SIGNIFICANCE ANALYSIS OF SST ANOMALY EXPERIMENTS: USE OF GUESS PATTERNS

Hans von Storch
Meteorologisches Institut der Universität Hamburg
Hanburg, Federal Republic of Germany

1. INTRODUCTION

The significance of multicomponent signals is nowadays often assessed by an introductory series of univariate considerations in the grid point space (originally suggested by Chervin and Schneider) and by the performance of a subsequent 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 done by projecting the original data on 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 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). I will show how this basic version may be applied successfully to midlatitude 500 mb height topography data simulated by the NCAR Climate Community Model (Section 3).

Firstly, some circulation statistics derived from an extended range "control run" is objectively compared with respective circulation statistics based on DWD analyses of Januarys 1967 to 1984 to assess whether the simulated flow is systematically different from the real flow (Section 4).

Secondly, the longitudinal distribution of 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. (Section 5)

Finally, I discuss a particular version of the general concept, which makes use of one problem-dependent guess pattern derived from similar but independent observations or simulations. This approach seems to me rather robust and simple: the projection into a one-dimensional subspace allows one to apply any univariate test. A final univariate analysis to characterize the reponse is no longer necessary. This procedure yields not only a statement concerning the significance, i.e. the stability of the response, but also an assessment of 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 above sketched general technique using problem-independent guesses. (Section 6)

This particular version is used in order to study the NCAR CCM's response to El Niño SST conditions in the equatorial Pacific and to assess whether this response is similar to the 500 mb height topography observed in the El Niño Januarys 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

Since the quantities considered are generally given as fields, multivariate techniques have to be used to assess the (statistical) significance of a response pattern. However, multivariate tests are likely to fail to detect an existing nonzero signal if too many (noisy) parameters (e.g. grid point values) are used to describe the atmospheric state. A simple example demonstrating this fact is given by Storch and Kruse (1985).

A first step to reduce the dimension of the problem can be an averaging over certain areas or a restriction on certain areas of interest. As a second step, the spatially distributed data may be expanded into a fast-converging series of orthonormal functions (to be precise: discrete functions = vectors). Since only the first few of the expansion coefficients are connected with significant variance, only the first few need to be retained.

After the dimension of the problem has been reduced in this way, a multivariate test has to be done to assess whether the null hypothesis of no signal is in conflict with the data. If so, the test's outcome is the statement that the probability of a zero signal is less than, say, 5%. This statement is taken as a statistical proof of a nonzero signal, but it is no more than a proof of its existence. Nothing is said about its character, about its spatial distribution or about its intensity.

The following thought experiment shows that one is not allowed to interpret each detail of the mean signal as "significant", i.e. as sample-independent properties of the true nonzero signal. Assume the considered area is divided into two subareas denoted by A and B. In subarea A there is a nonzero signal, while in B there occurs only random variations. If the signal in A is sufficiently strong, the significance analysis applied to the complete area A+B will assess the signal to be nonzero in A+B. If we plot the mean signal, we will find a well-defined signal in A and some structures in B, which are exclusively due to chance.

Thus, the multivariate tests have to be followed by an analysis where (in the grid point space or in some spectral space) the experimental samples exhibit patterns systematically different from the control samples. 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, control states. This analysis may be done by ordering both the control and the experimental data and calculating the Mann-Whitney test statistic locally. Another way is to estimate roughly a 95% band of the control ensemble; that is for each grid point an interval which will contain about, say, 95% of all control states. Together with the 95% band, all individual experimental states are plotted. In that way one may easily demonstrate where the experimental samples show structures stably different from the control samples. This procedure and the resulting diagrams are used throughout this study.

3. NCAR CCM EXPERIMENTS AND AVAILABLE DATA

Using the NCAR Climate Community Model (CCM) two experiments were performed in the perpetual January mode, each integrated over a 200 day spin-up period and an experimental range 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 while in the second experiment a doubled Rasmusson and Carpenter (1983) 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 split up into a series of 30 Januarys in the following way. Firstly, the time series is broken into adjacent 40 days intervals. Secondly, the first 10 days of each subinterval are disregarded and the remaining 30 days are considered to form one "January". The 10 day gap between each January should be sufficient to obtain statistically independent Januarys . In that way, a total of 30 control Januarys and 30 "El Niño" Januarys is obtained.

I restrict myself to the extratropical Northern Hemisphere characterized by the 500 mb height topography.

4. VERIFICATION OF NCAR CCM 500 MB HEIGHT CLIMATOLOGY

The above mentioned basic strategy is used to compare some Northern Hemisphere circulation statistics derived from the control experiment and from DWD analyses of Januarys 1967 to 1984. (The method and its results obtained for the ECMWF T21 and T40 GCM and for the Hamburg University Model are described in detail by Storch et al. (1985).)

The first step of the strategy -the reduction of degrees of freedom- is done by expanding into an EOF series and using the first 5 EOF coefficients, only. For the multivariate test -the second step- we use the generalized Mann/Whitney test. The univariate analysis is done as sketched in Section 1 by plotting an estimated 95% band of the observed state and all individual simulated states.

The considered Northern Hemisphere circulation statistics calculated from daily 500 mb height data characterize the "stationary" (= January mean) and "transient" (= January daily variability) state.

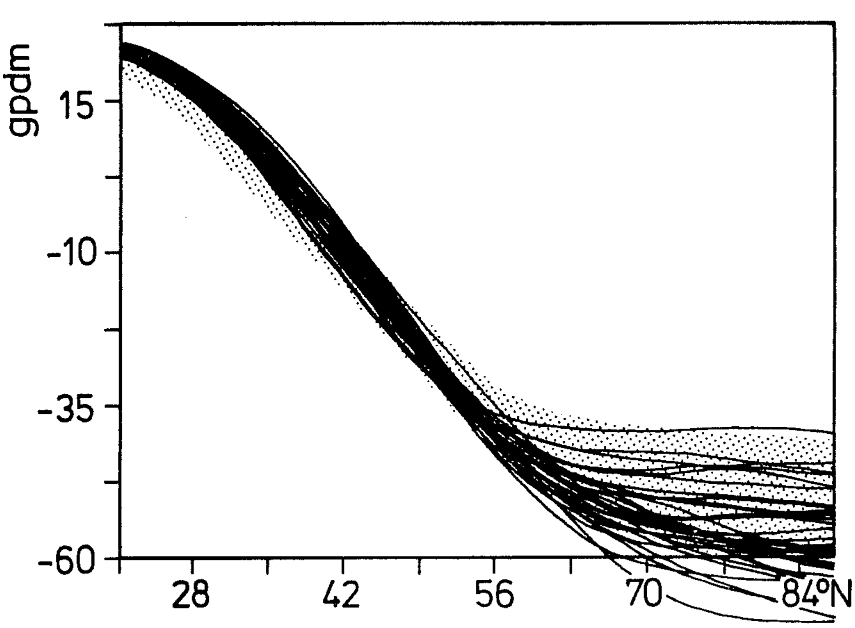


Figure 1.a: Verification of Northern Hemisphere NCAR CCM 500 mb height climatology. Univariate analysis of $[\bar{z}]$. Dotted: 95% band derived from DWD analyses of Januaries 1967-84 Lines: Individual NCAR CCM generated states.

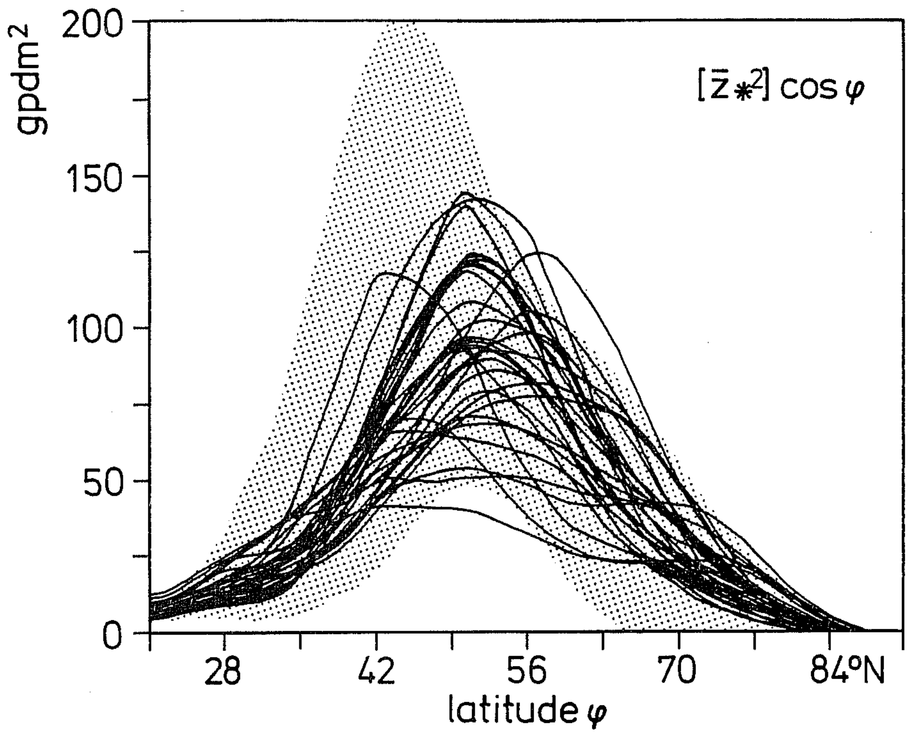
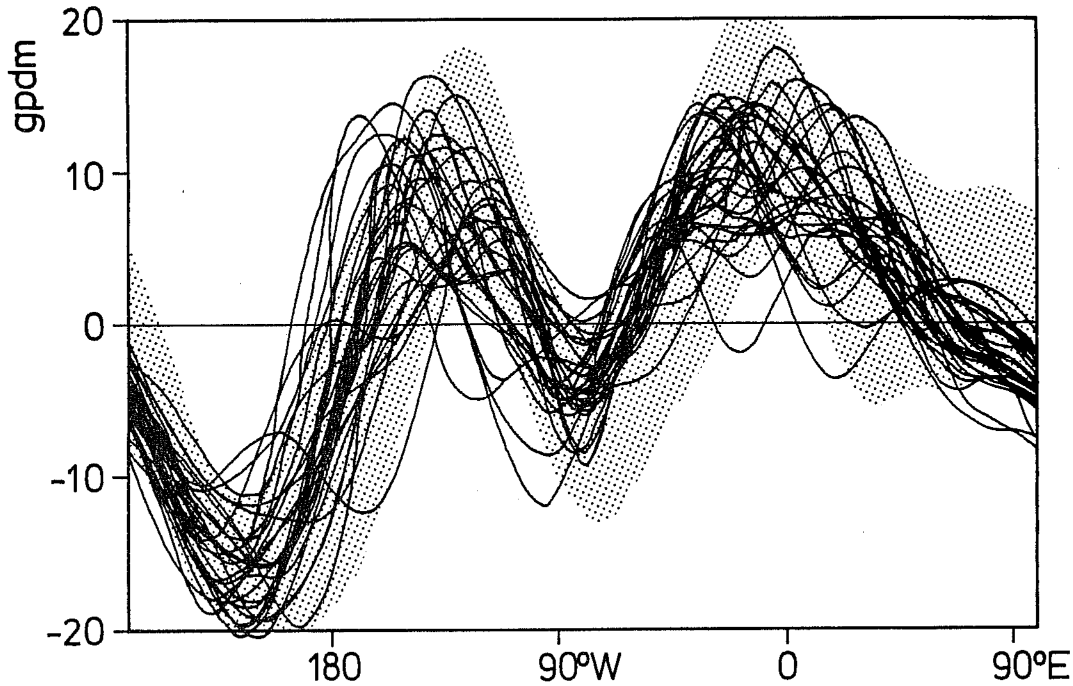


Figure 1.b:
Verification of Northern Hemisphere NCAR CCM 500 mb height climatology. Univariate analysis of zonally averaged variance due to stationary eddies, $[\bar{z}^{*2}] \cos\varphi$.
Dotted: 95% band derived from DWD analyses of Januarys 1967-84
Lines: Individual NCAR CCM generated states.

Figure 1.c:
Verification of Northern Hemisphere NCAR CCM 500 mb height climatology. Univariate analysis of the 30-60°N mean of stationary disturbances, $\{\bar{z}^*\}$
Dotted: 95% band derived from DWD analyses of Januarys 1967-84
Lines: Individual NCAR CCM generated states.



4.1 The stationary circulation

The January mean circulation is characterized by three parameters, namely:

* the zonal monthly mean, $[\bar{z}]$

* the zonally averaged variance due to stationary eddies, weighted with the cosine of latitude: $[\bar{z}^{*2}]\cos(\varphi)$

* the 30-60°N average of stationary eddies: $\{\bar{z}^*\}$

The first two parameters are functions of latitude, the third a function of longitude.

The result of the multivariate test in the 5-dimensional subspace spanned by the first 5 EOFs is that the GCM generated states are significantly different from the respective observed states, whichever of the three parameters is used.

The univariate analysis is displayed in Figure 1, showing the 95% band estimated from the 18 DWD-analysed Januarys and the 30 individual CCM-generated states. According to Figure 1.a the CCM midlatitude gradient of $[\bar{z}]$ is much too steep, which is unfortunately connected with a too intense zonal flow in the middle troposphere. The intensity of stationary eddies is somewhat too weak and shifted northwards(Figure 1.b). The longitudinal distribution, $\{\bar{z}^*\}$, exhibits a well developed Pacific stationary disturbance, but the Atlantic system is rather weak compared to that observed(Figure 1.c).The differences with respect to the stationary disturbances appear to be of minor severity, especially if compared with the performance of other GCMs (Figure 2): apparently the NCAR CCM is at the upper bound of present-day GCMs ability to reproduce actual climate.

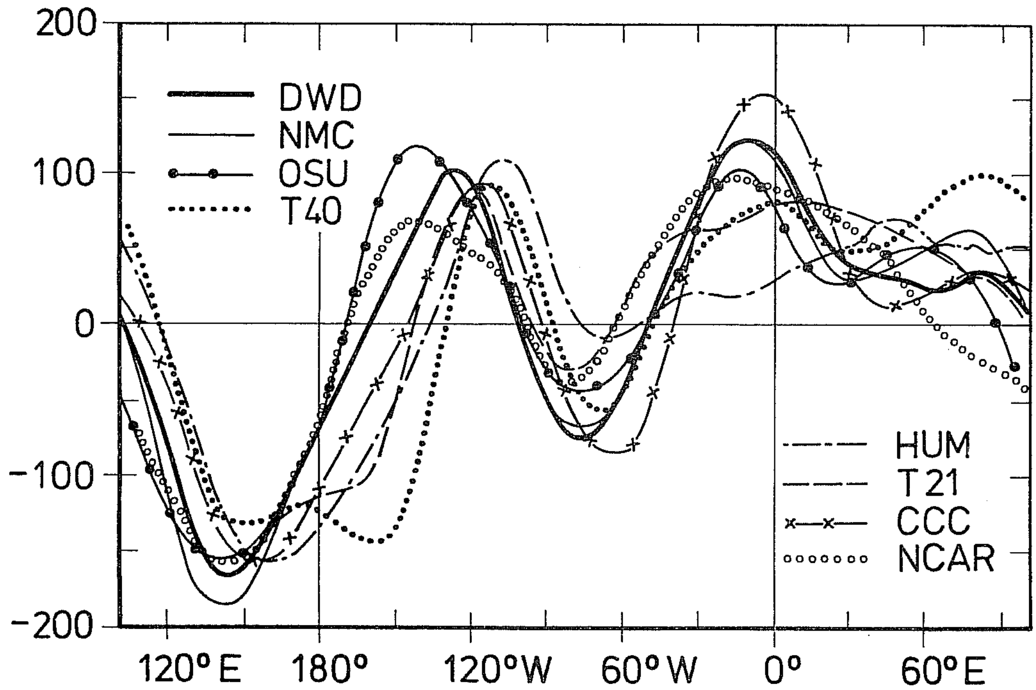
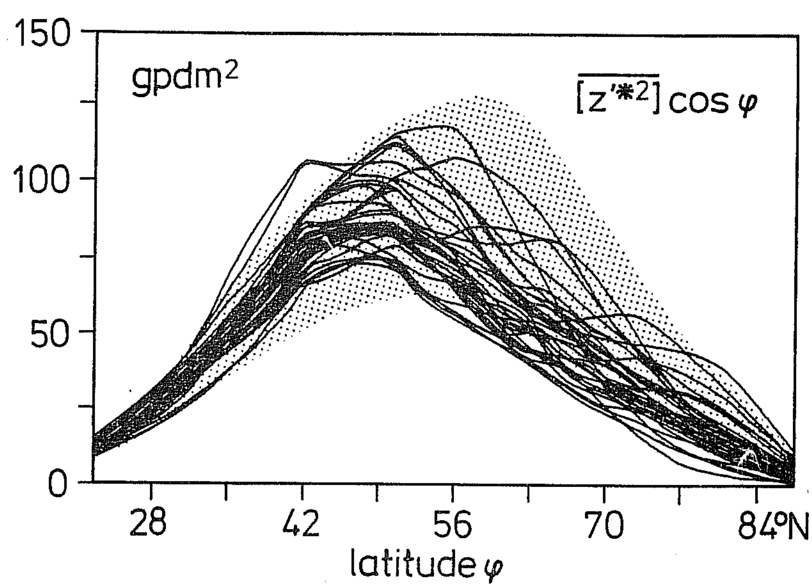


Figure 2:
Intercomparison of NCAR CCM January 500 mb height climatology with observed and by other GCMs simulated climatology.
Quantity: 30-60°N mean of stationary disturbances, $\{\bar{z}^*\}$
(DWD = DWD analyses 1967-85; NMC = NMC analyses 1956-66; T21 & T40 = ECMWF; HUM = Hamburg University; CCC = Canadian Climate Centre; OSU = Oregon State University)

Figure 3.a:
Verification of Northern Hemisphere NCAR CCM 500 mb height climatology.
Univariate analysis of the zonally averaged variance due to transient eddies, $[\overline{z'^{*2}}] \cos(\varphi)$
Dotted: 95% band derived from DWD analyses of Januarys 1967-84
Lines: Individual NCAR CCM generated states.



4.2 The transient component

The January daily variability is summarized by two parameters:

* the zonally averaged variance due to transient eddies weighted with the cosine of latitude: $[\overline{z'*^2}]\cos(\varphi)$

* the 30-60°N average of transient variance: $\{\overline{z'^2}\}$

The first parameter is a function of latitude and the second a function of longitude.

The result of the multivariate test is that the GCM-generated level of daily variability differs significantly from the observed level. According to Figure 3.a, this is connected with an underestimation of transient variance north of, say, 50°N, while the longitudinal distribution is practically sufficiently well simulated, even if the overall level is simulated slightly too low (Figure 3.b). An comparison with other GCMs shows that the NCAR CCM simulates the transients favorably (Figure 4).

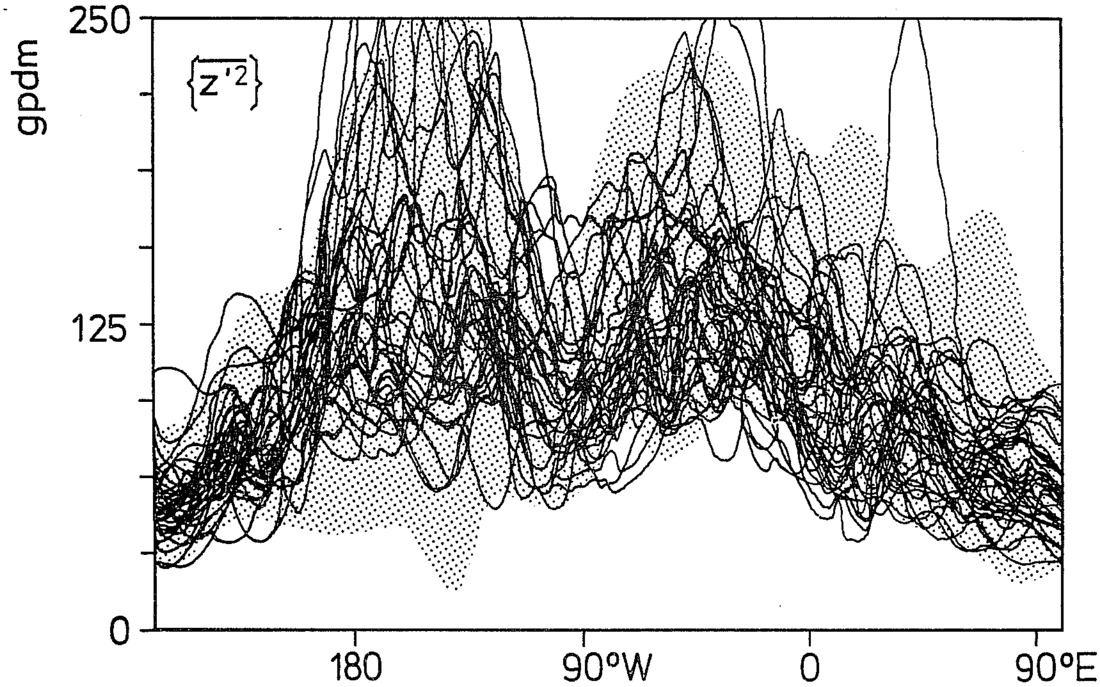
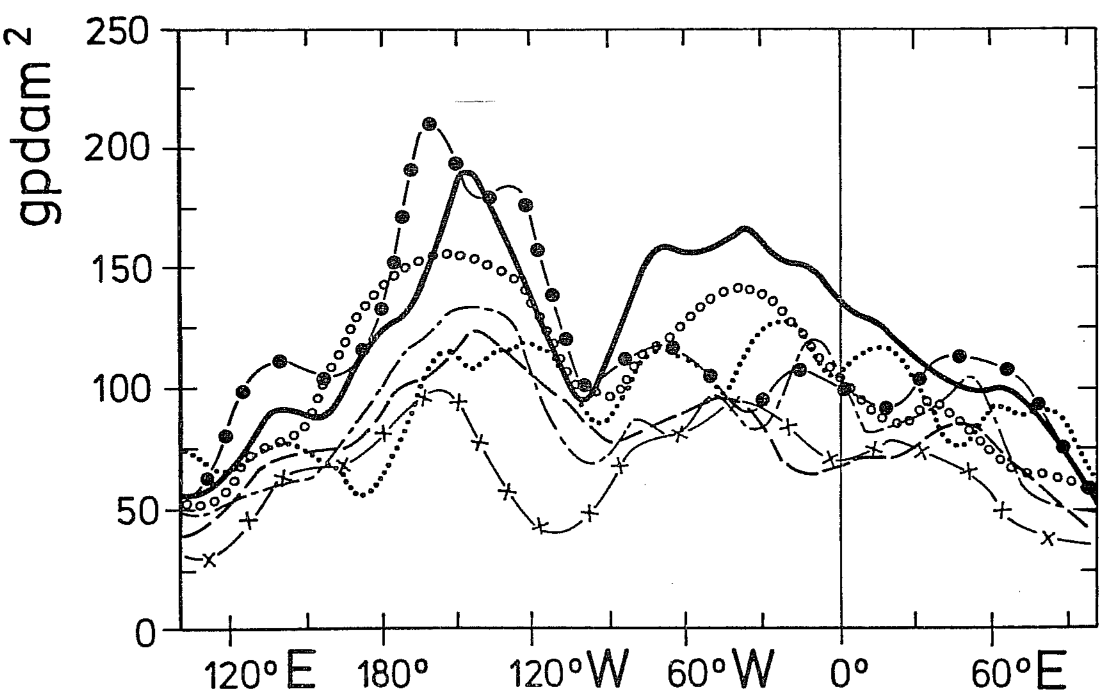


Figure 3.b:
Verification of Northern Hemisphere NCAR CCM 500 mb height climatology. Univariate analysis of the 30-60°N mean of variance due to transient eddies, $\{\overline{z'^2}\}$
Dotted: 95% band derived from DWD analyses of Januarys 1967-84
Lines: Individual NCAR CCM generated states.

Figure 4:
Intercomparison of NCAR CCM January 500 mb height climatology with observed and by other GCMs simulated climatology.
Quantity: 30-60°N mean of variance due to transient eddies: $\{\overline{z'^2}\}$
(The heavy solid curves shows $\{\overline{z'^2}\}$ derived from DWD analyses of Januaries 1967-84; the open circles belong to the NCAR CCM; for the other curves, see Figure 2.)



5. EL NINO EXPERIMENT SIGNIFICANCE ANALYSIS WITHOUT A-PRIORI GUESSES

If no-apriori guesses 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 and transient eddy circulation statistics defined in Section 4 were tested: statistically significant signals were detected in the stationary component only.

Figure 5 displays the univariate analysis of the response in terms of the 30-60°N mean. Apparently one of the 30 experimental Januarys exhibits a longitudinal distribution totally different from the other 29. Apart from this 'strayshot', a clear tendency towards a lowered 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 variance due to stationary eddies was classified as being significantly different in the El Niño experiment. The univariate analysis (Figure 6) shows, however, that this statistical significance is not connected with physical relevance.

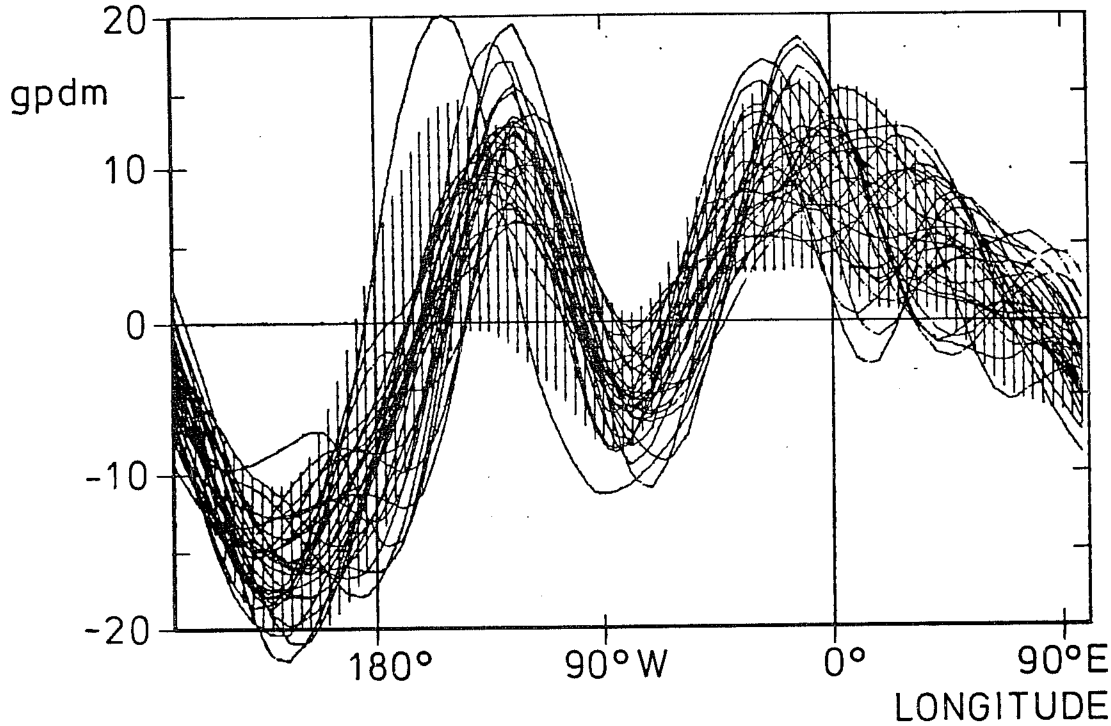
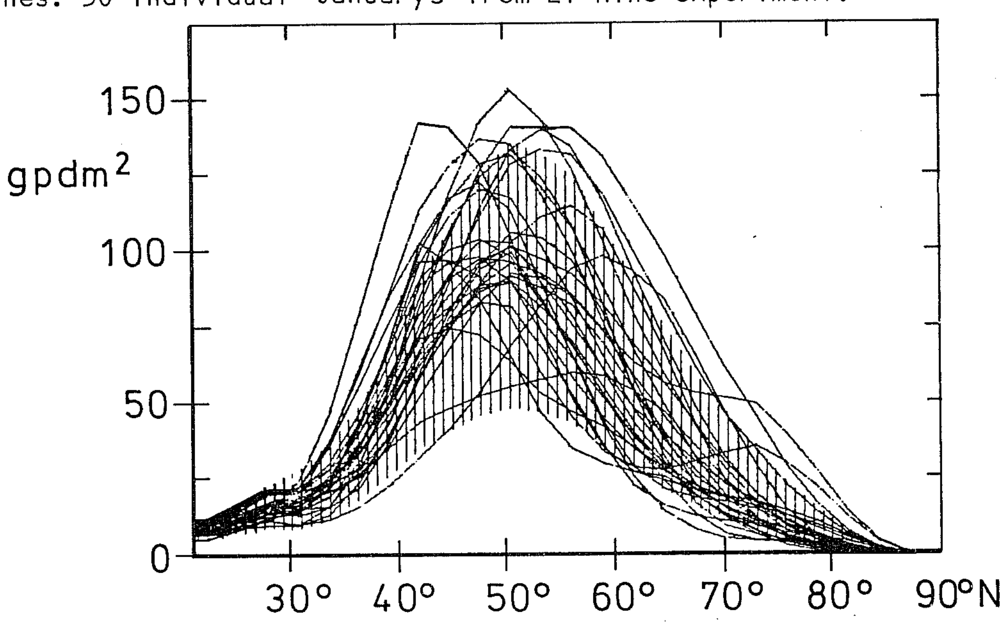


Figure 5:
Intercomparison of Northern Hemisphere NCAR CCM 500 mb height data simulated in the control and in the El Niño anomaly experiment.
Univariate analysis of the 30-60°N mean of stationary disturbances, $\{\bar{z}^*\}$
Shaded: 95% band derived from 30 control run Januarys .
Lines: 30 individual Januarys from El Niño experiment.

Figure 6:
Intercomparison of Northern Hemisphere NCAR CCM 500 mb height data simulated in the control and in the El Niño anomaly experiment.
Univariate analysis of the zonally averaged variance due to stationary eddies, $[\bar{z}^{*2}]\cos(\varphi)$
Shaded: 95% band derived from 30 control run Januaries.
Lines: 30 individual Januarys from El Niño experiment.



6. THE PARTICULAR VERSION OF THE SIGNIFICANCE ANALYSIS STRATEGY

If an a-priori notion of 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 first step -reduction of degrees of freedom- the complete signal denoted by S is split up into one component denoted by P parallel to the a-priori fixed guess vector G, and one component Q perpendicular to G:

$$S = P + Q$$

with $P = (S,G)G$ and $(Q,G) = 0$

Here, (,) is defined to be the standard dot product. The guess vector G is assumed to be normalized, i.e. $(G,G) = 1$.

6.1 The parallel component

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, the mean difference P of the parallel component simulated in the control and the anomaly experiment may be assessed by any standard univariate test as to whether it 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 Fourier-coefficients f derived from the individual control or experimental fields.

The advantage of this procedure is that, in addition to the significance of the complete signal, the similarity of the signal with an a-priori fixed pattern is also considered.

A drawback of this concept of considering the "component parallel to the guess" is the fact that the guess pattern is phase fixed. Thus, it can happen, at least theoretically, that the guess pattern coincides perfectly with the actual signal except for a small phase shift. In that case the projection of the guess pattern would not lead to the (correct) assessment that signal and guess pattern are highly coherent.

6.2 The perpendicular component

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 are in 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 identified as being 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 way, 3 different types of guesses are available, namely:

(A) The 500 mb height topography anomalies observed during the Januarys (mature phase; year "+1") of the last few strong El Niño events, i.e. Januarys 1973, 1977 and 1983. These anomalies were derived from the data already used for verifying the NCAR CCM in Section 4.

(B) The 500 mb height topography anomaly simulated in an extended range GCM experiment by Cubasch (1985) with the T21 GCM of ECMWF. Similar to the experiment described in Section 3, Cubasch performed a series of "normal" January simulations and "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 coworkers in the NCAR CCM experiment.

The ECMWF T21 GCM generated response pattern evolved to be significantly nonzero and, furthermore, to be highly coherent with the January 1983 anomaly pattern (Storch and Kruse, 1985).

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

It has to be admitted that in most applications not enough samples are available to allow for a subdivision into two sufficiently large subsamples. Therefore, guess (C) will not, in general, be available.

7.2 Results

7.2.1 NCAR CCM experiment guess: (C)

As was to be expected a-priori, the guess derived from every second experimental sample yields the highest significance with respect to the parallel component. The mean generalized Fourier-coefficient f(S) amounts to 164 units. Figure 7 shows the complete signal and the guess pattern calculated from the other half of the data. Both maps are very similar with respect to pattern and magnitude.

To demonstrate 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 counted. According to this, the probability of a positive (S,G) is only 43% in the control run but up to 97% in the El Nino SST anomaly experiment.

The signal perpendicular to the guess pattern turned out to be not

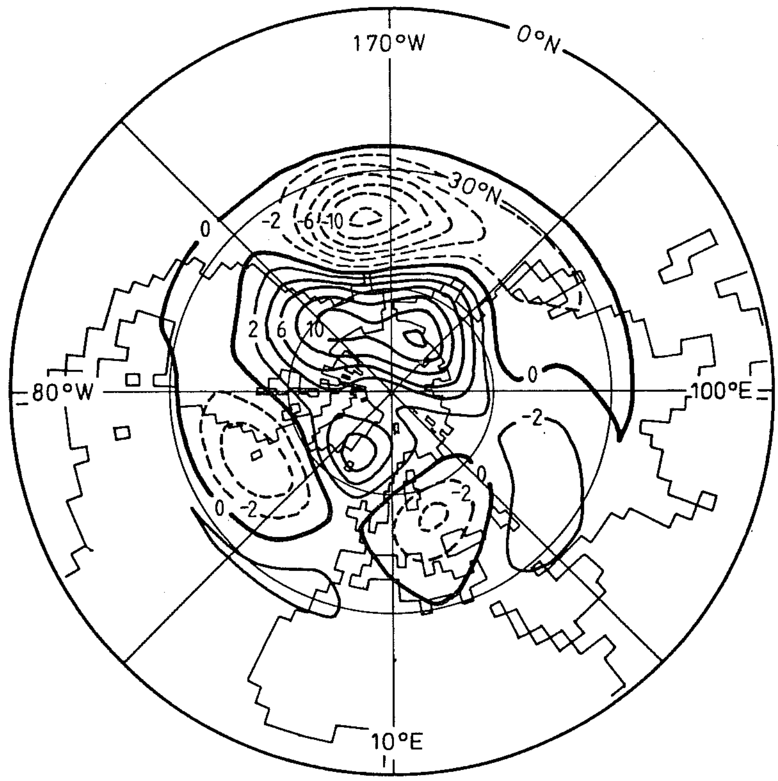
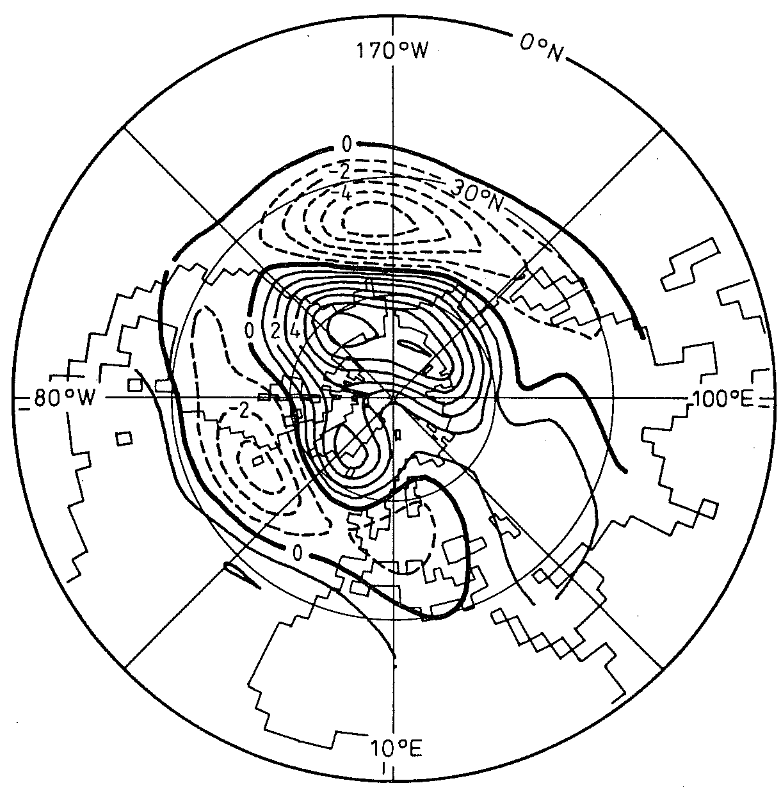


Figure 7.a:
Complete signal of NCAR CCM El Niño experiment in terms of 500 mb height: control minus experimental data. Spacing: 2 gpdm.

Figure 7.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 gpdm.



significantly nonzero. According to the univariate analysis displayed in Figure 8, 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 (Figure 9) has a magnitude one order less than the complete or the parallel signal (Figure 7).

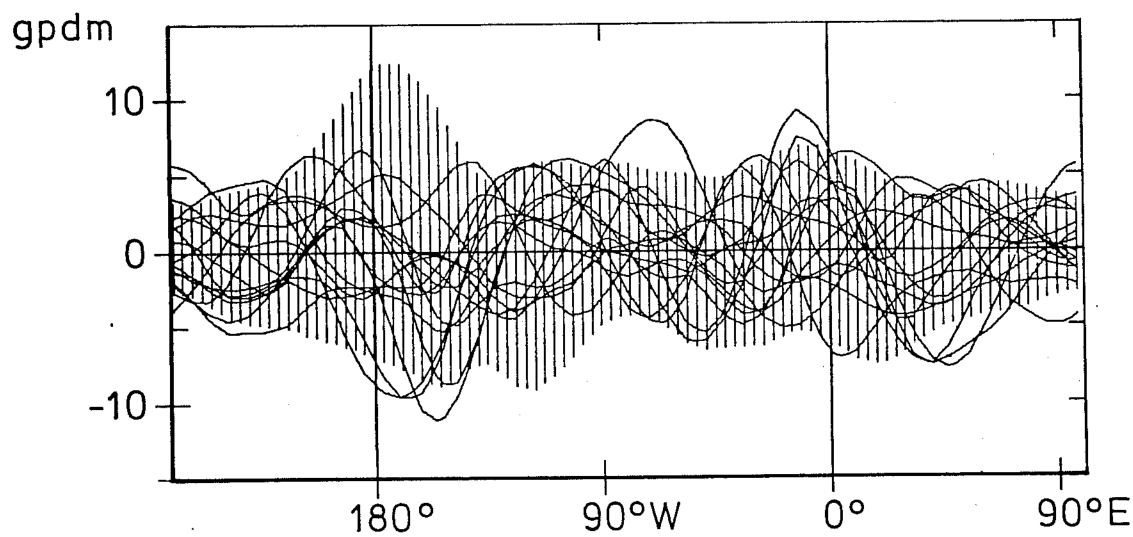


Figure 8: Comparison of NCAR CCM data simulated in the control and in the El Niño anomaly experiment. Component perpendicular to guess (C) built up from half of the experimental data.
Univariate analysis of the 30-60° N mean of stationary disturbances, $\{\bar{z}^*\}$
Shaded: 95% band derived from 30 control run Januarys.
Lines: 15 individual Januaries from El Niño experiment.

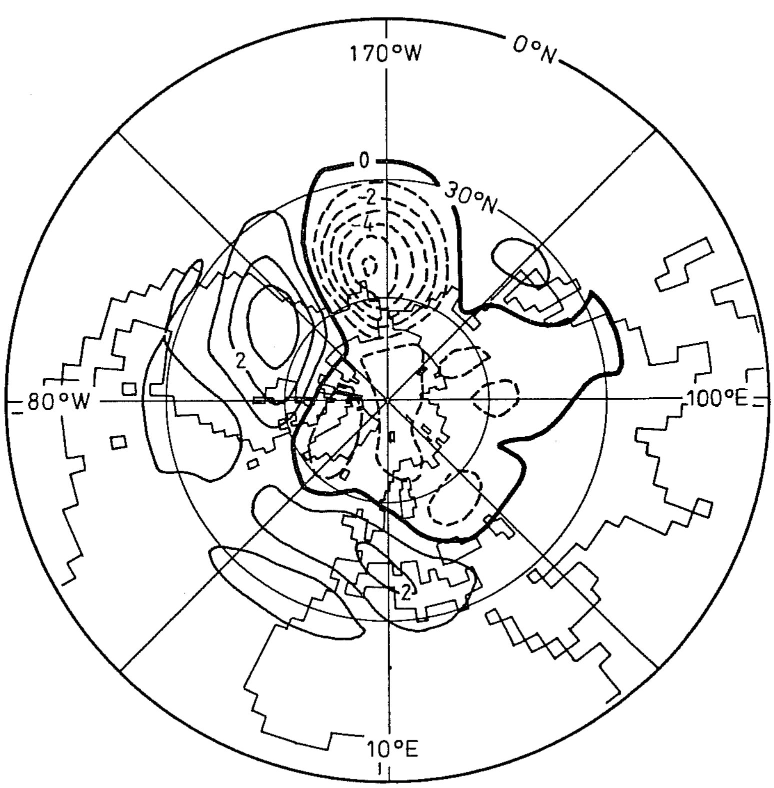


Figure 9:
Signal (derived from half of the experimental data) perpendicular to the guess (C) build up from the other half of the experimental data ("perpendicular signal").
Spacing: 1 gpdm.

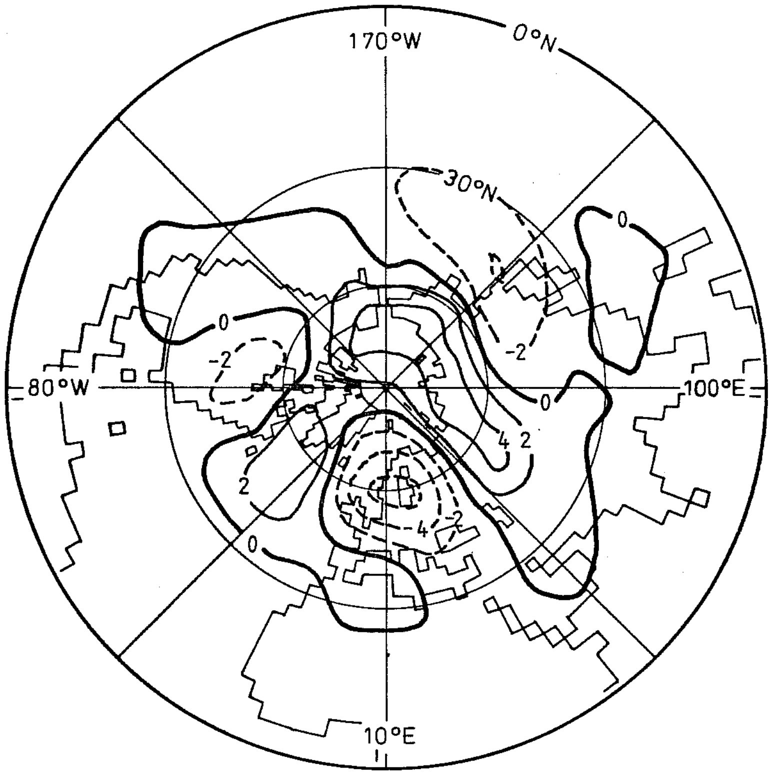
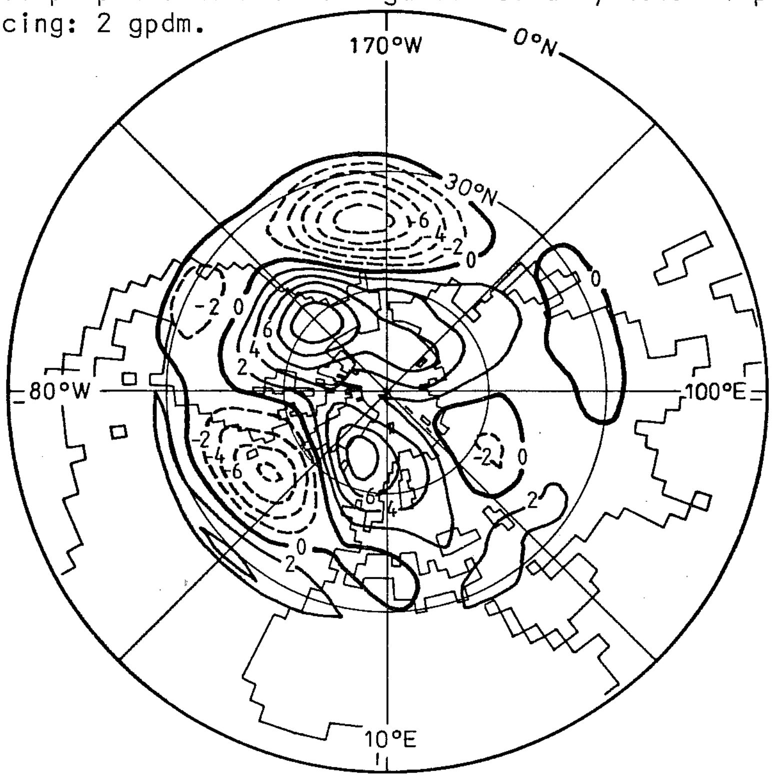


Figure 10.a:
Projection of the signal "control - anomaly" on the guess pattern "January 1973" derived from DWD analyses Januarys 1967-84 ("parallel signal"). Spacing: 2 gpdm.

Figure 10.b:
Complete signal perpendicular to the guess "January 1973" ("perpendicular signal"). Spacing: 2 gpdm.



7.2.2 Observed guesses: (A)

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

The January 1973 SST anomaly was quite intense in terms of Wright's (1984) SST index: 1.74°C. It yields a successful guess, but with reversed sign: the projection of the GCM signal is significantly negative, namely -88 units on average. Figure 10 shows both the parallel and the perpendicular component: the parallel one exhibits 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 7 and turns out to be significantly nonzero. According to Figure 11 there is a stable perpendicular signal stretching from the dateline downstream to the Atlantic.

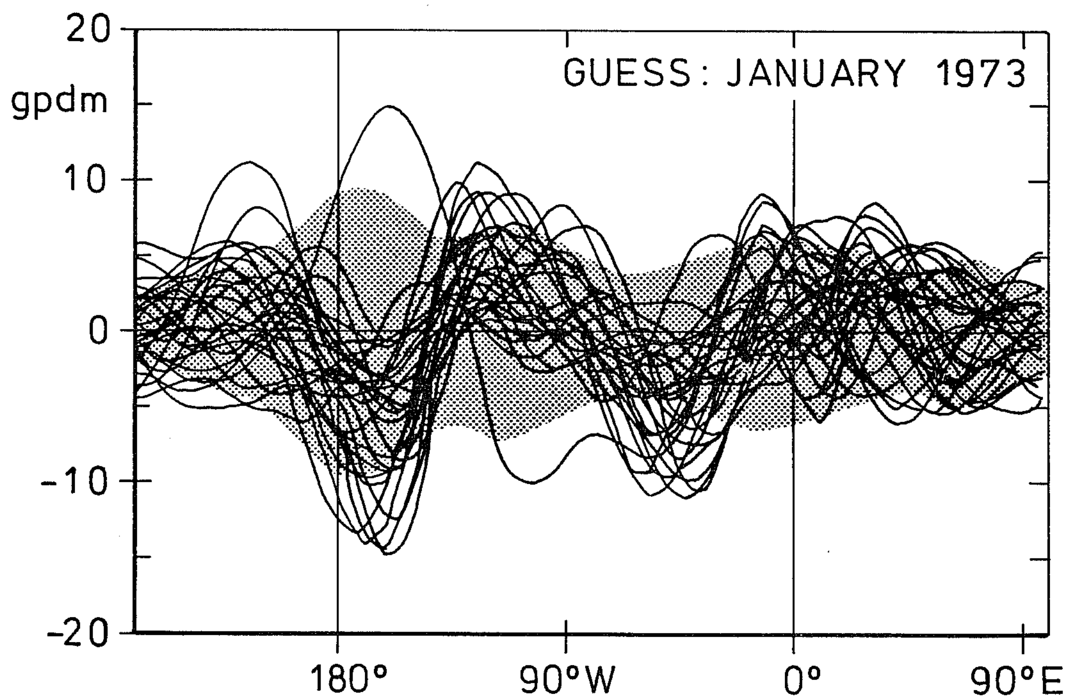


Figure 11: Intercomparison of NCAR CCM data simulated in the control and in the El Niño anomaly experiment. Component perpendicular to guess "January 1973" derived from DWD analyses.
Univariate analysis of the 30-60°N mean of stationary disturbances, $\{\bar{z}^*\}$
Dotted: 95% band derived from 30 control run Januaries.
Lines: 30 individual Januaries from El Niño experiment.

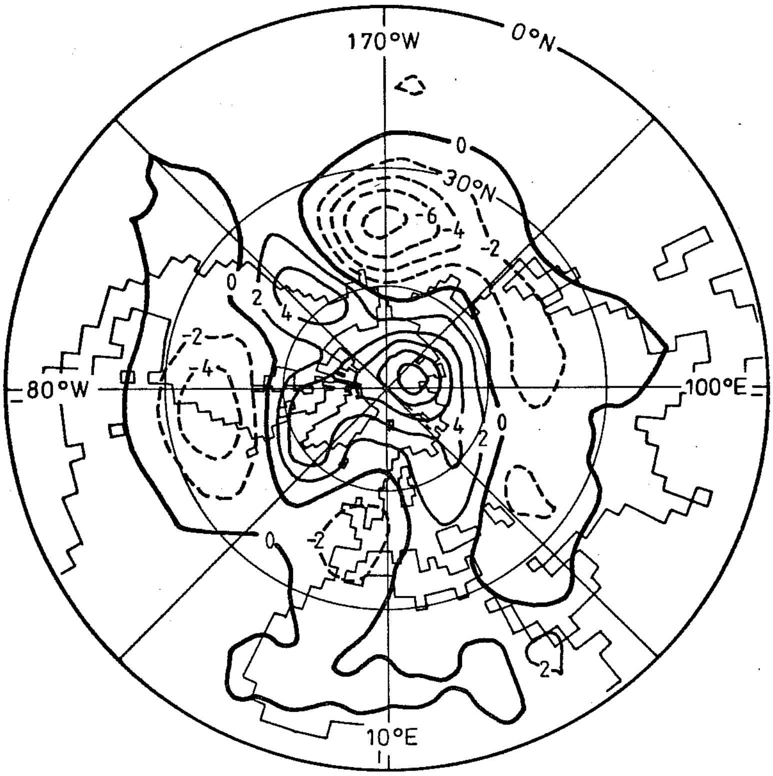


Figure 12.a:
Projection of the signal "control - anomaly" on the guess pattern "January 1977" derived from DWD analyses Januarys 1967-84 ("parallel signal"). Spacing: 2 gpdm.

Figure 12.b:
Complete signal perpendicular to the guess "January 1977" ("perpendicular signal"). Spacing: 2 gpdm.

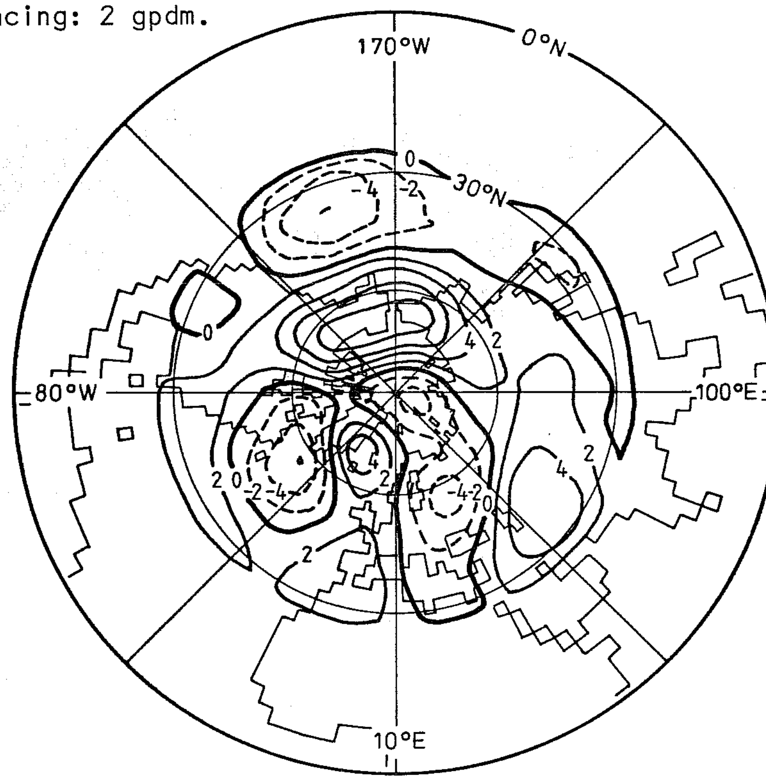
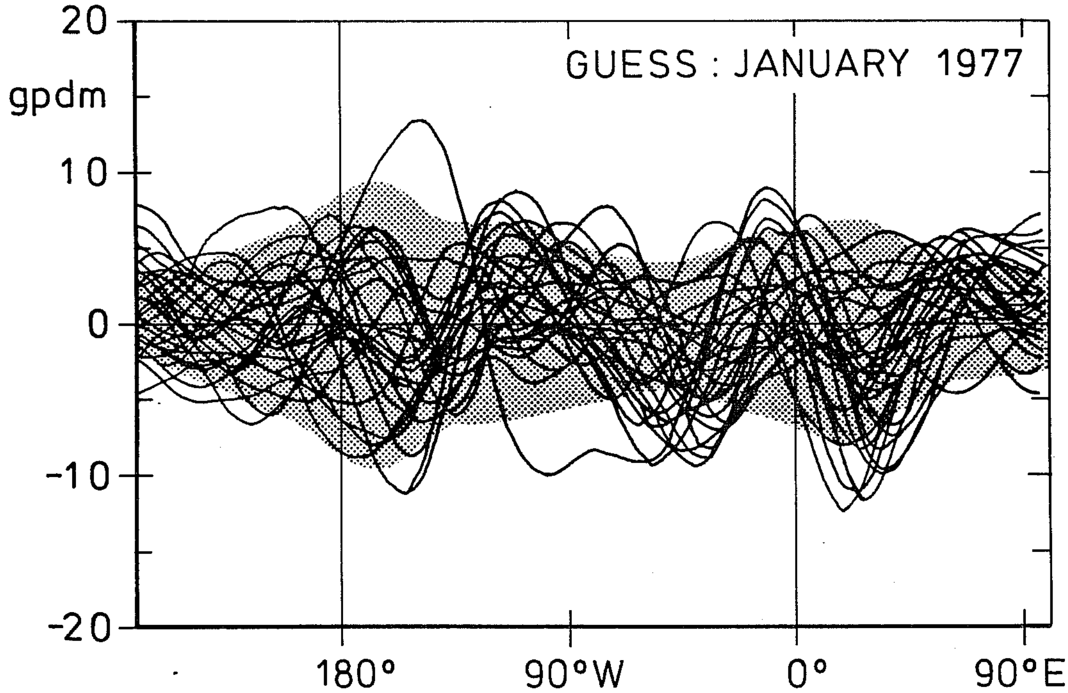


Figure 13: Intercomparison of NCAR CCM data simulated in the control and in the El Niño anomaly experiment. Component perpendicular to guess "January 1977" derived from DWD analyses.
Univariate analysis of the 30-60°N mean of stationary disturbances, $\{\bar{z}^*\}$
Dotted: 95% band derived from 30 control run Januarys.
Lines: 30 individual Januaries from El Niño experiment.



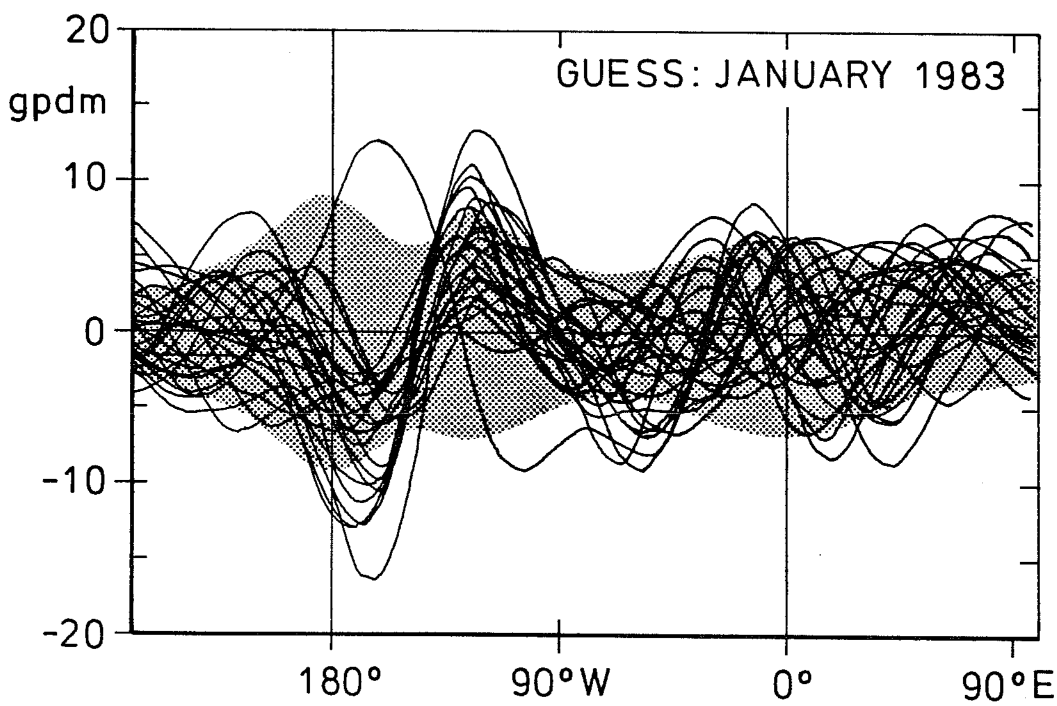
January 1977 was associated with a somewhat weaker SST anomaly - Wright's SST index is 0.85°C. It yields a very successful signal with a mean projection of +122 units. The parallel component (Figure 12a) is very similar to the full signal (Figure 7), even with respect to details. The perpendicular component (Figure 12b) is still significantly nonzero, but this significance is - according to Figure 13- not connected with disturbances common to all or to most experimental samples.

The most intense El Niño event in the records is the 1982/83 event. In January 1983, Wright's SST index is 2.55°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 that observed in the preceding Januarys.

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 is nearly zero, namely -12 units. The perpendicular signal (Figure 14) is practically identical to the complete signal.

Figure 14: Intercomparison of NCAR CCM data simulated in the control and in the El Niño anomaly experiment. Component perpendicular to guess "January 1983" derived from DWD analyses.
Univariate analysis of the 30-60°N mean of stationary disturbances, $\{\bar{z}^*\}$
Dotted: 95% band derived from 30 control run Januarys.
Lines: 30 individual Januarys from El Niño experiment.



7.2.3 T21 ECMWF GCM response pattern: (B)

A similar experiment performed by Cubasch with the ECMWF T21 GCM gives rise to a signal significantly antiparallel to the NACR CCM signal: the mean projection amounts to -81 units. As can be deduced from Figure 15, the mean (anti-) parallel signal is considerably shorter than the perpendicular signal.

The negative sign of the mean Fourier coefficient f(S) is not that unexpected, since the 1983 guess pattern and the ECMWF T21 GCM response were found to be highly coherent (Storch and Kruse, 1985). On the other hand, it was stated above that the 1983 guess CAR CCM are (insignificantly) negatively correlated.

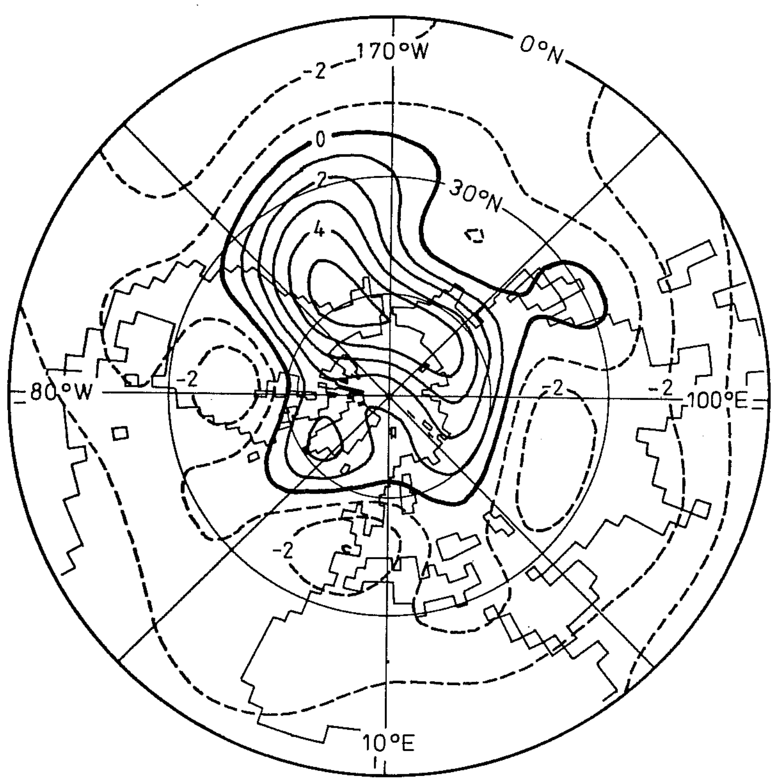
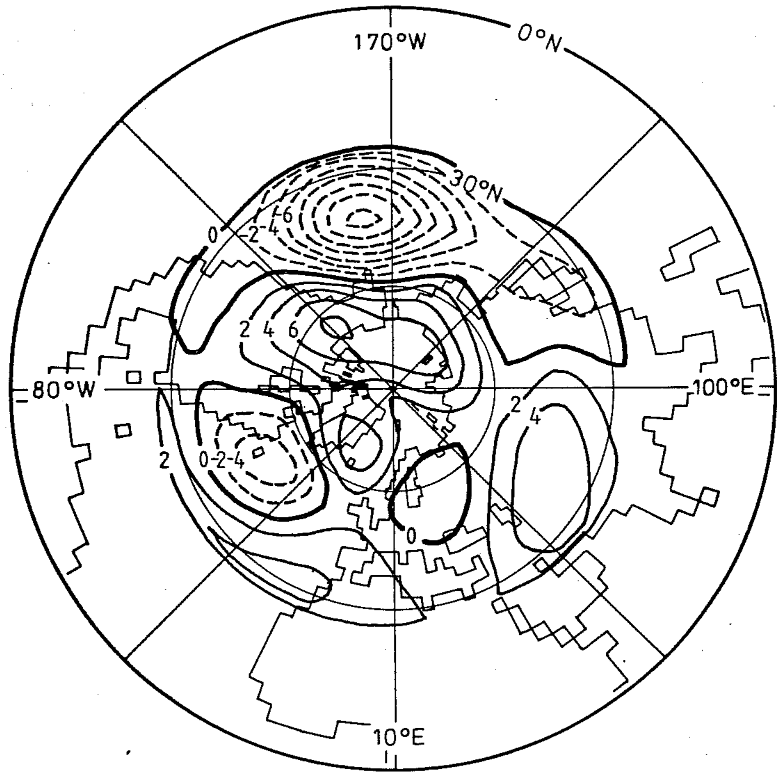


Figure 15.a:
Projection of the NCAR CCM signal "control - anomaly" on the guess pattern "ECMWF T21 GCM El Niño experiment" (parallel signal"). Spacing: 2 gpdm.

Figure 15.b:
300 mb height signal perpendicular to the guess "ECMWF T21 GCM El Nino experiment" ("perpendicular signal"). Spacing: 2 gpdm.



8. CONCLUSION

From the results presented in this paper, it is concluded that:

1. Method: The proposed method to assess the statistical significance of response patterns simulated by GCM experiments performs satisfactorily. The particular version, which makes use of problem dependent guess patterns, allows an assessment of 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 sufficiently large for it to be subdivided into one subset for establishing the guess pattern and a second subset to perform 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, a too weak stationary disturbances in mid latitudes and a somewhat too weak eddy in the Atlantic sector. The level of transient eddy variance appears to be favorable compared to that simulated by other GCMs.

3. 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 coherence 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 January 1983, the NCAR CCM simulation contains no valuable predictive global information.

4. Similarity with ECMWF T21 GCM El Niño response: The use of the response pattern derived from a similar El Nino SST anomaly experiment performed with the ECMWF T21 GCM, yields that part of the NCAR CCM response may be explained by the ECMWF T21 GCM response pattern. However, the correlation is unexpectedly negative.

ACKNOWLEDGEMENT: I wish to thank Maurice Blackmon, who supplied me with the simulated data.

REFERENCES

Blackmon, M.L.; J.E. Geisler; 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.

Conover, W.J., 1971: Practical nonparametric statistics. - New York London Sydney Toronto Wiley

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.

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.; W.Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. - Mon. Wea. Rev. 11, 46-59.

Rasmusson, E.; T. Carpenter, 1981: Variations in tropical SST and surface wind fields associated with the Southern Oscillation / El Nino. - 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 acidental result: The mean 1983 January 500 mb height field significantly different from its 1967-81 predecessors. - Beitr. Phys. Atmos. 57, 440-444.

Storch, H.v ; H. A. Kruse, 1985: The extratropical atmospheric response to El Niño events - a multivariate significance analysis. - Tellus 37A, 361-377

Storch, H.v.; E. Roeckner; U. Cubasch, 1985: Intercomparison of extended range January simulations with general circulation models: Statistical assessment of ensemble properties. - Beitr. Phys. Atmos. 58, 477-498.

Wright, P.B., 1984: Relationships between indices of the Southern Oscillation. - Mon. Wea. Rev. 112, 1913-1919.