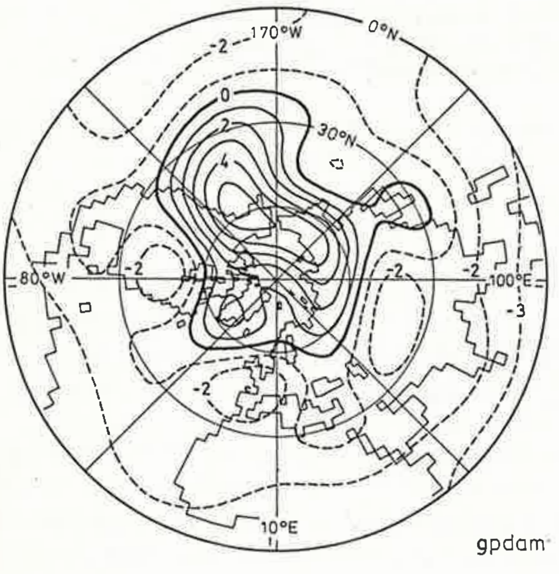
C T21 ECMWF GCM Response Pattern

The similar experiment performed by CUBASCH with the ECMWF T21 GCM gives rise to a signal significantly antiparallel to the NCAR CCM signal: the length of the mean projection amounts to 50 % of (A). As can be deduced from Figure 14, the mean (anti-) parallel signal is considerably shorter than the perpendicular signal.

The negative sign of the mean Fourier coefficient f(S) is not completely unexpected, since the January 1983 guess pattern and the ECMWF T21 GCM response were found to be highly coherent (STORCH and KRUSE, 1985), whereas as pointed out above, the 1983 guess pattern and the NCAR CCM response are (insignificantly) negatively correlated.



• **Figure 13** Northern hemisphere extratropical response parallel (a) and perpendicula (b) to the guess pattern "January 1983" derived from DWD analyses Januaries 1967–84. Spacing: 2 gpdam.



• **Figure 14**
Projection of the NCAR CCM signal "control – anomaly" on the guess pattern "ECMWF T21 GCM El Niño experiment" ("parallel signal"). Spacing: 1 gpdam.

8 Conclusion

From the results presented in this paper it is concluded:

1. **Method:** The proposed method of assessing the statistical significance of response patterns simulated by GCM experiments performs satisfactorily. This particular version, which makes use of problem dependent guess patterns, allows for an assessment as to whether the GCM generated response pattern

is similar to independent observations or simulations. The most powerful results are obtained if the experimental sample set is large enough to be subdivided into one subset for establishing the guess pattern and a second subset for performing the test.

2. **NCAR CCM climatology**: The Northern Hemisphere January 500 mb height climatology is simulated adequately, even though a number of systematic errors were detected: a too steep gradient of the zonal monthly mean (Figure 1), too weak stationary disturbances in midlatitudes (Figure 2) and a somewhat too weak stationary disturbance in the Atlantic sector (Figure 3). The level of transient eddy variance appears to be quite close to observed if compared to that simulated by other GCMs (Figure 6).

3. **Stability of the NCAR CCM response** to an **El Niño SST anomaly**: The NCAR CCM responds in a characteristic way to the anomalous boundary condition in the equatorial Pacific, namely with an eastward shift of the quasistationary patterns in the western hemisphere (Figure 7). In spite of the large natural variability of the (model) atmosphere these characteristics are present in almost all of the experimental Januaries.

4. **Predictive skill of NCAR CCM** with respect to **El Niño SST anomalies**: The use of extratropical anomaly flow patterns observed during the last few El Niño events results in a high coherency between the model simulation and the anomaly flow observed in January 1977. With the January 1973 flow, the model result is negatively correlated. With respect to the January 1983, the NCAR CCM simulation contains no valuable predictive global information (Figure 15).

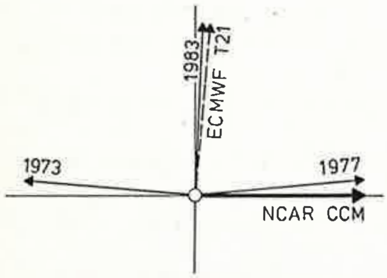


Figure 15
Schematic of the relative directions of the guess vectors "January 1973", "January 1977", January 1983", and "ECMWF T21 GCM response", and of the "NCAR CCM response".

Since the GCM responds stably to the imposed SST anomaly, this behaviour is not related to the model's variability. It is suggested that the larger variability of the real atmosphere, as inferred from the 500 mb height mean January anomalies in 1973, 1977 and 1983, may be related to the interannual variability of the intensity and the pattern of the El Niño SST anomaly. It may also be the result of simultaneously acting anomalies of other external parameters (aerosols, snow cover, SST anomalies outside the equatorial Pacific). Also, one might speculate whether the model's local response in terms of convective heating is too efficient, which could be related either to its strength or to its vertical distribution.

An extratropical response of the NCAR CCM similar to the observed January 1983 anomaly and consequently dissimilar to the observed January 1977, would have been much easier to understand. Then, the similarity could have been related to the similarity of the strength of the prescribed and the observed SST anomalies. The dissimilarity could have been due to the fact that the weakness of the equatorial ENSO 76/77 SST anomaly caused a possibly not statistically significant extratropical response. Also, there was a strong North Pacific SST anomaly in January 1977.

5. **Similarity with ECMWF T21 GCM El Niño response**: Using the response pattern derived from a similar El Niño SST anomaly experiment performed with the ECMWF T21 GCM we found that part of the NCAR CCM response may be explained by the ECMWF T21 GCM response pattern. However,

the correlation is unexpectedly negative. That means that the two considered GCMs, ECMWF T21 and NCAR CCM, are in spite of their similarities (low horizontal resolution, spectral numerical treatment) considerably different, likely in their parameterization of subgrid scale processes, e.g. convection and interaction of flow and orography.

Acknowledgement

I wish to thank Maurice BLACKMON, who supplied me with the NCAR CCM data. Ulla KIRCHER's and Ilka BRECHT's help with typing the manuscript, and Marion GRUNERT's thoroughly prepared diagrams are gratefully acknowledged. The anonymous reviewers' comments have been useful in improving the manuscript's clarity.

References

BLACKMON, M. L., J. E. GEISLER and E. J. PITCHER, 1983: A general circulation model study of January climate anomaly patterns associated with interannual variations of equatorial Pacific sea surface temperature. J. Atmos. Sci. **40,** 1410–1425.

CUBASCH, U., 1985: The mean response of the ECMWF global model to the El Niño anomaly in extended range prediction experiments. Atmosphere-Ocean **23,** 43–66.

FU, C., H. F. DIAZ and J. O. FLETCHER, 1986: Characteristics of the response of sea surface temperature in the Central Pacific associated with warm episodes of the Southern Oscillation. Mon. Wea. Rev. **114,** 1716–1738.

HASSELMANN, K., 1979: On the signal-to-noise problem in atmospheric response studies. Meteorology of the Tropical Oceans, Roy. Met. Soc., 251–258.

LIVEZEY, R. E., 1985: Statistical analysis of general circulation model climate simulation: sensitivity and prediction experiments. J. Atmos. Sci. **42,** 1139–1149.

LIVEZEY, R. E. and W. Y. CHEN, 1983: Statistical field significance and its determination by Monte Carlo techniques, Mon. Wea. Rev. **11,** 46–59.

RASMUSSON, E. and T. CARPENTER, 1981: Variations in tropical SST and surface wind fields associated with the Southern Oscillation/El Niño. Mon. Wea. Rev. **110,** 354–384.

STORCH, H. v., 1982: A remark on Chervin-Schneider's algorithm to test significance. J. Atmos. Sci. **39,** 187–189.

STORCH, H. v., 1984: An accidental result: The mean 1983 January 500 mb height field significantly different from its 1967–81 predecessors. Beitr. Phys. Atmosph. **57,** 440–444.

STORCH, H. v. and A. KRUSE, 1985: The extratropical atmospheric response to El Niño events – a multivariate significance analysis. Tellus **37 A,** 361–377.

STORCH, H. v. and E. ROECKNER, 1983: Verification of general circulation models applied to the Hamburg University GCM. Part I: Test of individual climate states. Mon. Wea. Rev. **111,** 1965–1976.

STORCH, H. v., E. ROECKNER and U. CUBASCH, 1985: Intercomparison of extended range January simulations with general circulation models: Statistical assessment of ensemble properties. Beitr. Phys. Atmosph. **58,** 477–498.

STORCH, H. v. and F. W. ZWIERS, 1987: Recurrence analysis of climate sensitivity experiments. Max-Planck-Institut für Meteorologie, Report **3,** 41 pp.

WRIGHT, P. B., 1984: Relationships between indices of the Southern Oscillation. Mon. Wea. Rev. **112,** 1913–1919.