Eigenvector Analysis of Some Observed and Model-Generated Climatological Fields

Michael Dennis and Jerome Spar

The City College

July 1980

Technical report, Grant NGR 33-013-086, NASA Goddard Space Flight Center. This research was conducted at the Goddard Institute for Space Studies (GISS).

Introduction

The decomposition of two-dimensional synoptic fields into principal components is a well-known application of the method of finding the characteristic values (eigenvalues) and characteristic vectors (eigenvectors) of a matrix, and is generally referred to in the meteorological literature as empirical orthogonal function (EOF) analysis. (See, e.g., Lorenz, 1956; Kutzbach, 1967; Kidson, 1975.) EOF analysis has been used mainly to compact very large masses of synoptic meteorological data into tractable form for time series and correlation studies and for statistical weather prediction. The method may also be used to compare synoptic maps. In this study it is applied to some of the monthly mean fields computed with a general circulation model and to the corresponding observed climatological patterns.

An efficient computer program for the calculation of the eigenvectors by the method of asymptotic singular decomposition (Jalickee and Kiepczynski, 1977), which was provided by J.B. Jalickee of NOAA, has been used to compute the principal components of the data matrices representing, on an 8° x 10° latitude-longitude grid, global maps of various observed and model-generated climatological fields. The variables analyzed are 500 mb geopotential height (G500), 850 mb temperature (T850), and sea-level pressure (SLP). The observed climatology was obtained from the National Center for Atmospheric Research (NCAR), and the model-generated climatology is from a 5-year simulation run with one version (No. 540) of the coarse-mesh

(8° x 10°) GISS climate model (Hansen, et al., 1980). It is assumed that the two climatologies are indeed comparable, i.e., that they are both representative of stable long-period averages. Results are presented below for the month of January.

Method

Singular decomposition of an M x N data matrix, $\underline{A}$, of rank L ($\leq$ M) may be written:

$$A_{ij} = a_1 f_{1i} g_{1j} + a_2 f_{2i} g_{2j} + \text{---} + a_L f_{Li} g_{Lj} ,$$

where the coefficients, a, are square roots of the eigenvalues of the symmetric square matrices, $\underline{A}\,\underline{A}^T$ and $\underline{A}^T\,\underline{A}$, whose eigenvectors are f and g, respectively, the latter being left and right orthonormal singular vectors of $\underline{A}$. Each vector product, $f_{ri} g_{rj}$, yields an M x N matrix, which, when multiplied by its singular value, a_r, represents one component of the data field. The coefficients, a, and hence the component matrices, are arranged in descending order of magnitude, and the expansion is optimal in the sense of least squares. The method of computation is a generalization of the power method in which an iterative correction of a first guess is calculated in conjunction with Gram-Schmidt orthonormalization (Jalickee and Klepczynski, 1977).

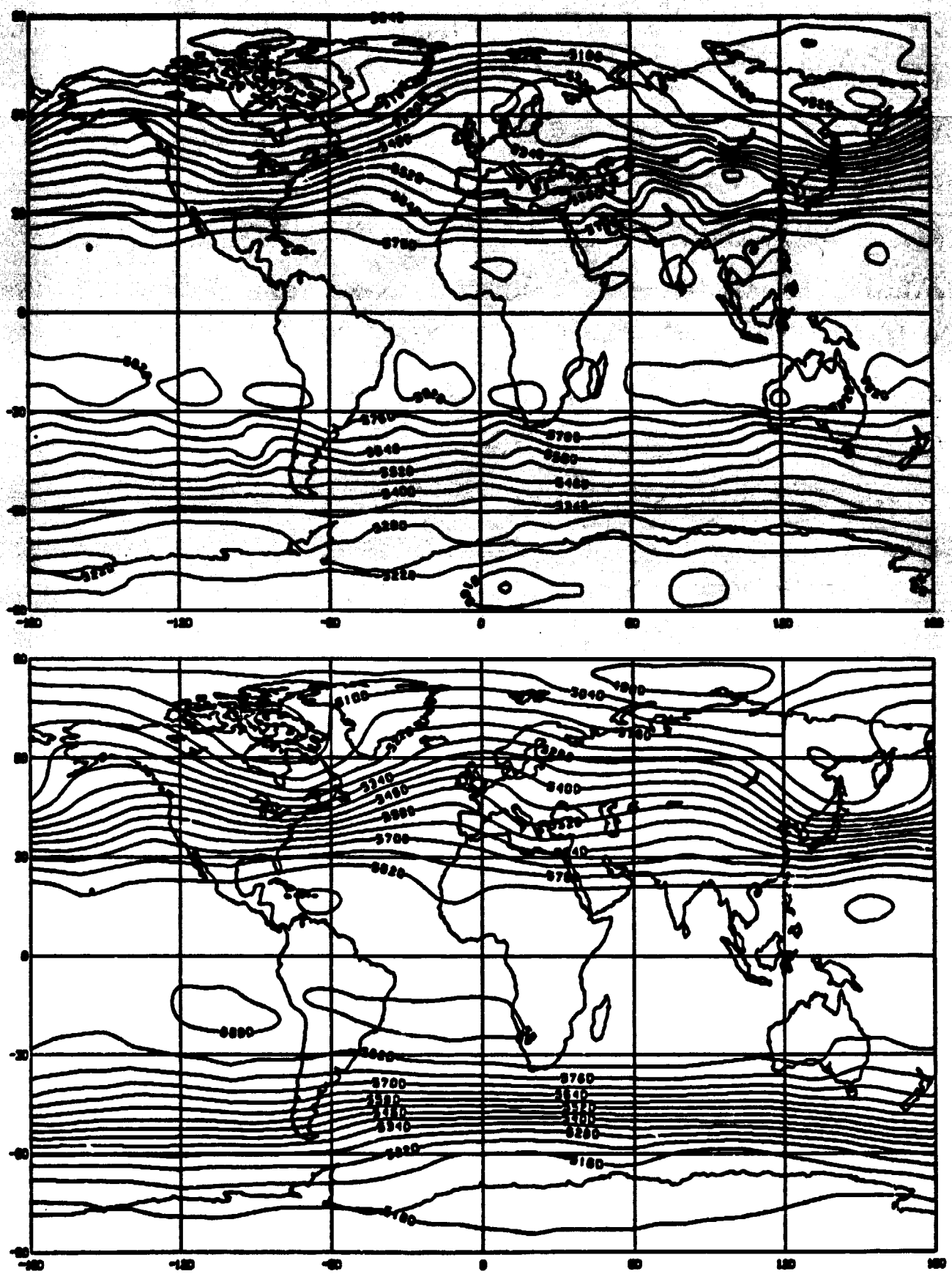

fig 1a. 500 mb. geopotential height. Model generated climatology (top), observed climatology (bottom).

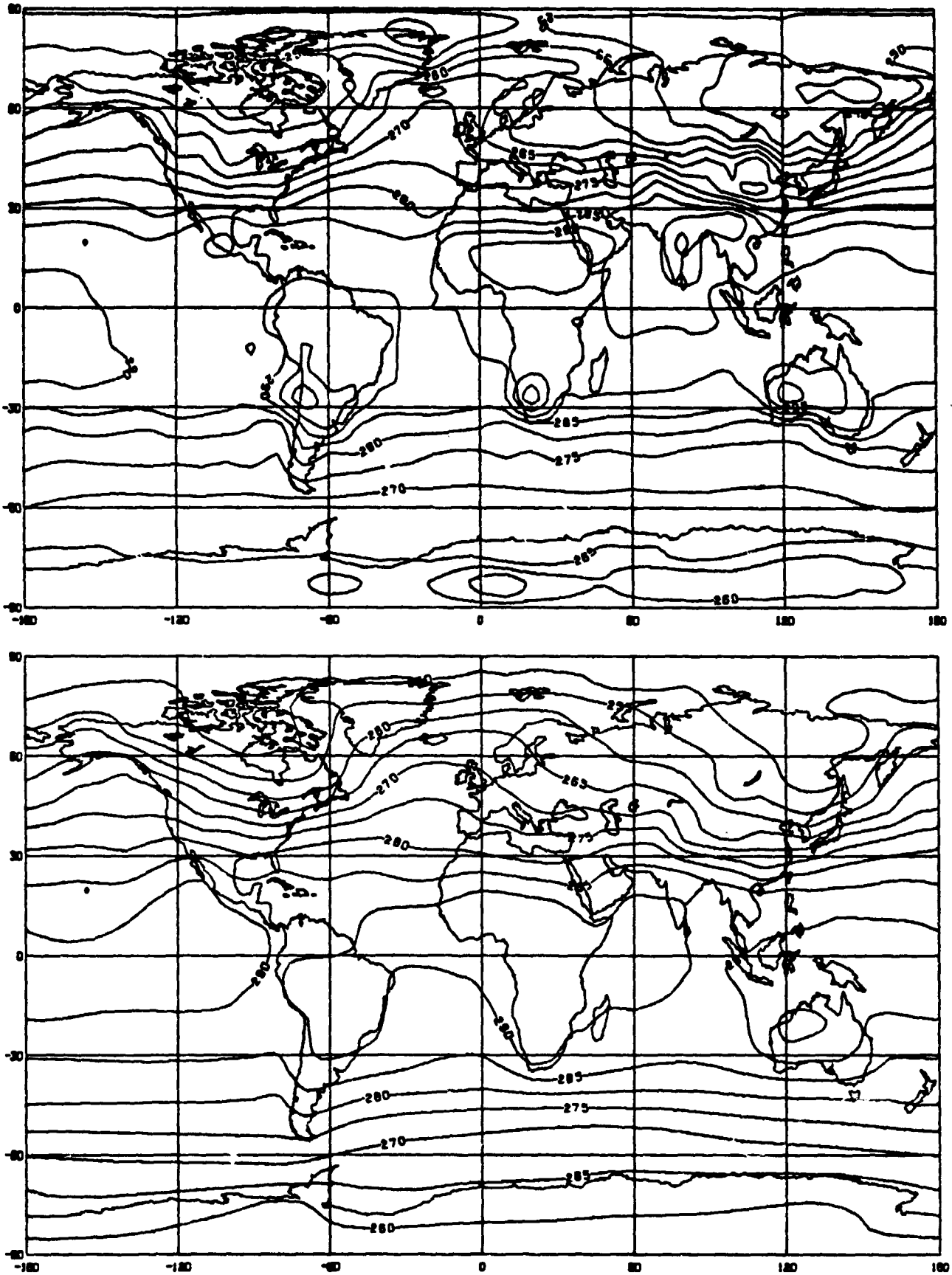

fig 1b. 850 mb. temperature (Kelvin). Model generated climatology (top), observed climatology (bottom).

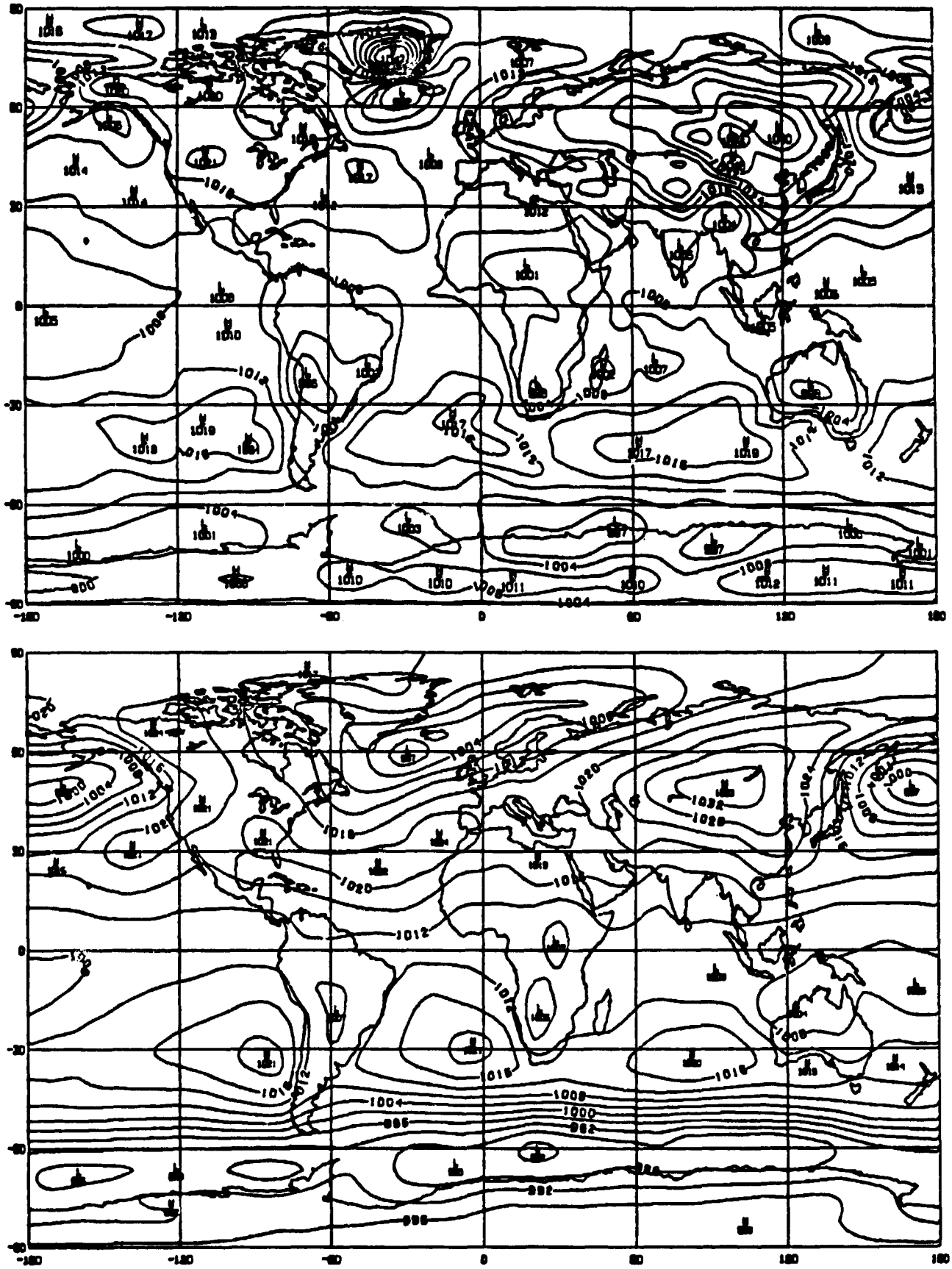

fig 1c. Sea level pressure (mb). Model generated climatology (top), observed climatology (bottom).

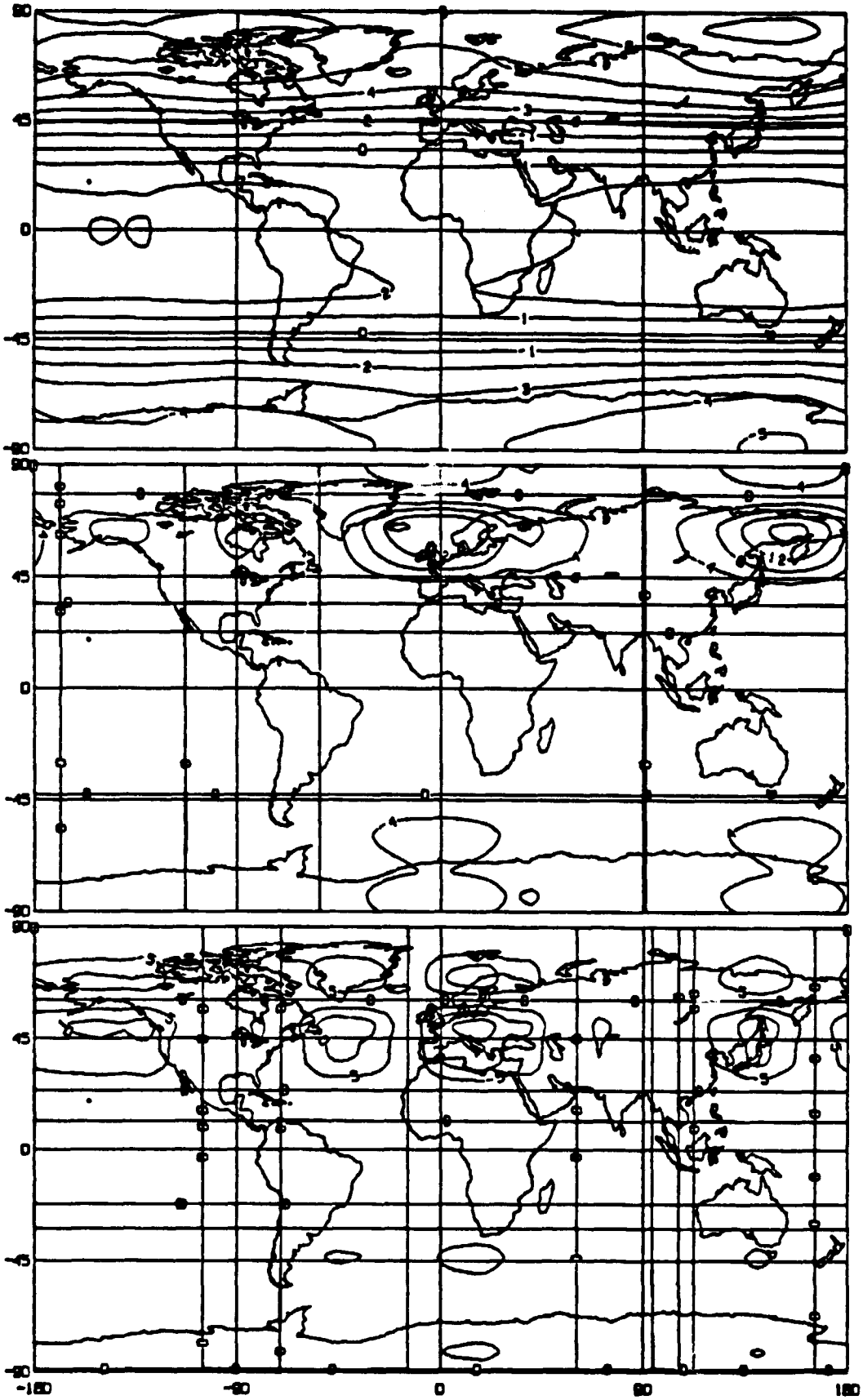

fig 2. First, second and third components (top to bottom) of model climatological 500 mb. geopotential height.

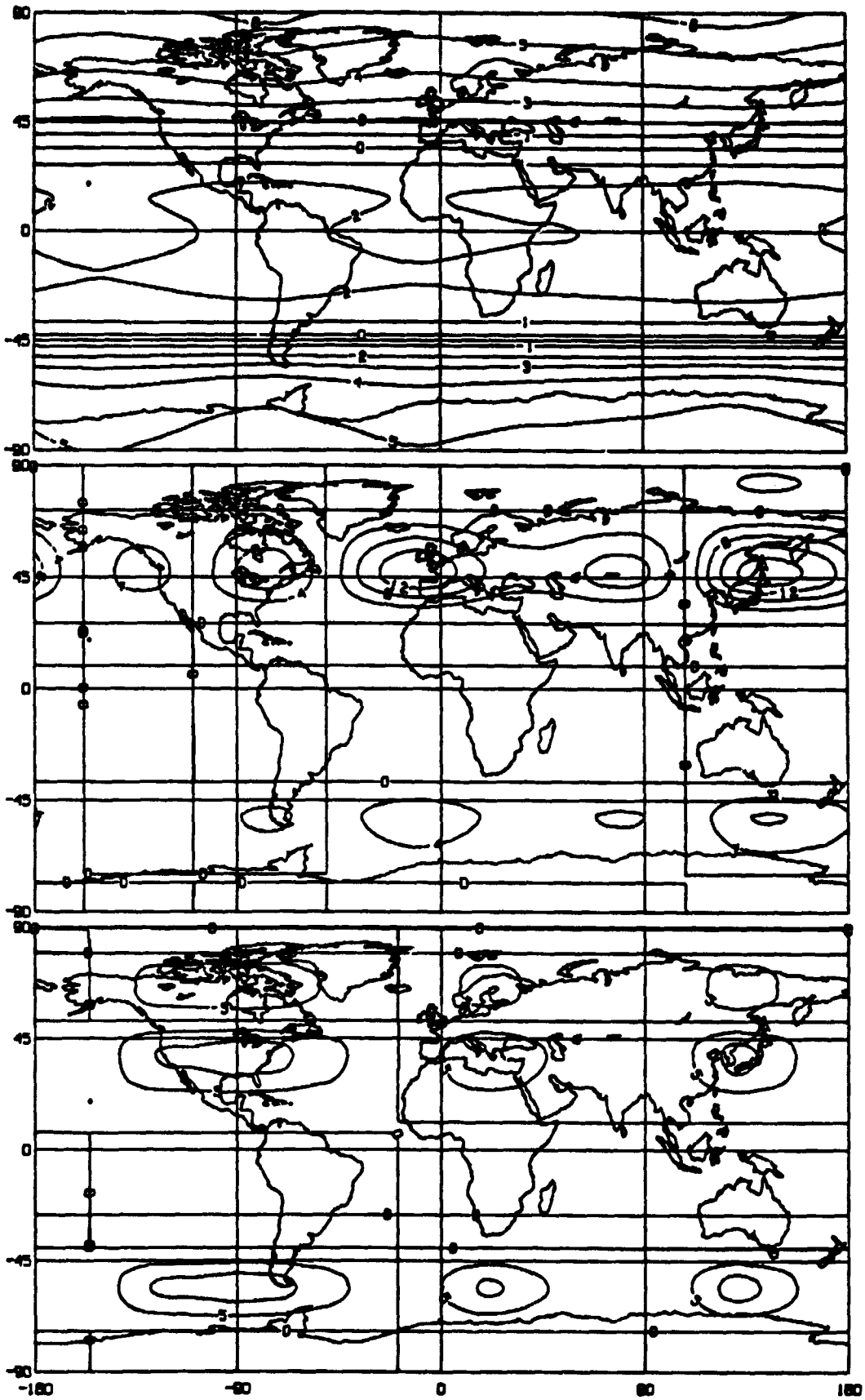

fig 3. First, second and third components (top to bottom) of observed climatological 500 mb. geopotential height.

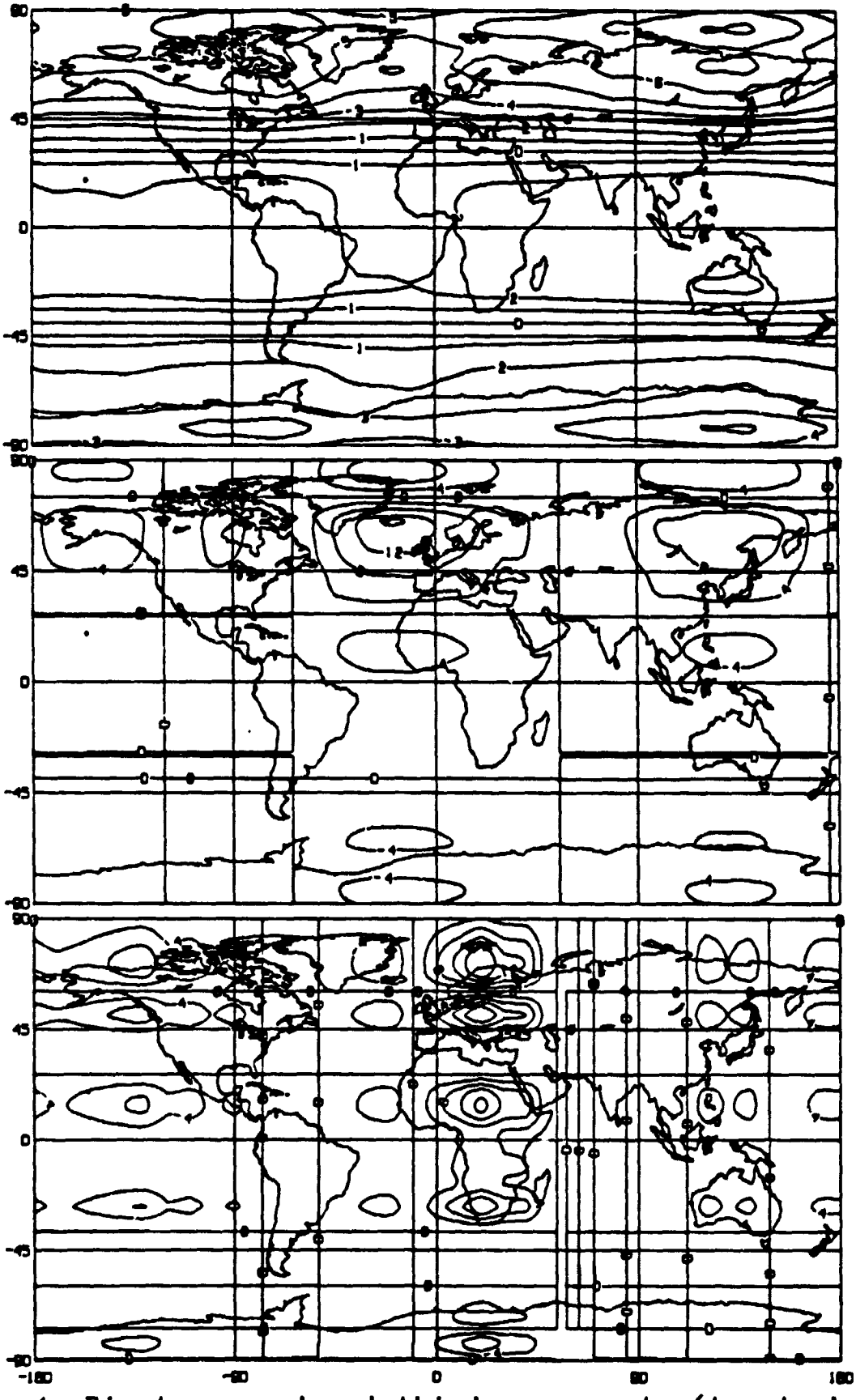

fig 4. First, second and third components (top to bottom) of model climatological 850 mb. temperature

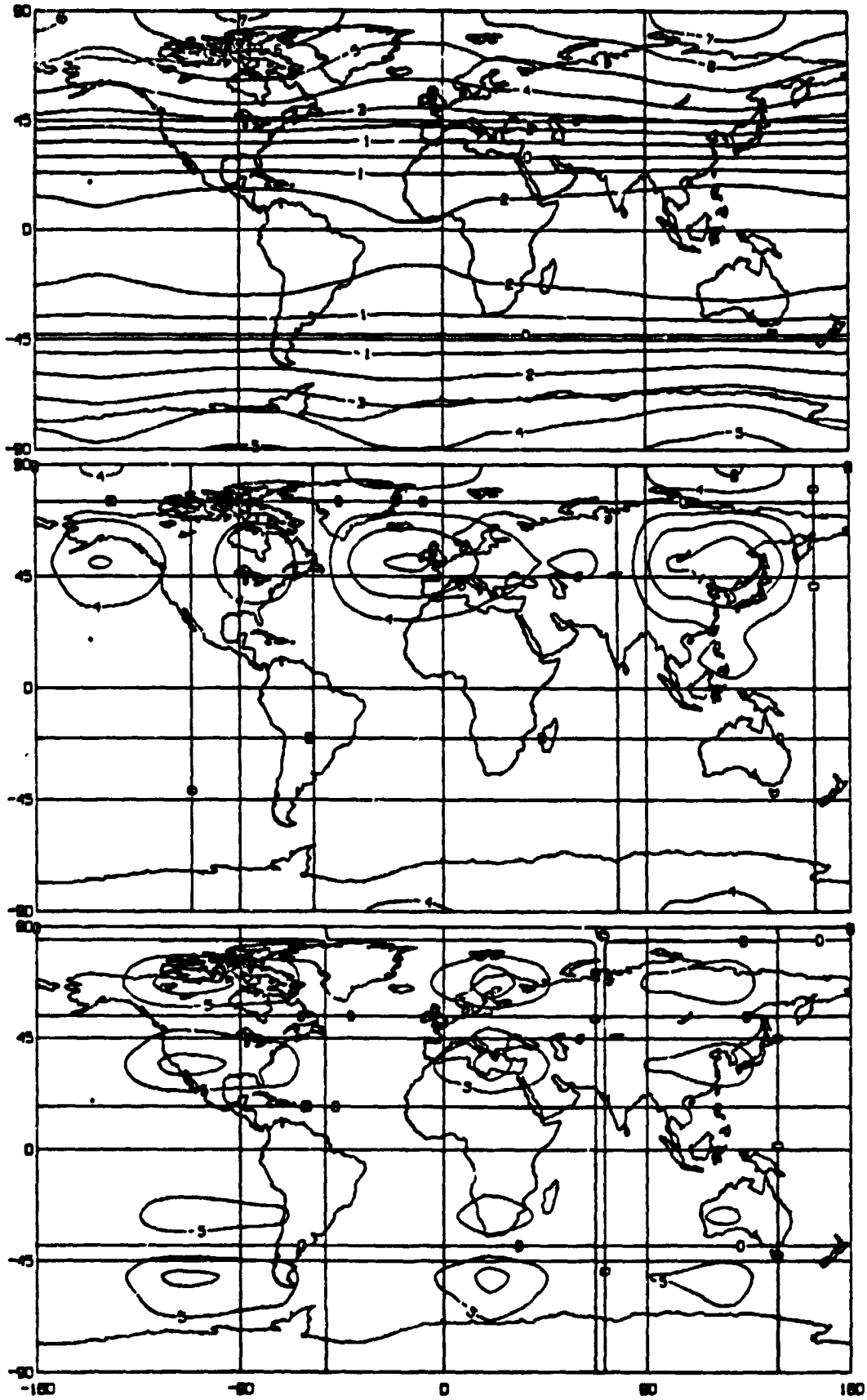

fig 5. First, second and third components (top to bottom) of observed climatological 850 mb. temperature

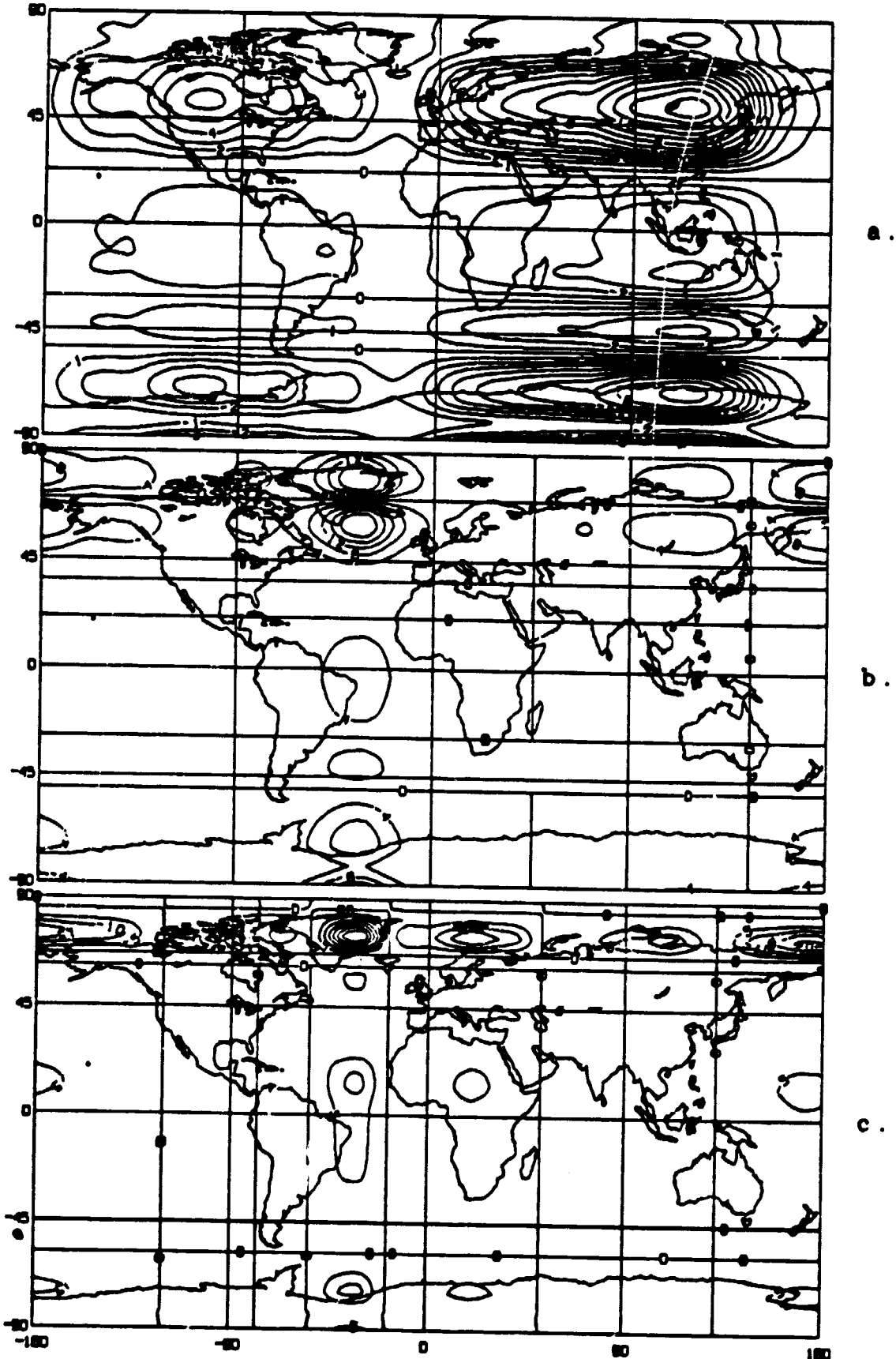

fig 6. First, second and third components (top to bottom) of model climatological sea level pressure

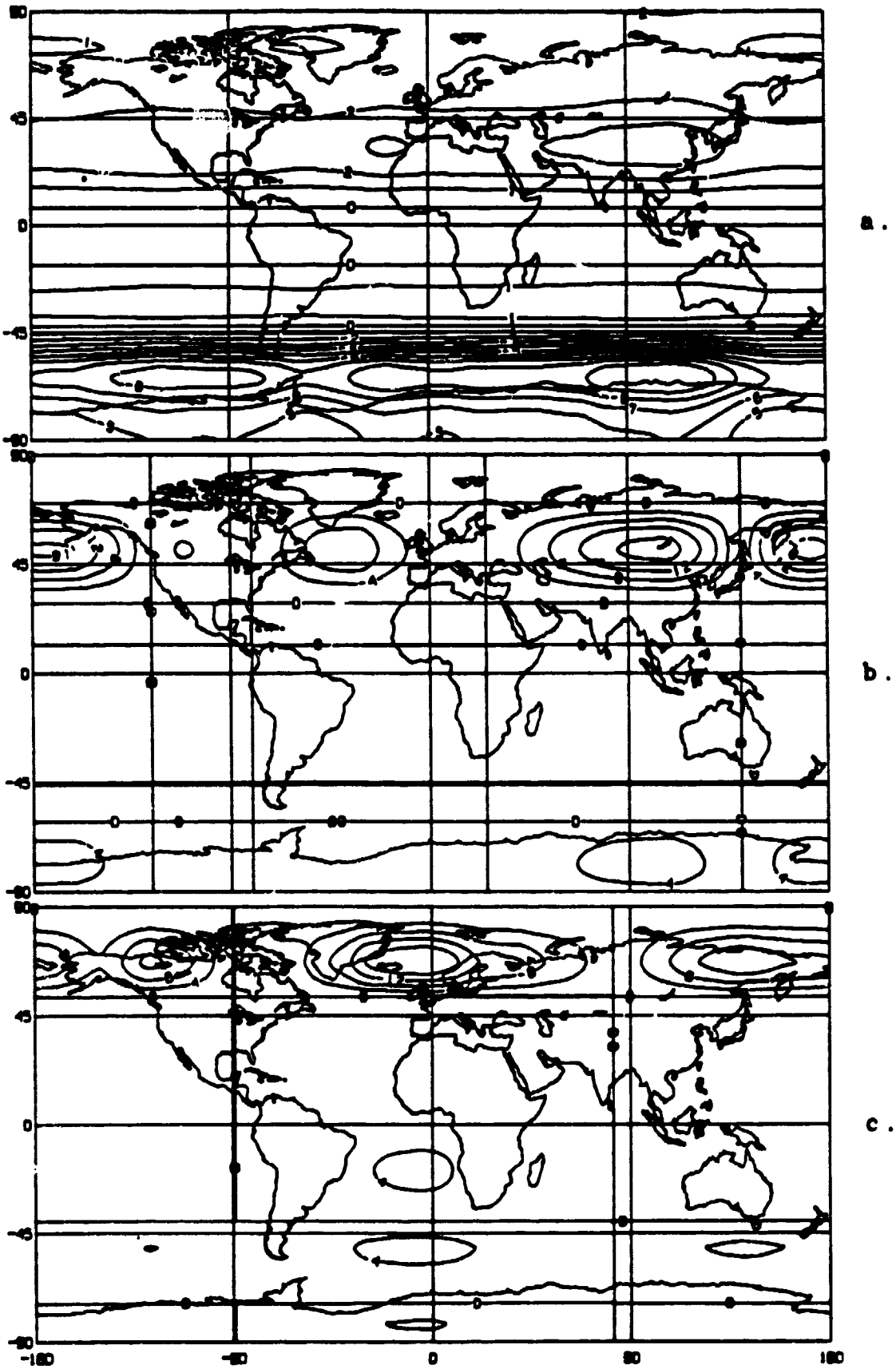

fig 7. First, second and third components (top to bottom) of observed climatological sea level pressure

first five eigenvalues, a, in units of m, K and mb, are listed in Table 1, together with the cumulative explained variances.

To assess the influence of the pole points, the eigenvectors were recalculated with the pole points excluded. Figure 8 shows the first component of SLP for the model (top) and observed (bottom) January climatologies, and Table 2 lists the eigenvalues and explained variances for the 22 x 36 data matrix.

Differences between the model and observed climatologies (fig. 1), which are most obvious in SLP and least apparent in G500, as well as similarities between corresponding maps, are exposed more clearly by the maps of the principal components (figs. 2-8). To interpret the latter quantitatively, it should be noted that the isopleths in figs. 2-8 are labelled with values that should be multiplied by 10^{-2}, and that the actual deviation from the mean is found by multiplying these values by the tabulated values of a. Thus, for example, the positive maximum of +13 over Siberia in fig. 6(a) represents a deviation from the global mean of 162.9 mb x 0.13, i.e. +21 mb.

From fig. 1 (c) it appears that the model simulates the 500 mb geopotential height field rather well. This impression is at least partly supported by figs. 2 and 3, where it can be seen that the mainly meridional structure of the first component of the model G500 field (which accounts for 97.4% of

Table 1. Eigenvalues (a_r), in m, K, and mb, of G500, T850, and SLP, respectively, and cumulative explained variances (E), in percent, for model (M) and observed (C) January climatologies. M x N = 24 x 36.

r	M		C	
	a_r	E (%)	a_r	E (%)
		G500		
1	9696.0	97.4	10738.7	96.9
2	1065.4	98.9	1113.7	99.0
3	489.7	99.4	579.0	99.5
4	305.5	99.5	479.3	99.8
5	230.7	99.6	296.2	99.9
		T850		
1	487.6	93.7	476.1	95.0
2	70.6	96.2	65.9	98.2
3	42.7	97.3	33.6	99.1
4	34.1	98.3	20.3	99.5
5	28.0	98.9	16.5	99.7
		SLP		
1	162.9	65.3	281.3	77.8
2	71.8	75.7	95.7	90.9
3	54.9	81.6	60.7	94.9
4	49.3	90.7	40.6	98.0
5	29.8	93.5	25.1	99.0

Table 2. Same as Table 1, but without pole points.

M x N = 22 x 36.

r	M		C	
	a_r	E	a_r	E
		G500		
1	8740.6	97.5	9427.0	96.8
2	996.9	99.0	1103.2	99.0
3	487.2	99.4	578.8	99.5
4	302.7	99.6	436.7	99.8
5	228.9	99.7	294.8	99.9
		T850		
1	449.5	94.0	417.8	95.2
2	66.3	96.6	61.4	98.3
3	42.5	97.6	33.6	99.2
4	34.1	98.4	20.3	99.5
5	28.0	99.0	16.1	99.7
		SLP		
1	157.5	64.3	266.8	70.4
2	68.1	75.5	95.3	88.3
3	53.8	83.4	60.7	94.0
4	49.1	90.8	40.6	97.6
5	28.8	93.7	25.0	98.9

its spatial variance) does indeed resemble that of the observed climatological G500 field (which explains 96.9% of its variance). The second component (contributing 1.5% and 2.1% to the explained variances of the model and observed climatologies, respectively) reflects a pattern of east coast troughs and west coast ridges in the Northern Hemisphere, which is found in both G500 fields. However, this component also reveals some significant differences between the model and observed patterns, with the longitudinal wave in the model-generated field located at a higher latitude than in the observed climatology, and with a weaker east coast trough over North America in the model. The third components of the two G500 fields also exhibit their principal differences over the western hemisphere.

The model simulation of T850 (fig. 1 (b)) is slightly less successful than that of G500. Again, the first components of both temperature fields (figs. 4(a) and 5 (a)) exhibit patterns that are mainly meridional and quite similar, although they explain a smaller part of the variance (93.7 and 95.0 percents) than at the 500 mb level. The second components of the 850 mb temperature fields, which reflect the distribution of cold continents and warm oceans in the Northern Hemisphere, are in fairly good agreement except for a northward displacement of the longitudinal wave in the model and the weaker influence of North America. However, the third components, which account for about 1% of the total variance, show striking

differences over Europe and Africa, as well as in the western hemisphere, apparently as a result of excessive heating in the model over the tropical and summer hemisphere continents (fig. 1 (b)).

From figs. 6 and 7, as well as fig. 1, it is obvious that the model's simulation of the sea-level pressure field is less than satisfactory. The first component of the model climatology (fig. 6 (a)) displays a cellular structure in the Southern Hemisphere, as well as the dominating continental highs of the Northern Hemisphere, whereas the first component of the observed SLP field (fig. 7 (a)) is mainly meridional. Of course, the first component explains only 65.3 and 77.8 percent, respectively, of the variance of the model and observed climatological fields. The second component adds another 10.4 and 13.1 percent, respectively, to the explained variances of the SLP fields, and, in the case of the observed climatology, it is this component that accounts for the continental highs and oceanic lows of the Northern Hemisphere. Both the second and third component of the model-generated SLP field are strongly influenced by the spurious Greenland high. However, the major discrepancy between the model and observed SLP fields is the predominantly cellular structure of the former as contrasted with the more meridional structure of the latter. This is shown not only by the striking differences between the two maps of first components, but also by the explained variances in Table 1 (and 2), where it can be seen that the first 5 components explain 99.0% of the

variance of the observed SLP field but only 93.5% of the "noisier" model-generated field.

The effects of the pole points on the EOF analysis may be seen by comparing fig. 8 (first components of model and observed SLP, with pole points excluded) with figs. 6 (a) and 7 (a) (first components of model and observed SLP, respectively, including pole points) and Table 2 with Table 1. From the figures, there appear to be no discernable differences between the principal components resulting from the pole points. A comparison of the corresponding maps (not shown) for G500 and T850, as well as those of the second and third components of all three variables, also reveals no significant effects of the pole points. On the other hand, the eigenvalues listed in Table 2 for the 22 x 36 matrices are consistently smaller than those in Table 1, indicating that exclusion of the pole points does alter slightly the magnitudes of the gradients of the component fields, although the patterns are hardly affected. Exclusion of the poles also reduces the percent of variance of the observed climatological SLP field explained by the first component from 77.8 to 70.4 percent, and raises the contribution of the second component (reflecting the continent vs. ocean pattern) from 13.1 to 17.8 percent. These relatively minor results of excluding (or retaining) the pole points do not, however, alter the basic results of the eigenvector analysis.

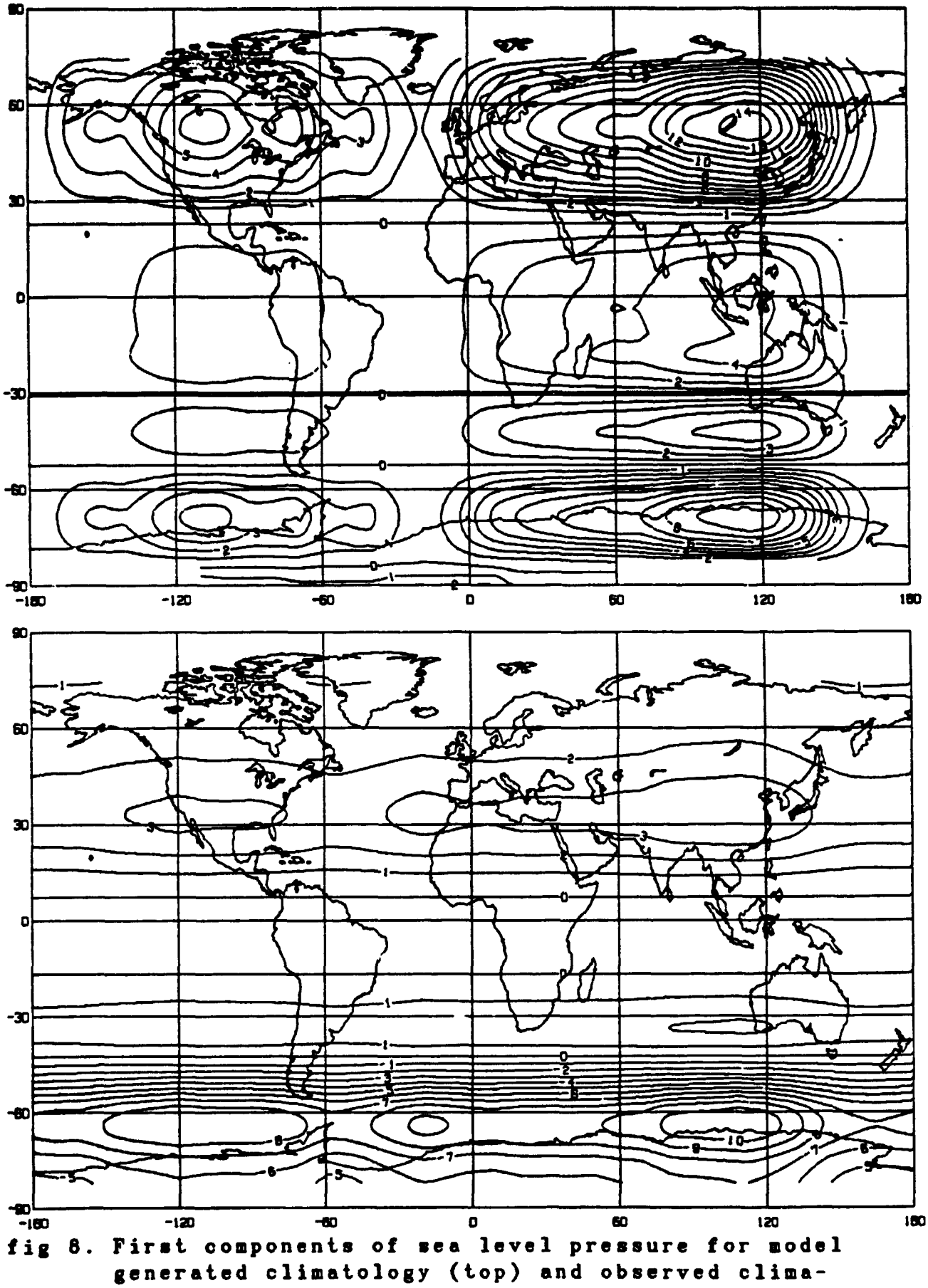

fig 8. First components of sea level pressure for model generated climatology (top) and observed climatology (bottom). Pole points are excluded making array 22x36.

Conclusions

Eigenvector analysis of observed and model-generated synoptic climatological fields is a useful diagnostic aid in the evaluation of global model simulations. The maps of any climatological variables are readily decomposed into a few orthogonal components which account for a large fraction of the total spatial variance, and the observed and model-generated patterns may then be compared in terms of these principal components of the original data matrices.

The GISS climate model simulation of a climatological 500 mb geopotential height field for January exhibits a smooth and predominantly meridional character, similar to that of the NCAR observed climatology, with 99.5% of the spatial variance explained by the first 3 or 4 components, and 97% of the variance accounted for by the mainly meridional first component in both cases. Differences are revealed by the second and higher components. Similar results are found for the 850 mb temperature field, but with a smaller fraction of the total spatial variance explained by the primarily meridional first component. The effects of the continents of the Northern Hemisphere on the longitudinal structure of these two fields, as reflected in the second components, are shifted poleward in the model simulation compared with the NCAR climatology, indicating a possible modeling problem.

The January climatological sea-level pressure field generated by the model exhibits more of a small scale,

cellular structure than that of the NCAR climatology. This is indicated not only by the first component, which is cellular for the model and meridional for the NCAR climatology, but also by the fact that the first 5 components explain 99.0% of the variance of the observed climatological field and only 93.5% of that of the model-generated map. The cellular structure of the model SLP field is possibly due to the short period (5 years) on which it is based, and will undoubtedly be smoothed by a longer run of the model. However, the weak meridional gradient in the Southern Hemisphere in the model simulation, which appears to be largely responsible for the non-meridional character of the first component, may indicate a more basic problem.

References

Hansen, J. and collaborators, 1980: An efficient three dimensional global model for climate studies. GISS, unpublished.

Jalickee, J.B. and W.J. Klepczynski, 1977: A method for compacting navigation tables. Navigation: Journal of the Institute of Navigation, 24, 125-131.

Kidson, J.W., 1975: Eigenvector analysis of monthly mean surface data. Mon. Wea. Rev., 103, 177-186.

Kutzbach, J. 1967: Empirical eigenvectors of sea level pressure, surface temperature and precipitation complexes over North America. J. Appl. Meteor., 6, 791-802.

Lorenz, E., 1956: Empirical orthogonal functions and statistical weather prediction. Sci. Rep. No. 1, Contract AF19 (604)-1566. Dept. of Meteorology, MIT, 49 pp.