Tellus (1984), *36A*, 336–347

EOF analysis of 500 mb geopotential: a comparison between simulation and reality

By J. P. VOLMER, M. DEQUE* and D. ROUSSELET, *Direction de la Météorologie, 77 rue de Sèvres, 92100 Boulogne Billancourt, France*

(Manuscript received November 15, 1982; in final form January 30 1984)

ABSTRACT

Two long-range integrations (6 and 10 years) have been carried out with the European Center for Medium Range Weather Forecasts (ECMWF) spectral model using a T21 truncation. The parameterization of the physical processes in the ground and in the sea-ice differ in the two experiments. No major difference has been found in the mean climate generated by the simulations. In both experiments, the chronology of the atmosphere simulated by the model is studied through principal component analysis of the 500 mb geopotential height field. The two sets of results are compared with each other and with reality.

The differences between the results of the two experiments are mainly in the annual cycle which is, however, correctly described. The non-seasonal signals are also described quite realistically: spatial patterns (blocking phenomena with splitting of the jet stream in winter, Greenland see-saw) as well as their persistence (autocorrelation) and their percentage of variance are found to be very satisfactorily captured.

1. Introduction

During the last 10 years, long integrations have been made with atmospheric models. Some of them span several years, and the long-range behaviour of the simulated atmospheres such as the annual cycle (Déqué and Royer, 1982) and the interannual variability (Manabe and Holloway, 1975; Schlesinger and Gates, 1979) have been studied. This has been done by examining synoptic patterns as in Cubasch (1981) or by using a statistical approach as in Lau (1981) and Volmer et al. (1983). The last paper will be referred to as VDJ in the following. The statistical method used is the principal component analysis (PCA) introduced in meteorology by Fukuoka (1951) and Lorenz (1956).

The ECMWF spectral model was integrated over 6 years in our first experiment and over 10 years in the second one. The parameterization of the physical processes in the ground and in the sea-ice differ in the two experiments. In this paper, we compare the long-range behaviour of the atmosphere in the results obtained from the two experiments with the real atmosphere. As we did not have a several-years sample of global observed data, we used the analysed data of 1980 (from ECMWF) and a 20-year sample of observed data from a limited area, viz. Europe and Atlantic (from Direction de la Météorologie, France).

The paper is organized as follows. The design of the two experiments is described in Section 2. In Section 3, we study the annual cycle on global data; in Section 4, the main non-seasonal signals on global data after subtracting the mean year are investigated. The annual cycle on the limited area is examined in Section 5, and Section 6 describes the analysis of the main non-seasonal signals on the limited area.

2. Description of the experiments

We have used a 15 level, T 21 (i.e. truncated up to zonal and meridional wavenumber 21), sigma coordinate version of the ECMWF global spectral

*Météorologie Nationale, EERM/CNRM, Avenue Coriolis, 31057 Toulouse Cedex, France.

model. The headlines of this version and the choice of the truncation are developed in VDJ.

In the first experiment, as a result of a coding error, the surface temperature over frozen sea (not predicted) was not updated and the sea ice was not allowed to melt thereafter. Moreover, the temperature and the moisture in the deep soil, used for fluxes in the ground when computing the surface balance were kept constant over land at the November climatological values. Some boundary conditions that should have had a seasonal variation were thus constant throughout the year. In the second experiment, these deficiencies were corrected. Furthermore, the representation of mountains differs slightly in the two experiments. In the first experiment, we used a spectral orography from an old Geophysical Fluid Dynamics Laboratory (GFDL) numerical prediction model. In the second experiment, we used a more accurate orography, calculated from a regular latitude longitude grid (48 latitudes, 96 longitudes) with a spectral smoothing. The difference can be as much as several hundred meters over the Antarctic and Greenland.

The starting date for the first experiment was 15 November, 1979, and 11 November, 1979 for the second. The model was run until 1 January, 1980, that is for 46 and 50 days, respectively, in order that the initial conditions could be ignored. In the first experiment, a 6-year simulation was carried out and the results were stored every second day of the integration. In the second experiment, a 10-year simulation was carried out and the results were stored every day. The first experiment will be hereafter referred to as E1, the second as E2. The climatologies (i.e. mean January and mean July meteorological fields) generated by the model are similar and presented in VDJ.

In this paper, we shall concentrate on the time evolution of the large-scale characteristics of the 500 mb geopotential height field (referred to as Z in the following), which describes the main features of the general atmospheric circulation. This field is given by its spectral coefficients and can be expanded in:

$$Z(x,t) = \dot{Z}(\bar{x}) + Z_1(t)\,C_1(x) + \cdots + Z_k(t)\,C_k(x) + R_k(x,t), \quad (1)$$

where x is the spatial coordinate, t is the time, and $\dot{Z}(x)$ is the time average of Z. The set of functions C_i is chosen to ensure the smallest possible mean squared truncation error ($R_k(x, t)$). The functions C_i are called empirical orthogonal functions (EOFs) and are calculated by principal components analysis (PCA), and the coefficients $Z_i(t)$ are called principal components (PCs). For further details, see VDJ.

3. PCA on the global data

We performed PCAs on the 6 years of E1 and on the 10 years of E2. For E2, we also performed a PCA on the first 6 years, where the data every second day were ignored in order to attain the same conditions as in E1. As expected, neither the maps of the EOFs, nor the variances of the PCA were changed significantly. As can be seen in Fig. 1, the first EOFs (C_1) show a contrast between northern (NH) and southern (SH) hemisphere. The time evolutions of the corresponding PCs are close to a sinusoidal curve with periods of 1 year and phases of 1 month delayed with respect to the solar cycle. We also performed PCAs on a single year for E1 and E2, and noticed a remarkable stability, at least for C_1. Since we had the ECMWF analysis for the whole year 1980, we compared the first EOFs (C_1) obtained in E1, E2 and the real data. This allowed us to measure how well the model is able to simulate the annual cycle (for 1980 and E2 see Fig. 1; for E1 see VDJ).

In Table 1, the values of the total variances and of the variances of Z_1 for the real data, E1 and E2 are displayed. We can see that the variance of Z_1 is larger in E2 than in E1, but the variance of the residual components is the same. The amplitude of the annual cycle of the 500 mb height field is improved when deep soil temperatures and moistures and sea ice extent follow a seasonal cycle. This can be verified if we look at mean January minus mean July maps of the 850 mb temperature field (not shown): the large features of this field look like those of C_1 and we have a climatology available. We can see that the contrast over continents is better in E2 than in E1. The delay of 1 month of the 500 mb height cycle with respect to the solar cycle is correct. These results show that the annual cycle is satisfactorily taken into account by the model.

It is now of interest to study the non-seasonal

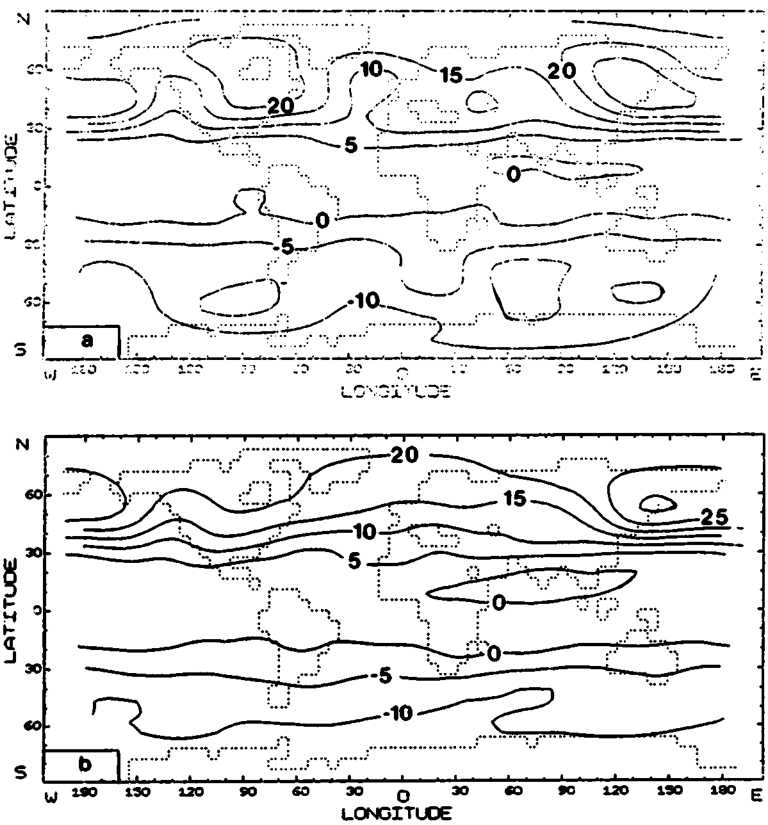


Fig. 1. Distribution of C_1 observed in 1980 (a), simulated in E2 (b).

Table 1. *Total variance, variance of the first PC and residual variance for global raw data for the year 1980 (observed), E1 and E2 (simulated)*

	1980	E2	E1
total variance (m^2)	11,100	8,500	6,600
variance of Z_1 (m^2)	5,420	5,130	3,230
residual variance (m^2)	5,580	3,370	3,370

signals. In order to achieve this, we have separated the annual cycle from the other signals by subtracting the mean year Z^M from Z as described in VDJ (for E1 we use the mean year of E1, and for E2 we use the mean year of E2).

4. PCA after subtracting the mean year

We thus performed a PCA on the field $Z' = Z - Z^M$. As could be expected, the information is less confined in the first components. The first two PCs are degenerate and have a percentage of variance of 4.5% (see appendix A of VDJ for the separation techniques). Fig. 2 shows the corresponding EOFs C'_1 and C'_2. The maps are very similar to those obtained in El as can be seen in VDJ. There is a decoupling between the NH and the SH and both EOFs are related to the increase or the decrease of the pole to equator gradient.

The variances of C'_1 and C'_2 in both E1 and E2 have practically the same value, 120 m^2. If the amplitude of those waves is not changed, their persistences are different. E2 provides a more persistent 2nd PC than E1, as shown in Fig. 3. This second EOF, which is located in the NH may be caused by thermal forcings (over the continents and the sea ice) or orographic forcings (for instance over Greenland). These forcings have been modified in E2. Moreover, the weak senescence effect found in VDJ does not appear in E2. In VDJ, we related some episodes which were anomalous in their amplitudes or their durations. In E2, the

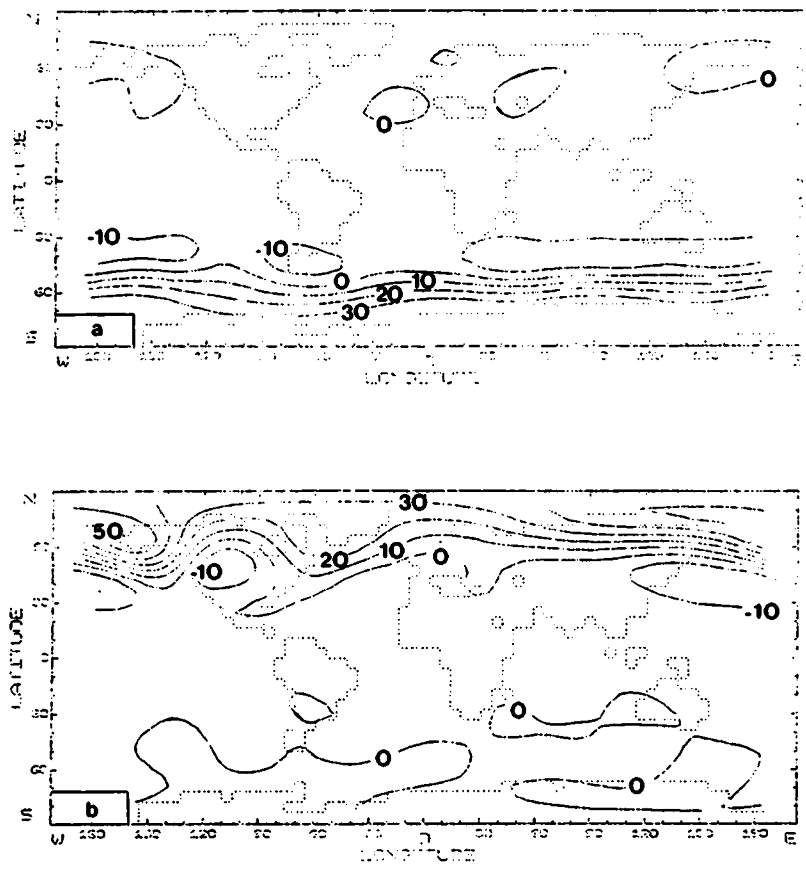


Fig. 2. Distribution of C'_1 (a) and C'_2 (b) simulated in E2.

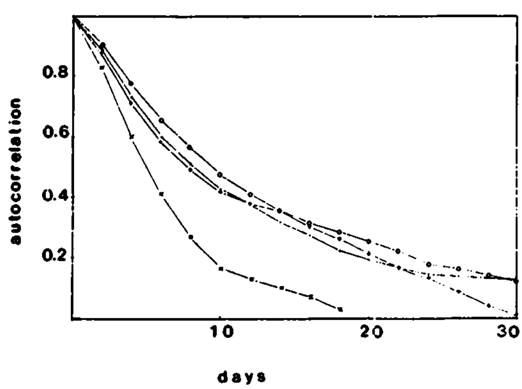


Fig. 3. Autocorrelation for lags up to 30 days (PCA after subtracting the mean year). ·——·: C'_1 simulated in E1; x——x: C'_2 simulated in E1; O——O: C'_1 simulated in E2; +——+: C'_2 simulated in E2.

longest episode with a constant sign for a component lasted for 116 days for Z'_2 from 9 September year 8 to 2 January year 9. It was thus an event in the NH.

5. PCA on data in a limited area

We had at our disposal a 20-year sample (from 1961 to 1980) of observed values of Z over a limited area. This area includes most of the eastern coast of North America, the North Atlantic and Europe. The data are given on an 89-point grid. Most of the points correspond to stations of the radiosonde network. This grid is completed by some points over the Atlantic Ocean where the values have been interpolated. In E1, the total

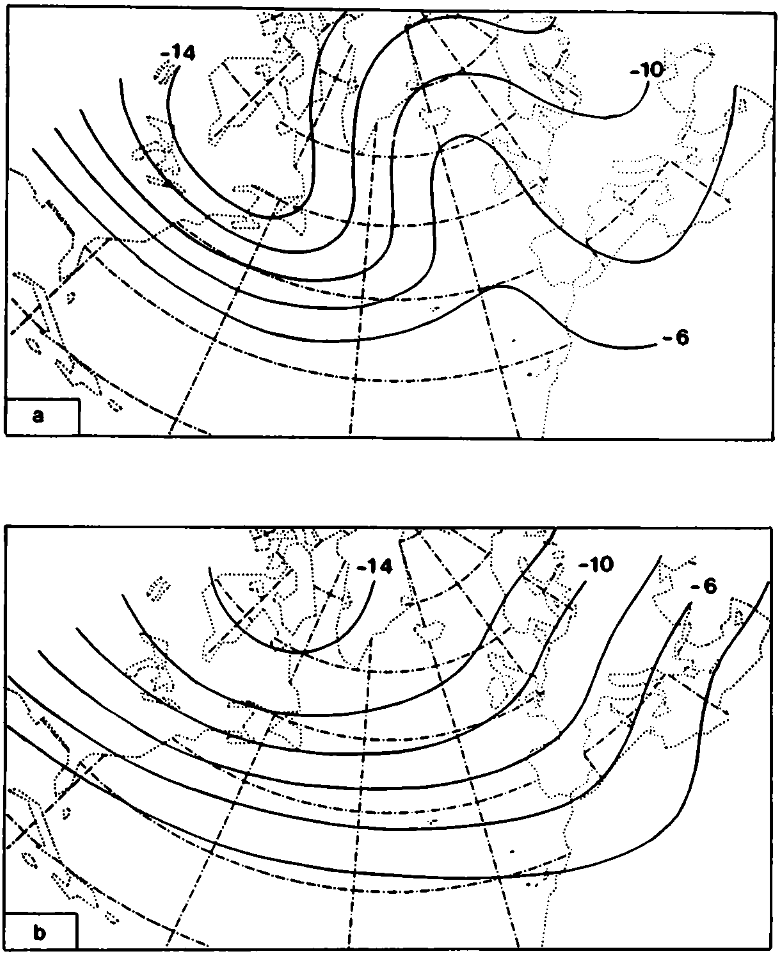


Fig. 4. Distribution of C_1 observed over 20 years (a), simulated in E2 (b).

variance is half that of the real one; in E2 it is two thirds. We first performed a PCA on the raw values. The first EOF (Fig. 4) looks like the first EOF obtained on the global data (Fig. 1) if we consider only the Atlantic and Europe area. The percentage of variance for Z_1 is 52%, 66% and 51% for the real data, E2 and E1, respectively. The increase of total variance in E2 has been exclusively absorbed in the annual cycle. The time evolution of Z_1 is quasi-sine shaped. The other EOFs mainly describe the non-seasonal effects and will be studied in section 6.

We then studied the mean year which is defined by 365 or 366 averages over N years of the height field, $N = 20$ for the real data set, $N = 6$ for E1 and $N = 10$ for E2. We performed a PCA on the mean year. As could be expected, the information is mainly contained in the first component. The percentages of variance of Z_1^Y are 93% for the real data, 94% for E2 and 85% for E1. The corresponding EOFs are quasi-identical to those obtained on the raw data (Fig. 4). The second PC explains 2% of the total variance. However, the time evolution of Z_2^Y remains a quasi-sinusoid in quadrature with Z_1^M. This shows that C_2^M is still significant.

Fig. 5 shows the map of C_2^Y. As we can see, they describe a contrast between high and middle

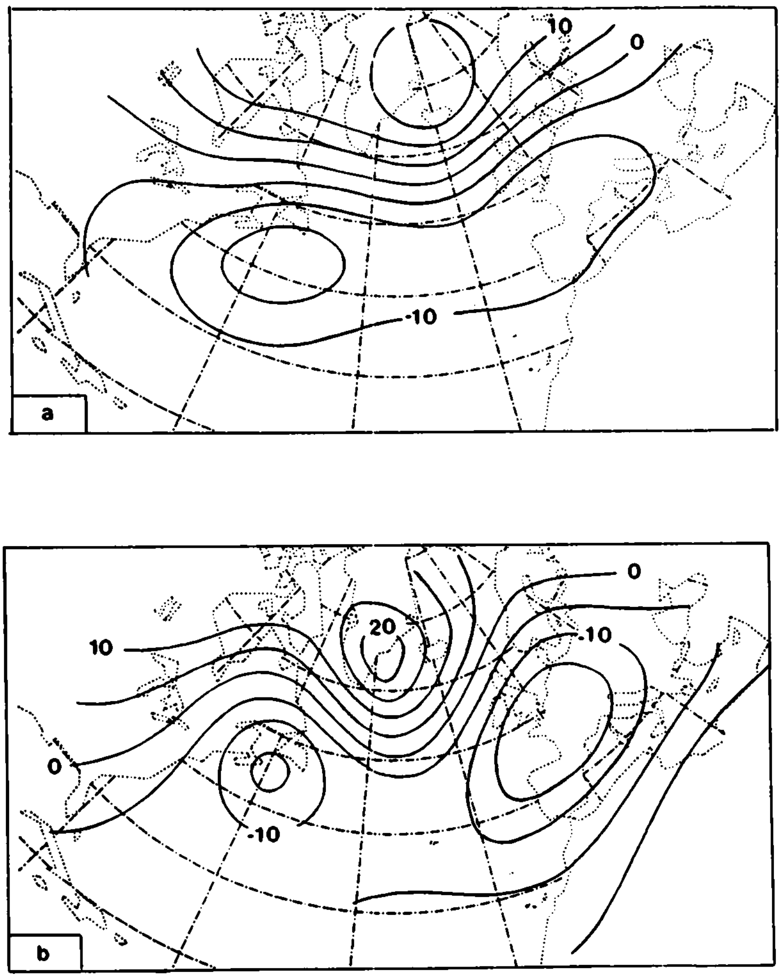


Fig. 5. Distribution of C_2^V observed over 20 years (a), simulated in E2 (b).

latitudes. To interpret C_2^V, we approximate Z^V with the first two EOFs as follows:

$$Z^V(x,t) = \bar{Z}(x) + a_1 \cos \omega(t - t_0) \cdot C_1^V(x) + a_2 \sin \omega(t - t_0) \cdot C_2^M(x), \quad (3)$$

where $\omega = 2\pi/365$, $t_0 = 20$ (time is expressed in days). Because the whole variance is contained in Z_1^V, a_2 is very small if compared to a_1. Hence, for each point x we have:

$$Z^V(x,t) = \bar{Z}(x) + A \cos(\omega(t - t_0) + \varphi), \quad (4)$$

where

$$\varphi \simeq \tan \varphi = \frac{a_2 C_2^M(x)}{a_1 C_1^M(x)}, \quad A \simeq a_1 C_1^V(x). \quad (5)$$

We can see that C_1^V describes the spatial distribution of the amplitude of the annual cycle, whereas C_2^V describes its phases (at least the sign of its phase, $C_1^V(x)$ being everywhere positive). This can be seen in Fig. 6 which displays the maps of amplitude and phase for the real data. We find good agreement for both E1 and E2.

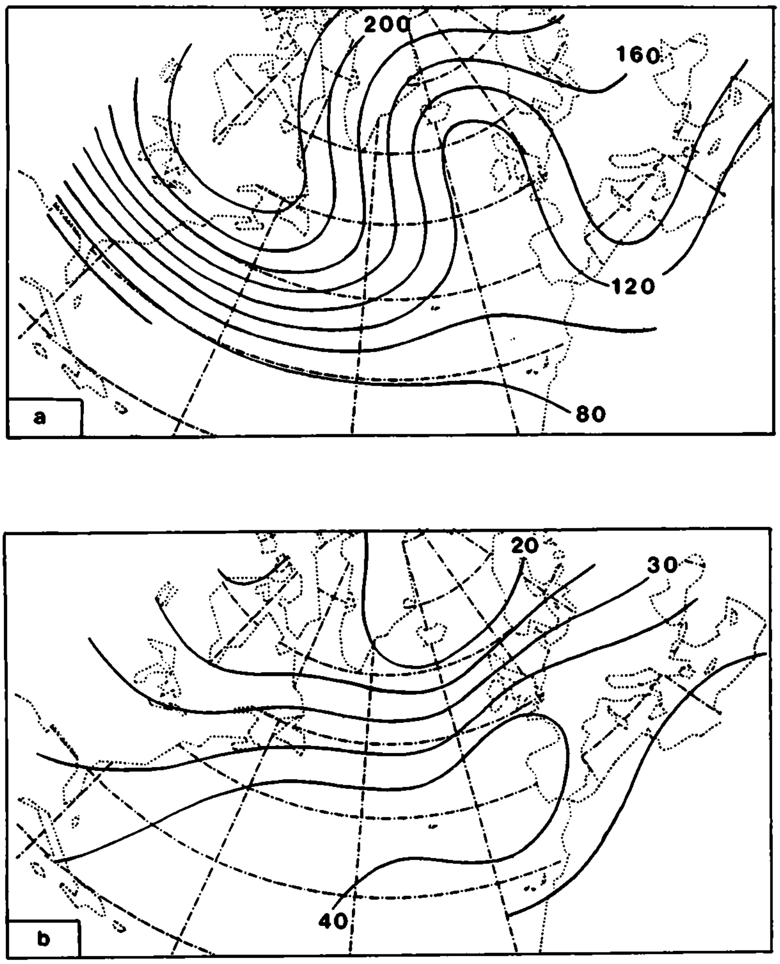


Fig. 6. Distribution of amplitude (a) (contour interval 20 m), and of phase (b) (contour interval 5 days) of the observed annual cycle of 500 mb geopotential height.

6. PCA after subtracting the mean year (limited domain)

In order to study the non-seasonal signals, we used the same method as in Section 4 for the global data. We subtracted the mean year from the raw values $Z' = Z - Z^M$. We carried out a PCA on this sample for E1, E2 and the real data set. We note a good resemblance between C'_1, C'_2, C'_3 (EOFs of this new data set) and C_2, C_3, C_4, (EOFs of the raw data set presented in the last section), respectively. The first EOF, C'_1, is shown in Fig. 7. Its main feature is a contrast between a minimum centred on Greenland and a maximum centred on western Europe. This pattern is known in the literature as the "Greenland see-saw" (Van Loon and Rogers, 1978). As Z^M is mostly zonal, when Z'_1 is positive, there is a low over Greenland and a high over Europe. This implies a northward circulation over western Europe with warm oceanic air and a geopotential gradient stronger than usual. When it is negative, western Europe receives cold air from the Arctic regions. The effect of C'_1 is almost the same in summer as in winter, but in winter, the

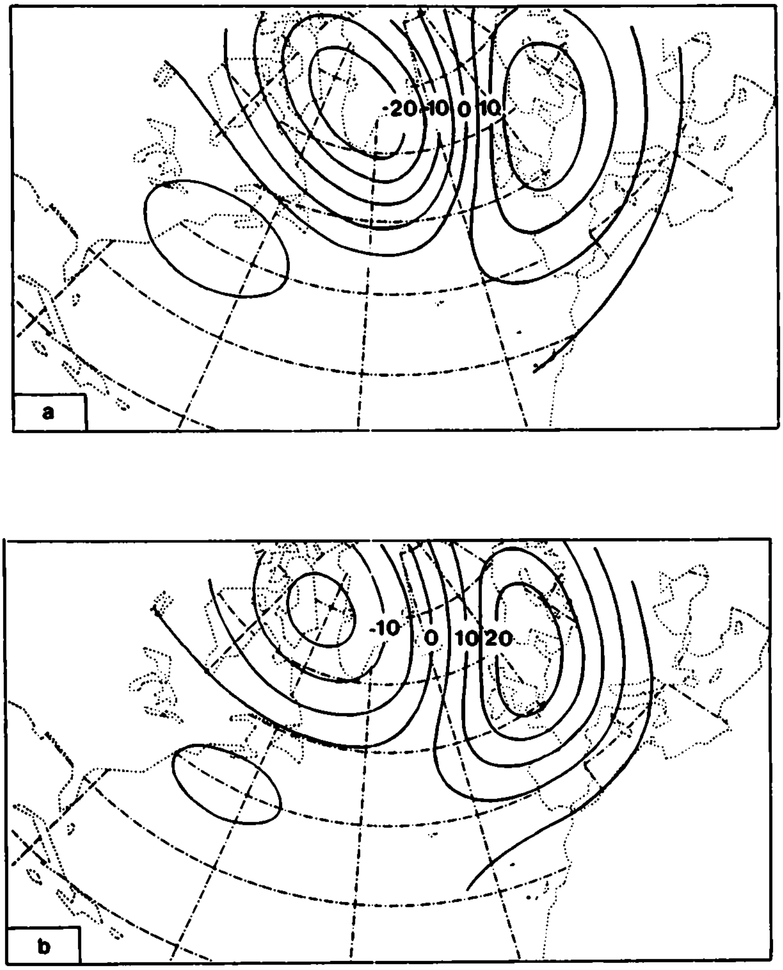


Fig. 7. Distribution of C_1' observed over 20 years (a), simulated in E2 (b).

variance of Z_1' is twice as large. The resemblance of C_1 in E2 and in the real sample is very satisfactory. In E1 (not shown) the minimum over Greenland was underestimated.

Fig. 8 shows the map of C_2'. It includes a north–south contrast which involves a displacement of the westerlies similar to the effect of the first two EOFs on global data without annual cycle. When Z_2' is positive, the gradient is stronger than usual over England. When it is negative the flow is split over northern Europe and a maximum of geopotential height appears over Norway.

The model reproduces correctly the north–south contrast, but it fails to accurately locate the centre of the negative structure which takes place over Greenland in the model (both E1 and E2) instead of the Norwegian Sea. Fig. 9 and Fig. 10 show the map of the third and fourth EOFs.

The patterns obtained in E2 look closely like the real ones. Those obtained in E1 (not shown) are quite different, which indicates that the resemblance between these patterns is not induced by the method used.

In Table 2, the values of the percentage of

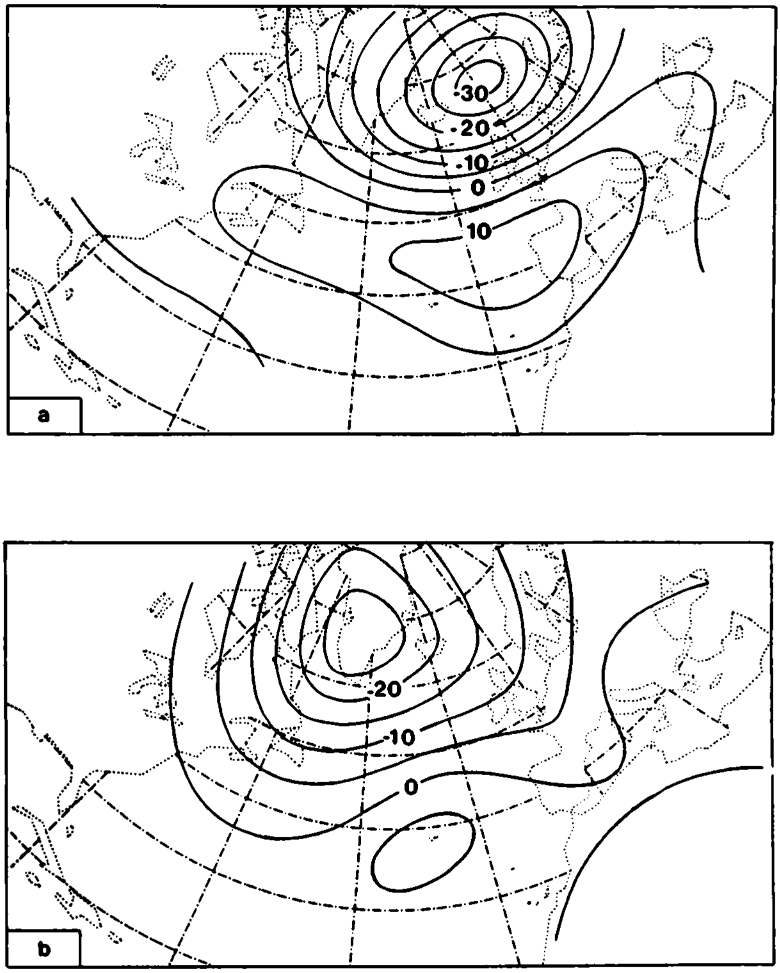


Fig. 8. Distribution of C_2' observed over 20 years (a), simulated in E2 (b).

variance for both E2 and the real atmosphere are displayed. The amount of information contained in the first 4 signals is larger for the real atmosphere than for the simulation. Table 2 also shows the values of the autocorrelations computed at a lag of 2 days and 6 days. These values indicate the persistences of the accidents corresponding to the EOFs. The results of E2 are consistent with the real atmosphere. The first PCs are persistent enough to produce the following effect. If we perform a PCA on 5-day, 10-day or 20-day averages, we get the same EOFs with a larger percentage of variance, because averaging filters the PCs of higher order.

7. Summary and concluding remarks

The PCA on raw global data has shown the large impact of a change in the parameterization of the deep soil and the sea ice boundary conditions. The changes between E1 and E2 have brought about a drastic improvement in the global vari-

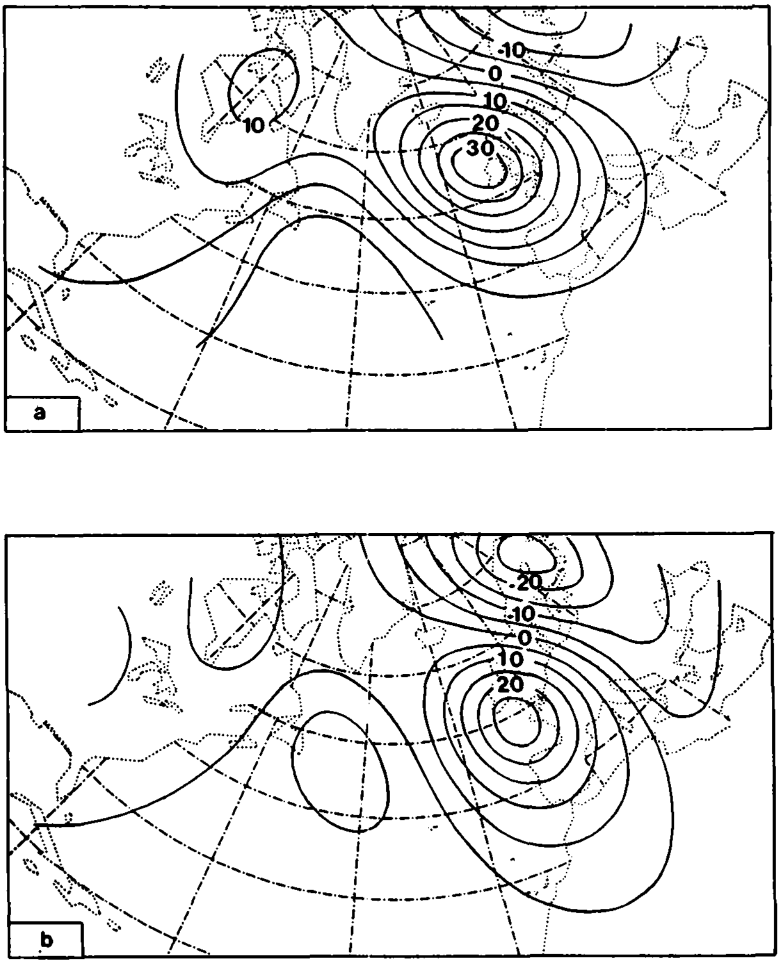
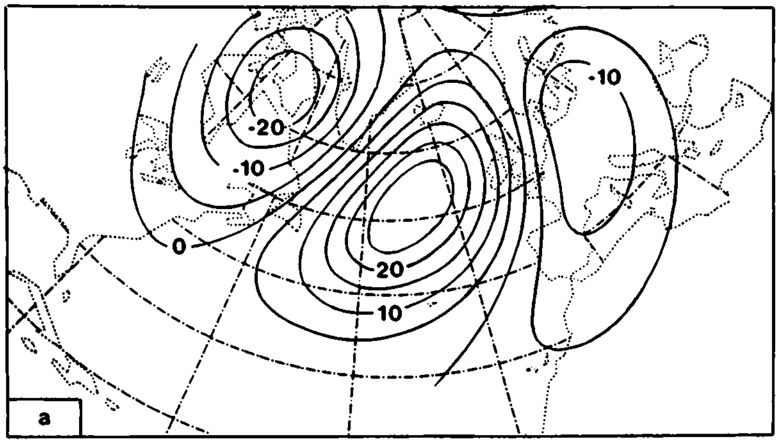


Fig. 9. Distribution of C'_1 observed over 20 years (a), simulated in E2 (b).

ance. As these changes consist mostly of introducing a seasonal variation in these processes, the improvement occurred as expected mainly in the annual cycle. This result is confirmed by PCA on observed and simulated data for the Europe Atlantic domain. The annual cycle is the major component of the long-range behaviour of the atmosphere. Almost the whole annual cycle is explained by only one sinusoidal function. The other functions are mostly responsible for the varying phase shifts in the different areas.

The PCAs made on the data after subtracting the mean year show good agreement between reality and simulation, especially for E2. The first EOF is related to the "Greenland see-saw", the second to blocking over western Europe. The corresponding PCs are rather persistent and express long-range signals. Since the model does not take long-range non-seasonal variations of external forcings, such as the sea surface temperature and the spatial coverage of sea ice into account, it is not surprising to find a lack of variance in the non-seasonal phenomena (about one half).

Investigations are presently being carried out by modellers to couple the sea surface temperature to upper air conditions. These non-seasonal long-range signals, which are responsible for a large part of the inter-annual variability, are present in the simulation.



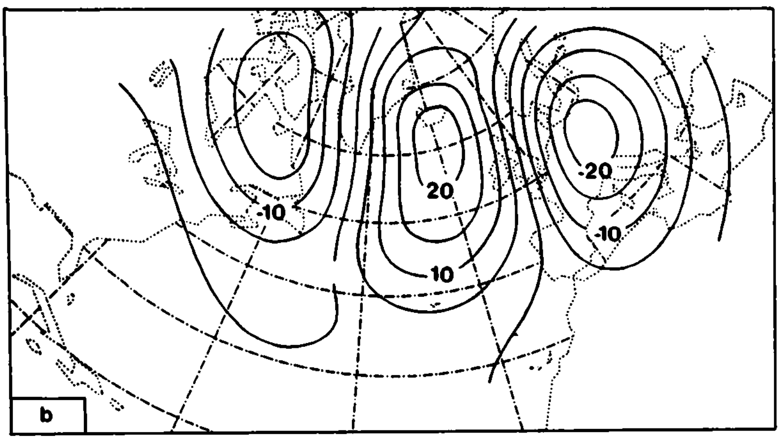


Fig. 10. Distribution of C'_4 observed over 20 years (a), simulated in E2 (b).

Table 2. *Percentage of variance, 2-day and 6-day autocorrelations of the first four PCs on Europe-Atlantic after removing the mean year for reality and E2*

	Percentage of variance		2-day autocorrelation		6-day autocorrelation	
	reality (%)	E2 (%)	reality	E2	reality	E2
Z'_1	16.3	12.6	0.74	0.69	0.27	0.27
Z'_2	12.5	12.2	0.76	0.80	0.25	0.40
Z'_3	10.2	7.5	0.71	0.61	0.20	0.08
Z'_4	8.6	7.3	0.71	0.64	0.29	0.16

8. Acknowledgements

Special thanks are due to U. Cubasch whose contribution in running the experiments was decisive. We are grateful to L. Bengtsson and J. Lepas for their encouragement and support. We had also stimulating discussions with M. Jarraud and J. F. Geleyn.

REFERENCES

Cubasch, U. 1981. Preliminary assessment of long-range integrations performed with the ECMWF global model. *ECMWF Technical memorandum* no. 28. (Available from ECMWF, Shinfield Park, Reading, Berks, RG2 9AX, England.)

Déqué, M. and Royer, J. F. 1982. Modélisation du cycle annuel à l'échelle globale. Etude de janvier et de juillet. *Note de travail de l'EERM no. 35.* (Available from Direction de la Météorologie, EERM, 77 rue de Sèvres, 92100 Boulogne, France.)

Fukuoka, A. 1951. A study of 10-day forecasts (a synthetic report). *Geophys. Mag. 22*, 117–208.

Lau, N. C. 1981. A diagnostic study of recurrent meteorological anomalies appearing in a 15-year simulation with a GFDL general circulation model. *Mon. Wea. Rev. 109*, 2287–2311.

Lorenz, E. N. 1956. Empirical orthogonal functions and statistical weather prediction. *Scientific report no. 1.* Dept of Meteorology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

Manabe, S. and Holloway, J. L., Jr. 1975. The seasonal variation of the hydrologic cycle as simulated by a global model of the atmosphere. *J. Geophys. Res. 80*, 1617–1649.

Schlesinger, M. E. and Gates, W. L. 1979. Performance of the Oregon State University two-level atmospheric general circulation model. *GARP Publ. Serv. no. 22*, WMO/ICSU, Geneva, Switzerland.

Van Loon, H. and Rogers, J. C. 1978. The see-saw in winter temperature between Greenland and northern Europe. *Mon. Wea. Rev. 106*, 296–316.

Volmer, J. P., Déqué, M. and Jarraud, M. 1983. Large-scale fluctuations in a long-range integration of the ECMWF spectral model. *Tellus 35A*, 173–188.