# Toward a Dynamical Understanding of Planetary-Scale Flow Regimes

JOHN MARSHALL*

*Space and Atmospheric Physics Group, Department of Physics, Imperial College, London, United Kingdom*

FRANCO MOLTENI**

*European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom*

(Manuscript received 14 November 1991, in final form 30 September 1992)

ABSTRACT

A strategy for diagnosing and interpreting flow regimes that is firmly rooted in dynamical theory is presented and applied to the study of observed and modeled planetary-scale regimes of the wintertime circulation in the Northern Hemisphere. The method assumes a nonlinear dynamical model of the atmospheric motion, and determines a subspace of the phase space of the model in which multiple quasi-stationary solutions of the equations of motion are likely to be located. The axes that generate this subspace are the vectors that possess the smallest amplitude of the time derivative computed from a linearized version of the model, using the time-mean state of the system as a basic state. These vectors are called here "neutral vectors," and are shown to be eigenvectors of a self-adjoint operator derived from the linearized model.

As a prototype of a dynamical system with quadratic nonlinearity relevant to atmospheric dynamics, the three-variable convection model that generates the well-known Lorenz attractor is first investigated. It is shown that the presence of two unstable stationary solutions, which determine the shape of the attractor, generates a strong bimodality in the projection of the state vector of the system onto the most neutral vector, once a proper time filter is used on the data.

To apply this method to the analysis of atmospheric low-frequency variability, a three-level quasigeostrophic model in spherical geometry is adopted as the dynamical model. Neutral vectors are computed using the observed mean atmospheric state in winter as a basic state; alternative basic states, in which the eddies in the time-mean state are partially or fully removed, are also used in sensitivity experiments. The spatial patterns of the leading neutral vectors are relative insensitive to variations in some model parameters, but are strongly controlled by the form of the basic state; such dependence can be understood in terms of linear planetary-wave theory. The neutral vectors of the wintertime climatology are then used to analyse a 32-winter sample of observed atmospheric fields. It is found that the time series of the projection of these fields onto one particular neutral vector has a significantly bimodal probability density function, suggesting the existence of (at least) two separate flow regimes associated with anomalies of opposite sign. The two regimes are hemispheric in extent, and are close to some of the clusters found in previous studies that made use of empirical orthogonal functions.

Finally, it is shown that, if an appropriate forcing function is employed, the quasigeostrophic model is able to generate a very realistic climatology in a long nonlinear integration and, furthermore, two regimes similar to the observed ones. Again, these regimes can be identified by the presence of bimodality in the probability density function of the projections of model fields onto neutral vectors. Modeled and observed regimes have not only similar spatial patterns but also an almost identical distribution of the residence time.

## 1. Introduction

From a dynamical point of view, the atmosphere is a nonlinear system with an infinite number of degrees of freedom. Although hydrostatic and geostrophic balance places constraints on the possible modes of atmospheric variability, the number of states that the atmosphere can assume is infinitely large. Yet it is a matter of common synoptic experience that the atmosphere frequently adopts "preferred" flow patterns that are both persistent and recurrent. This has led to the notion of "weather regimes" or "flow regimes," and to the search of a rational basis to identify and study them.

Although efforts to objectively define regimes date back to the late 1940s–early 1950s (Baur 1947; Namias 1950; Rex 1950), substantial progress has only recently been achieved in this field. Flow regimes in the wintertime circulation over northern midlatitudes have been identified on the basis of multimodality in the one-dimensional or multidimensional probability density function of circulation indices (Sutera 1986;

* Current affiliation: Center for Meteorology and Physical Oceanography, Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts.

** Also affiliated with the Space and Atmospherics Group, Imperial College, London, United Kingdom.

*Corresponding author address:* Dr. Franco Molteni, European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berkshire, United Kingdom, RG2 9AX.

Hansen and Sutera 1987; Molteni et al. 1988, 1990; Kimoto 1989); cluster analysis (Mo and Ghil 1988; Legras et al. 1988); and minimization of an average time derivative (Vautard 1990). In addition, the search for quasi-stationary solutions has led to the definition of regimes not only in highly truncated models (such as those of Charney and Devore 1979; and Reinhold and Pierrehumbert 1982) but also in numerical models with a reasonable number of degrees of freedom (Legras and Ghil 1985; Mukougawa 1988; Vautard and Legras 1988; Vautard et al. 1988). Finally, flow regimes have been identified in general circulation models (GCMs) of the atmosphere (Hansen and Sutera 1990).

Despite such progress and a good agreement between diagnostic and modeling studies of localized regimes like blocking (see Vautard and Legras 1988), a fully objective definition and a satisfactory dynamical understanding have yet to be achieved in the case of planetary-scale regimes such as those investigated by Mo and Ghil (1988) and Molteni et al. (1990). Here difficulties arise because of the relatively large number of degrees of freedom needed to describe even the large-scale features of atmospheric variability. In observational studies, empirical orthogonal function (EOF) analysis is often used to reduce the dimensionality of phase space to a manageable size. Projections of atmospheric fields on the leading EOFs can subsequently be used in cluster analysis or probability density estimates in order to identify regimes. This approach has had considerable practical success. The leading EOFs explain a substantial proportion of the observed large-scale variability, and indeed resemble recurring atmospheric patterns. It is very difficult, however, to interpret EOFs dynamically, because they have no specific dynamical properties; consequently, there is no straightforward link between the statistical framework that we use to identify regimes and any dynamical theory that can explain them. Furthermore, the relatively simple numerical models in which quasi-stationary states and regimes have so far been identified dynamically cannot be easily used to interpret the observational results, because of the limited number of degrees of freedom represented. This prevents an unequivocal association between model regimes and observed planetary-scale regimes.

The purpose of this paper is to present a strategy for diagnosing and interpreting regimes in systems as multidimensional and complicated as the real atmosphere that is firmly rooted in dynamical theory. We show how to define a subspace of phase space with definite dynamical properties, in which quasi-stationary solutions are most likely to be located. The statistical methods that have been developed to analyze the projections on EOFs can be applied to the projections on this subspace, with the advantage of a more direct physical interpretation of any regimes identified.

The axes that generate this subspace will be referred to as *neutral vectors*. Given a plausible nonlinear dynamical model of the atmosphere, and a reference state that is likely to be an intermediate point between two quasi-stationary states, we linearize the model about the reference state and then define the neutral vectors as those (orthonormal) axes whose linearized time derivative possesses the smallest amplitude. (In stability theory, our neutral vectors correspond to the modes that have near-zero growth rate and phase speed.) The rationale behind our methodology is discussed in section 2, and its validity is demonstrated in the context of the well known Lorenz's three-variable convection model (Lorenz 1963) in section 3.

In section 4, we analyze the structure of neutral vectors computed with a three-level quasigeostrophic model, using basic states derived from the observed wintertime climatology. In section 5, we show that these neutral vectors do indeed describe important features of wintertime low-frequency variability in northern midlatitudes and that regimes can be identified by the presence of a highly significant bimodality in the projection of atmospheric fields onto one of them.

Section 6 demonstrates a further advantage of the neutral vector approach: if regimes are found in the projections of atmospheric fields onto neutral vectors computed by a linearized numerical model, the nonlinear version of that model can be used to test whether quasi-stationary solutions aligned along the *same* neutral vectors can explain the observed multimodality. A long nonlinear integration of our quasigeostrophic model produced a striking and very encouraging correspondence between regimes generated by a relatively simple dynamical model and regimes observed in the real atmosphere. Finally, our results are discussed in section 7, and possible extensions are suggested.

## 2. Neutral vectors and nonlinear stationary solutions

As a prototype of a nonlinear dynamical system relevant to atmospheric dynamics, let us adopt a quasigeostrophic (QG) model based on the time evolution of QG potential vorticity:

$$\frac{\partial q}{\partial t} = -J(\psi, q) - D(\psi) + S, \qquad (1)$$

where $q$ is potential vorticity (PV), $\psi$ streamfunction, $D(\psi)$ a linear operator that represents dissipative terms, $S$ a constant PV source, and $J$ the Jacobian of a two-dimensional field. Although Eq. (1) is the prognostic equation for the evolution of PV at a single level, our analysis also applies to a multilevel system of equations of the form of Eq. (1) (and with appropriate generalization to any forced dissipative system with quadratic nonlinearity). As is usual, we assume that the single-level or multilevel field of PV is a linear function of the (single-level or multilevel) streamfunction, which is invertible under appropriate boundary conditions.

Let us assume that Eq. (1) possesses two stationary solutions $(\psi_1, q_1)$ and $(\psi_2, q_2)$ that satisfy

$$-J(\psi_1, q_1) - D(\psi_1) + S = 0 \qquad (2a)$$

$$-J(\psi_2, q_2) - D(\psi_2) + S = 0. \quad (2b)$$

Defining the mean and the difference between the two steady states:

$$\bar{\psi} = \frac{\psi_1 + \psi_2}{2}, \quad \psi' = \frac{\psi_1 - \psi_2}{2}$$

$$\bar{q} = \frac{q_1 + q_2}{2}, \quad q' = \frac{q_1 - q_2}{2}$$

so that

$$\psi_1 = \bar{\psi} + \psi', \quad \psi_2 = \bar{\psi} - \psi'$$

$$q_1 = \bar{q} + q', \quad q_2 = \bar{q} - q'$$

and defining the operator:

$$L(\bar{\psi}, \psi') = -J(\bar{\psi}, q') - J(\psi', \bar{q}) - D(\psi'), \quad (3)$$

Eqs. (2a) and (2b) can be written, respectively:

$$-J(\bar{\psi}, \bar{q}) - D(\bar{\psi}) + S + L(\bar{\psi}, \psi') - J(\psi', q') = 0 \quad (4a)$$

$$-J(\bar{\psi}, \bar{q}) - D(\bar{\psi}) + S - L(\bar{\psi}, \psi') - J(\psi', q') = 0. \quad (4b)$$

Subtracting Eq. (4b) from Eq. (4a), one obtains

$$L(\bar{\psi}, \psi') = 0; \quad (5)$$

$L(\bar{\psi}, \psi')$ is simply the linearized time-derivative operator derived from Eq. (1) using the middle point between the two stationary solutions ($\bar{\psi}$) as a basic state, and applied to the vector $\psi'$ (the difference between the two steady solutions). A similar decomposition of the flow into $\bar{\psi}$ and $\psi'$ was proposed by Schneider (1988) to diagnose responses to variations in external forcing. Here, we use the technique to study stationary states determined by internal dynamics. Thus, if two stationary solutions exist satisfying Eq. (1) and the middle point between them is known, even if we are unable to solve the nonlinear Eq. (1), we can detect along which axis the two solutions are aligned by looking for a vector $\psi'$ that satisfies the linear Eq. (5). As mentioned in the Introduction, we shall call such a vector a "neutral" vector (this definition requires that such a vector not only has zero growth rate, but also has zero phase speed in the linear model). In practice, not even $\bar{\psi}$ is precisely known. In addition, $\psi_1$ and $\psi_2$ may be unstable stationary states. However, it has been shown by a number of investigations (e.g., Legras and Ghil 1985; Mo and Ghil 1987; Branstator and Opsteegh 1989) that, during numerical integrations of simple atmospheric models, the state vector of the system can spend a considerable amount of time in the neighborhood of stationary solutions even if they are unstable. In other words, some unstable stationary solutions can strongly influence the shape of the system attractor, so that flow regimes occur around them. This is not a recent idea; the famous Lorenz's strange attractor (Lorenz 1963) is composed of two spirals formed by trajectories that make loops around two slightly unstable stationary solutions. In Lorenz's system, the two solutions have symmetric properties and the mean point between the two steady states is relatively close to the time-mean state obtained by a long integration of the system, as we shall demonstrate in the next section.

We now hypothesize that, if the climate of a dynamical system can be seen as the alternation of two nearly equally populated regimes, which occur in the neighborhood of two unstable stationary solutions, then the linear operator $L$ computed using the time-mean state of the system as a basic state will still satisfy Eq. (5) to a useful degree of approximation. To the extent that this is true, Eq. (5) can be used to find the axis along which the two stationary states are aligned, given the time-mean state.

It is interesting to note that Eq. (5) implies that

$$J(\bar{\psi}, q') + J(\psi', \bar{q}) = -D(\psi'), \quad (6)$$

which says that the difference in the linearized PV advection in the two steady states is compensated by the difference in the dissipation term. If the last two terms on the rhs of Eq. (1) are combined to define a total PV forcing as

$$F(\psi) = -D(\psi - \psi^*), \quad (7)$$

where $D(\psi^*) = S$, Eq. (6) can be reexpressed as

$$J(\bar{\psi}, q_1 - q_2) + J(\psi_1 - \psi_2, \bar{q}) = F(\psi_1) - F(\psi_2). \quad (8)$$

Therefore, $\psi_1 - \psi_2$ can be formally seen as a linear response to the variation in PV forcing between $\psi_1$ and $\psi_2$ themselves. This suggest that, as far as the diabatic forcing is a function of the flow pattern, neutral vectors are likely to be excited in linear numerical models when the steady response to observed anomalies in diabatic forcing is computed. One should not conclude, however, that nonlinear dynamics is unimportant in the maintenance of multiple stationary solutions. The term $J(\psi', q')$ in Eqs. (4a) and (4b), which gives the self-interaction of the neutral vector, may be large, but is canceled out when the *difference* between the dynamical balance of $\psi_1$ and $\psi_2$ is considered, as a consequence of the quadratic nature of the Jacobian.

We now consider how to proceed, in practice, to look for stationary solutions of the type found in the Lorenz model in systems with a large number of degrees of freedom. Our discussion suggests that we should linearize the equations around the time-mean state ($\hat{\psi}$) of the system and look for axes $\psi'$ such that, given a suitable norm, $|L(\hat{\psi}, \psi')|$ is small. If $H$ is the linear operator that inverts PV into streamfunctions, we can obtain the time derivative of the streamfunction by applying $H$ to both sides of Eq. (1), and define for simplicity:

$$\hat{L}\psi' = HL(\hat{\psi}, \psi'). \quad (9)$$

We therefore look for axes that minimize the ratio:

$$\lambda^2 = \frac{|\hat{L}\psi'|^2}{|\psi'|^2} = \frac{\langle \hat{L}\psi', \hat{L}\psi' \rangle}{\langle \psi', \psi' \rangle}, \quad (10)$$

where $\langle , \rangle$ is an appropriate inner product. If $\hat{L}^*$ is the adjoint of $\hat{L}$ with respect to this inner product, we have

$$\lambda^2 = \frac{\langle \hat{L}^* \hat{L}\psi', \psi' \rangle}{\langle \psi', \psi' \rangle}. \quad (11)$$

Since $\hat{L}^*\hat{L}$ is a positive-definite, self-adjoint operator, it is straightforward to show that the axis for which $\lambda^2$ is minimized is the eigenvector of $\hat{L}^*\hat{L}$ with the smallest eigenvalue. The eigenvectors of $\hat{L}^*\hat{L}$ form an orthogonal basis, and for each eigenvector $\lambda^2$ is equal to the corresponding eigenvalue. In practice, when dealing with atmospheric models and data, we can consider as neutral vectors the eigenvectors for which $1/\lambda$ is greater than the typical decorrelation time for atmospheric fields. Such vectors, which have a sufficiently small growth rate and phase speed, define a subspace in which one may hope to find a statistical proof of the existence of stationary states and regimes.

It is important to stress that neutrality is a necessary but not sufficient property of the vector that joins two stationary states. Physically neutral vectors can be thought of as linearly resonant patterns that can be excited by weak forcing anomalies, and correspond to the singular vectors of $L$ studied by Navarra (1993). Whether a particular neutral vector does link two steady states depends on the full nonlinear balance of terms in the PV equation. In practice, these special vectors can be found by searching for bimodality in the probability density function of the projections of individual fields onto the neutral vectors. To illustrate our ideas in a familiar and well-defined context, we shall show how this procedure works in the case of Lorenz's attractor.

## 3. Analysis of Lorenz's dynamical system

As discussed in Lorenz's (1963) paper, the well-known Lorenz attractor is generated by a three-variable dynamical system, which is a highly truncated spectral model of Rayleigh–Benard convection. This simple, nonlinear dynamical system has long been considered as a prototype in which the problem of atmospheric predictability can be studied (see, for example, Moritz and Sutera 1981). Here, we are interested in the basic form of the Lorenz attractor; in particular, in the fact that its structure is determined by the presence of two unstable stationary solutions, so that the motion along the attractor is a combination of oscillations around either of the steady states and transitions from the neighborhood of one state to the other. The two spirals formed by the trajectories around the two steady states can be seen as the regimes of the Lorenz model, and the oscillations within them as the high-frequency transients that eventually cause the transitions from one regime to the other. In this section, we want to show how these regimes can be identified by a combination of the neutral vector analysis outlined in section 2 and appropriate statistical techniques such as time filtering and probability density estimation.

To do so, it is useful to start from a concise description of Lorenz's convection model. A layer of fluid of depth $H$ is considered, where a temperature difference between the lower and the upper surface is kept at a constant value. Assuming that no motion and no variation in the variables occur in the third dimension, the equations of motion can be expressed in terms of a two-dimensional streamfunction field $\psi$, and of the deviation $\theta$ of the temperature from a linear vertical profile. With appropriate boundary conditions, and letting $\psi_0$ and $\theta_0$ be appropriate normalization constants, Lorenz looks for solutions of the form:

$$\frac{\psi}{\psi_0} = X\sqrt{2}\sin\left(\pi a \frac{x}{H}\right)\sin\left(\pi \frac{z}{H}\right)$$

$$\frac{\theta}{\theta_0} = Y\sqrt{2}\cos\left(\pi a \frac{x}{H}\right)\sin\left(\pi \frac{z}{H}\right) - Z\sin\left(2\pi \frac{z}{H}\right). \quad (12)$$

Here, $X$ is an index of the direction and intensity of the convective motion, $Y$ of the horizontal temperature difference between the ascending and descending currents, and $Z$ of the deviation of the horizontally averaged temperature from linearity. The geometrical factor $a$ sets the horizontal wavelength of the streamfunction field.

From the original partial differential equations, the following equation can be derived for the derivatives of $X$, $Y$, $Z$ with respect to (dimensionless) time:

$$\dot{X} = -\sigma X + \sigma Y$$

$$\dot{Y} = -XZ + rX - Y$$

$$\dot{Z} = XY - bZ, \quad (13)$$

where $\sigma$ is the Prandtl number (kinematic viscosity/thermal conductivity), $r$ the ratio between the Rayleigh number and its critical value for the onset of convection, and $b$ a second geometrical factor equal to $4/(1 + a^2)$.

In addition to the trivial steady solution (0, 0, 0) for $r > 1$, Eq. (13) possesses two other stationary solutions:

$$X = Y = \pm[b(r-1)]^{1/2}, \quad Z = r - 1. \quad (14)$$

Let $C_1$ be the solution with positive $X$ and $Y$, $C_2$ the other one. These two solutions are stable if $r$ is lower than

$$r_0 = \sigma \frac{\sigma + b + 3}{\sigma - b - 1}. \quad (15)$$

Above this threshold for $r$, the two solutions become unstable, and the well-known strange attractor is generated by Eq. (13). The trajectory in phase space spirals alternately around $C_1$ and $C_2$, moving from the spiral around $C_1$ to the spiral around $C_2$ at irregular intervals. Physically, the switch between the neighborhood of $C_1$ and $C_2$ corresponds to a reversal in the sense of rotation of the convective cells, and in the phase of the horizontal temperature variations. Apart from this phase difference, the two steady states are physically equivalent, since the intensity of the convective motion and of the temperature gradient is the same. However, variations in the intensity of convection and in the temperature pattern occur within each oscillation around either $C_1$ or $C_2$.

The typical time of oscillation around $C_1$ and $C_2$ can be deduced by linearizing Eq. (13) around these two points. If, in general, $(X_0, Y_0, Z_0)$ is chosen as a basic state, the linearized dynamical system is given, in matrix notation, by

$$\frac{d}{dt}\begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix} = \begin{bmatrix} -\sigma & \sigma & 0 \\ (r - Z_0) & -1 & -X_0 \\ Y_0 & X_0 & -b \end{bmatrix}\begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix}. \quad (16)$$

The $3 \times 3$ matrix on the rhs of Eq. (16) is the analog of the operator $L$ defined in the previous section. Following Lorenz, we set $\sigma = 10$, $r = 28$, $b = 8/3$. If we then insert in Eq. (16) the values of $X_0$, $Y_0$, $Z_0$ corresponding to either $C_1 = (6\sqrt{2}, 6\sqrt{2}, 27)$ or $C_2 = (-6\sqrt{2}, -6\sqrt{2}, 27)$, we find one real, negative eigenvalue and two complex-conjugate eigenvalues for $L$. The complex eigenvalues correspond to an unstable normal mode with period $T = 0.62$; this mode accounts for the oscillatory motion around $C_1$ and $C_2$.

If we now numerically integrate Eq. (13) for $t = 630$ dimensionless units (corresponding to about 1050 oscillations) and disregard the spinup period $t < 30$, we obtain long time series for $X$, $Y$ and $Z$. We want to analyze these data on the basis of the concepts discussed in the previous section, and demonstrate that, without the need to solve the nonlinear equations explicitly, we can detect the existence of the two spirals (the two regimes of the Lorenz system), and locate their centers $C_1$ and $C_2$ in phase space to a good approximation.

First of all, we note that $\bar{C} = (0, 0, r - 1)$ is the middle point between $C_1$ and $C_2$, and that $(1, 1, 0)$ is a vector proportional to $C_1 - C_2$. By inserting the coordinates of $\bar{C}$ in Eq. (16) one can verify that $(1, 1, 0)$ is a neutral vector for $L$. In order to have the projection on this axis as one of the coordinates, we perform an orthogonal rotation of $X$ and $Y$ by defining

$$X' = \sqrt{0.5}(X + Y)$$
$$Y' = \sqrt{0.5}(X - Y). \quad (17)$$

Using these two coordinates together with $Z$, and the values of the parameters indicated above, we have $C_1 = (12, 0, 27)$ and $C_2 = (-12, 0, 27)$, so that the two solutions are aligned along the $X'$ axis. Table 1 gives the means and standard deviations of $X$, $Y$, $X'$, $Y'$, and $Z$ derived from our numerical integration, when no filtering is applied to the data, and when each time series is filtered by a running mean over a period $\delta t = 0.6$ (almost identical to the oscillation period of the unstable normal mode). As discussed by Mo and Ghil (1987), $X'$ is the coordinate with the largest variability in both the unfiltered and the filtered data. The unfiltered values of $X'$ and $Z$ have a comparable standard deviation, however, whereas in the filtered data the standard deviation of $X'$ is almost three times larger than the standard deviation of $Z$, and more than seven times that of $Y'$. One can conclude that most of the low-frequency variance is explained by $X'$.

Figure 1a shows the probability density function (PDF) for the filtered and unfiltered values of $X'$. No evidence of the existence of the two spirals can be deduced from the PDF of unfiltered data: the function has only one maximum at zero. A very different picture is obtained, however, when the high-frequency oscillations are filtered out. The PDF for the filtered $X'$ shows an impressive bimodality, with maxima at $+10$ and $-10$ ($C_1$ and $C_2$ are located at $X' = \pm 12$).

This statistical analysis clearly shows the existence of low-frequency regimes in the Lorenz attractor; however, it has been performed knowing a priori the axis along which $C_1$ and $C_2$ are aligned, and the period of the unstable normal mode of the system linearized about these solutions. We shall now try to obtain the same result without any a priori information on $C_1$ and $C_2$ by projecting the time series of the original variables $X$, $Y$, $Z$ obtained from the numerical integration onto neutral vectors.

From Table 1, one can see that the mean values of $X$ and $Y$ are very close to zero, and the mean value of $Z$ differs from its value in $C_1$ and $C_2$ by less than half the standard deviation of the unfiltered data. Following the notation of the previous section, let $\hat{L}$ be the linear

TABLE 1. Mean and standard deviation of unfiltered and filtered values of the variables in Lorenz's model.

| | $X$ | $Y$ | $X'$ | $Y'$ | $Z$ |
|---|---|---|---|---|---|
| Mean | 0.01 | 0.01 | 0.01 | 0.00 | 23.52 |
| Standard deviation (unfiltered) | 7.92 | 9.02 | 11.61 | 3.06 | 8.66 |
| Standard deviation (filtered) | 6.29 | 6.53 | 8.98 | 1.24 | 3.32 |

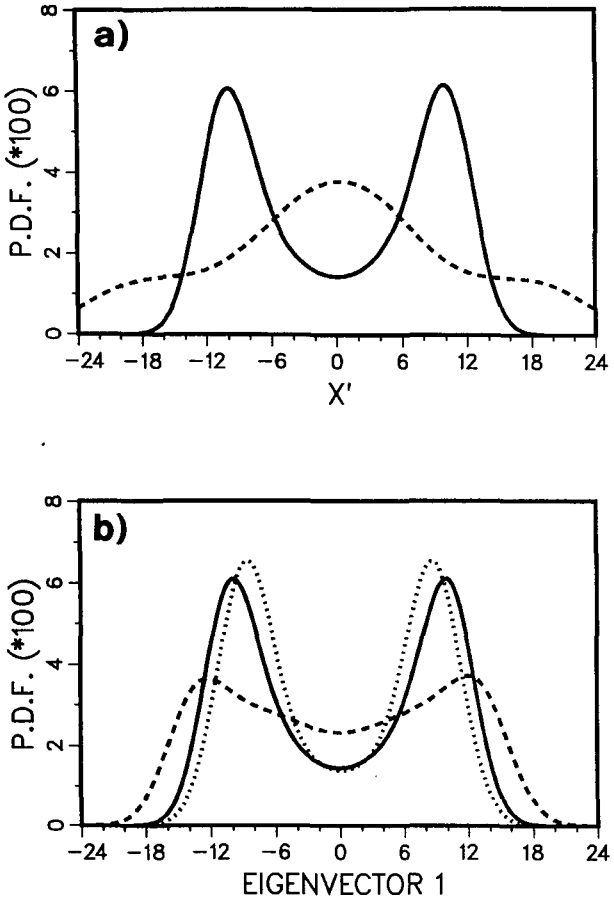


FIG. 1. (a) PDF of variable $X'$ in the Lorenz model; dashed line: unfiltered values; solid line: running means over $\delta t = 0.6$. (b) PDF of the projection on the first neutral vector of the Lorenz model, filtered by a running mean over $\delta t = 0.4$ (dashed line), $\delta t = 0.6$ (solid line), $\delta t = 0.8$ (dotted line).

operator obtained by using the time-mean state as the basic state in Eq. (16). The eigenvalues and (normalized) eigenvectors of $\hat{L}^*\hat{L}$ are listed in Table 2. The eigenvector with the smallest eigenvalue is very close to the $X'$ axis (the correlation is 0.998); the second eigenvector, with a slightly larger eigenvalue, is almost parallel to the $Z$ axis, whereas the third axis has a much larger eigenvalue.

The significance of the first eigenvector is confirmed by the PDF of the projection of $X$, $Y$, and $Z$ onto it. In the absence of indications from the normal-mode analysis, a simple visual inspection (or a spectral analysis) of the time series will reveal the approximate time scale of the high-frequency oscillation. We shall therefore try three filtering periods, namely, $\delta t = 0.4$, 0.6, and 0.8. The PDF of the projections onto the first eigenvector, filtered by a running mean over the above indicated periods, is shown in Fig. 1b. The bimodality along the first eigenvector is evident with all three filters; a filtering time longer than the optimal one reveals it more clearly than a shorter time, but the distance between the maxima is closer to the true distance between $C_1$ and $C_2$ when the shortest filtering period is chosen. When $\delta t = 0.6$ is used, the PDF of the projection onto the first eigenvector is almost indistinguishable from the PDF of the filtered $X'$ shown in Fig. 1a.

The neutral vector analysis has successfully identified the axis along which two stationary states are aligned in the Lorenz dynamical system, and the projection of time-filtered data onto the most neutral vector has clearly revealed the existence of regimes around these two states. It is important to note that we did not need to solve the full nonlinear equations to identify such axis. Because the climate of the Lorenz system is determined by the alternation of two equally populated regimes, the time-mean state provides a good approximation to the middle point between the stationary states in the definition of the linear operator $L$. We shall now apply our ideas to the real atmospheric circulation adopting a three-level quasigeostrophic model as our dynamical framework.

## 4. Neutral vectors of the wintertime climatology in quasigeostrophic dynamics

### *a. Formulation of the quasigeostrophic model*

The model used in this study for the search of neutral anomalies and quasi-stationary solutions in the atmospheric circulation is a spectral, three-level QG model, with global domain and pressure as the vertical coordinate. The basic equations from which it has been derived are briefly discussed in appendix A. The series of spherical harmonics used in the representation of horizontal fields has a triangular truncation at total wavenumber 21 (*T*21). The model integrates prognostic equations for QG PV at 200 hPa (level 1), 500 hPa (level 2), and 800 hPa (level 3) of the form:

$$\frac{\partial q_1}{\partial t} = -J(\psi_1, q_1) - D_1(\psi_1, \psi_2) + S_1 \qquad (18a)$$

$$\frac{\partial q_2}{\partial t} = -J(\psi_2, q_2) - D_2(\psi_1, \psi_2, \psi_3) + S_2 \qquad (18b)$$

$$\frac{\partial q_3}{\partial t} = -J(\psi_3, q_3) - D_3(\psi_2, \psi_3) + S_3, \qquad (18c)$$

TABLE 2. Eigenvalues ($\lambda^2$) and eigenvectors of $\hat{L}^*\hat{L}$ for the Lorenz model.

| $\lambda^2$ | $X$ | $Y$ | $Z$ |
|---|---|---|---|
| 5.62 | 0.673 | 0.738 | 0.048 |
| 7.11 | −0.032 | −0.036 | 0.999 |
| 215.46 | 0.739 | −0.674 | 0.000 |

where the index $i = 1, 2, 3$ refers to the pressure level. Here PV is defined as

$$q_1 = \nabla^2\psi_1 - R_1^{-2}(\psi_1 - \psi_2) + f \tag{19a}$$

$$q_2 = \nabla^2\psi_2 + R_1^{-2}(\psi_1 - \psi_2) - R_2^{-2}(\psi_2 - \psi_3) + f \tag{19b}$$

$$q_3 = \nabla^2\psi_3 + R_2^{-2}(\psi_2 - \psi_3) + f\left(1 + \frac{h}{H_0}\right), \tag{19c}$$

where $f = 2\Omega \sin\phi$, $R_1$ (=700 km), and $R_2$ (=450 km) are Rossby radii of deformation appropriate to the 200–500-hPa layer and the 500–800-hPa layer, respectively; $h$ is the (real) orographic height; and $H_0$ a scale height (set to 9 km).

In Eqs. (18a)–(18c), $D_1$, $D_2$, $D_3$ are linear operators representing the effects of Newtonian relaxation of temperature, linear drag on the 800-hPa wind (with drag coefficient depending on the nature of the underlying surface), and horizontal diffusion of vorticity and temperature. The exact form adopted for these operators is given in appendix A. Here, we note only that the temperature relaxation has a radiative time scale of 25 days; the linear drag damps the low-level wind on a spindown time scale of 3 days over the oceans, about 2 days over low-altitude land, and about 1.5 days over mountains above 2 km; a (strongly scale-selective) horizontal diffusion damps harmonics of total wavenumber 21 on a 2-day time scale.

The terms $S_1$, $S_2$, and $S_3$ are time independent but spatially varying sources of PV. Following Roads (1987), these can be determined by requiring that, when PV tendencies (i.e., time derivatives) are computed from Eq. (18) for a large number of observed atmospheric fields, the average values of these tendencies must be zero. In practice, the $S_i$ terms can be computed as the opposite of the average PV tendencies obtained by inserting observed streamfunction fields into a version of Eq. (18) in which these terms are omitted. This is equivalent to assuming that the sample of fields used in such a computation is representative of a statistically stable climatology.

In section 6, we shall discuss how to define $S_1$, $S_2$, and $S_3$ in a way compatible with the existence of two atmospheric flow regimes. The simple technique described above, however, was employed to obtain preliminary estimates of $S_i$, which were then used in nonlinear integrations aimed at determining the most appropriate values for the model parameters. (Note that the $S_i$ terms do not appear in the linearized version of the model used to compute neutral vectors.) Since we are interested in the wintertime variability in the Northern Hemisphere, the parameters were optimized for this region and period; the ensemble of observed fields used in the estimates of $S_i$ were ECMWF (European Centre for Medium-Range Weather Forecasts) analyses of streamfunction at the three pressure levels for each day in January and February 1984 to 1989.

A full description of the model climatology with the "two-regime" PV source will be given in section 6, where we shall discuss the results of a long nonlinear integration. It suffices to say here that, despite the simplicity of its assumed dynamics and physics, the model exhibits a very realistic mean state and a remarkably good simulation of wintertime extratropical variability; in particular, the use of a global domain allows a realistic description of the meridional structure of Northern Hemisphere planetary waves.

### *b. Linearization and definition of the basic state*

The QG model described in section 4a can be readily linearized in order to compute, for any given basic state, the operator $\hat{L}$ [defined by Eqs. (9) and (3)], which gives the time derivative of a streamfunction perturbation. Using the nonnull spherical harmonic coefficients as coordinates of the phase space of the model, the operator $\hat{L}$ can be written as a $1449 \times 1449$ matrix. From this matrix, the adjoint operator with respect to various inner products can also be computed. Here, an inner product has been chosen such that the squared norm of a streamfunction vector is proportional to its total kinetic energy; in other words, the norm gives the rms amplitude of the rotational wind.

As previously discussed, neutral vectors for the operator $\hat{L}$ will be identified as the eigenvectors of $\hat{L}^*\hat{L}$ with the smallest eigenvalues. Even though $\hat{L}^*\hat{L}$ is a symmetric, positive definite operator, this is a computationally demanding operation, given the size of the matrix. A posteriori checks on the eigenvectors have been performed, which have shown a high degree of numerical accuracy even in the computation of the eigenvector with the smallest eigenvalue.

In order to evaluate the sensitivity of the structure of the neutral vectors (and of the associated eigenvalues) to the basic state and to the assumed form of dissipative and forcing processes, neutral vectors have been computed for the following five combinations of model formulation and basic state.

1) The model has exactly the formulation given in section 4a and appendix A; the basic state is the average of the observed streamfunction in January and February 1984 to 1989, which we shall refer to as the observed climatology.

2) The model and the basic state are as in 1, but the time scale for temperature relaxation is reduced from 25 to 15 days.

3) The basic state is as in 1; the orography is neglected in the definition of PV at 800 hPa, and an average, spatially uniform drag coefficient for the 800-hPa wind is assumed, equal to $(2.5 \text{ days})^{-1}$); here, zonal asymmetries in the dynamics are only due to the basic state.

4) The model is as in 3; the basic state is derived from the observed climatology by progressively smoothing the eddy streamfunction field in the north-

ern subtropics, so that the stationary waves are retained in the northern extratropics, but the basic state is reduced to the zonal mean climatology in the tropics and in the Southern Hemisphere.

5) The model is as in 3, and the zonal mean climatology is assumed as a basic state.

The damping of the stationary waves in the basic state of case 4 was obtained by multiplying the eddy component of the streamfunction field by the latitude-dependent weight:

$$w(\phi) = 1 - (\cos\phi)^4 \quad \text{for} \quad \phi > 0$$
$$w(\phi) = 0 \quad \text{for} \quad \phi \leq 0.$$

It is interesting to consider the differences between the PV fields computed from the streamfunction of the climatological and the modified basic state. The 500-hPa PV of the two states is shown in Fig. 2. The removal of tropical eddies produces a uniform meridional gradient of PV in the northern subtropics, whereas regions of weak PV gradient exist at these latitudes in the climatological field, which can cause local reflection of Northern Hemisphere planetary waves (e.g., McIntyre 1982). We shall see that such changes in the tropical PV structure have a radical effect on the pattern of our neutral vectors.

### c. *Characteristic time and spatial pattern of the neutral vectors*

We shall now discuss the properties of the neutral vectors obtained with the combinations of basic state and model parameters listed in the previous subsection. If $\lambda^2$ is an eigenvalue of $\hat{L}^*\hat{L}$, the inverse of its square root $\lambda^{-1}$ gives the ratio between the norm of the eigenvector and the norm of its linear tendency. It can be interpreted as a characteristic time scale for the evolution of the eigenvector due to linear dynamics (irrespective of whether the perturbation is amplified, damped, or just advected). We shall use this parameter to quantify the neutrality of different eigenvectors.

The characteristic time $\lambda^{-1}$ is listed in Table 3 for eigenvectors 1–5, 10, and 20 computed from the model–basic state combinations 1–5. From the table,

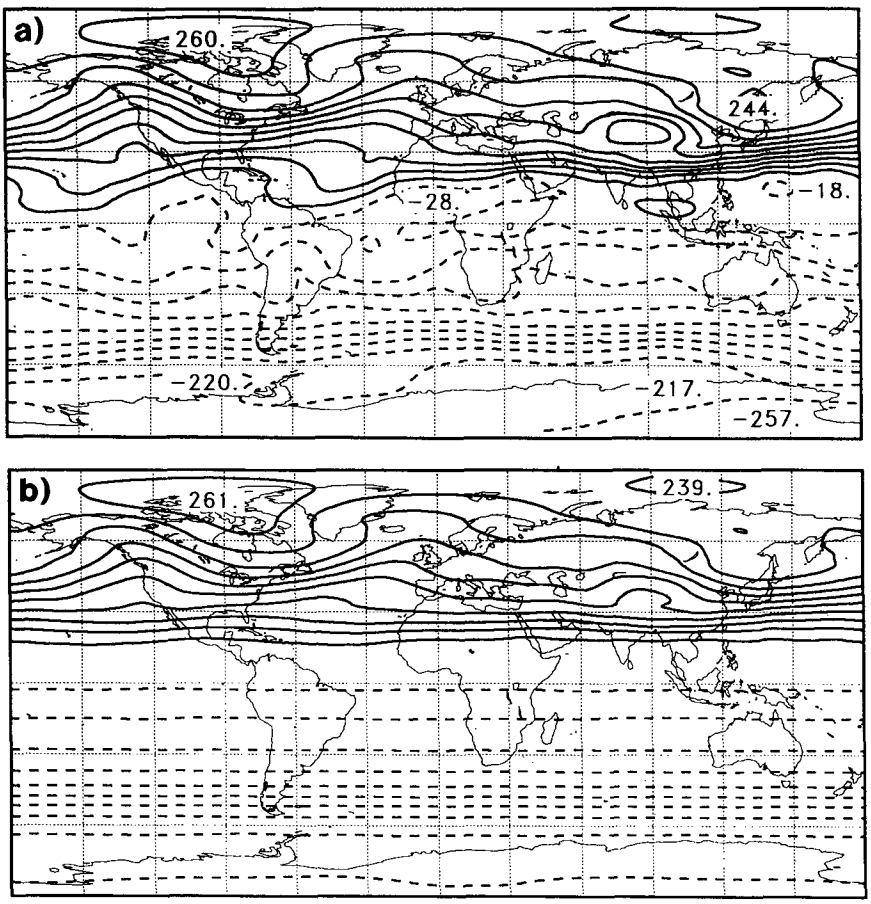


FIG. 2. Quasigeostrophic potential vorticity ($10^{-6}$ s$^{-1}$) at 500 hPa computed from (a) observed streamfunction climatology in January–February; (b) streamfunction climatology with eddies progressively damped in the tropics and the Southern Hemisphere.

TABLE 3. Characteristic times (days) of selected neutral vectors, for various basic states and model formulations.

| | Case 1 | Case 2 | Case 3 | Case 4 | Case 5 |
|---|---|---|---|---|---|
| Eigenvector 1 | 2398 | 917 | 1969 | 307 | 192 |
| Eigenvector 2 | 775 | 369 | 877 | 205 | 192 |
| Eigenvector 3 | 405 | 254 | 407 | 179 | 97 |
| Eigenvector 4 | 316 | 215 | 318 | 147 | 97 |
| Eigenvector 5 | 216 | 159 | 221 | 119 | 76 |
| Eigenvector 10 | 105 | 83 | 112 | 74 | 66 |
| Eigenvector 20 | 58 | 45 | 55 | 52 | 50 |

one can see that in each case neutral eigenvectors exist with $\lambda^{-1}$ at least one order of magnitude larger than the typical decorrelation time of large-scale fields (about 10 days). This is particularly true when the full climatology is used as a basic state: the damping of tropical stationary eddies (in case 4) increases considerably the magnitude of the linear tendency for the most neutral eigenvectors. The eigenvectors obtained from the zonal-mean basic state (case 5) have even larger tendencies. We therefore observe that the characteristic time of the leading neutral vectors increases with increasing amplitude and extent of the stationary eddies.

The removal of orography and of land–sea variation in the drag coefficient (case 3) has a marginal effect on the eigenvalues. The increase in the intensity of temperature relaxation (case 2) significantly affects the magnitude of the tendency of the two most neutral eigenvectors, but the change is about 25% or less from the fifth eigenvector onwards, and is definitely smaller than the change due to variations in the basic state. We want to point out that, as far as the existence of multiple quasi-stationary states is concerned, it matters little whether a particular neutral vector has a characteristic time of 200 or 150 days, as long as this time is much longer than the typical lifetime of large-scale regimes (about 10 to 20 days).

Let us now examine the spatial structure of the neutral vectors. The first (about) 10 eigenvectors obtained in cases 1, 2, and 3 are very similar to each other. The eigenvectors pertaining to the zonal mean flow have regular structures that bear little resemblance to observed patterns of low-frequency variability. Therefore, in the following we shall concentrate our attention of eigenvectors from cases 1 and 4, keeping in mind that their difference is due almost entirely to the change in the basic state.

The three-dimensional structure of the most neutral vector $\mathbf{E}_1$ in the real climatology analysis (case 1) is shown in Fig. 3; the fifth eigenvector from case 1, $\mathbf{E}_5$, which will prove to be a very important one in the subsequent analyses, is shown in Fig. 4; the first eigenvector $\mathbf{E}'_1$ of case 4 is shown in Fig. 5. In these figures, the vectors are normalized so that their rms wind amplitude is 5 m $s^{-1}$, which is comparable to the rms amplitude of the projections of actual fields on these vectors.

It is evident that $\mathbf{E}'_1$ has a more localized structure, both in longitude and in latitude, than either $\mathbf{E}_1$ or $\mathbf{E}_5$. The eigenvector $\mathbf{E}'_1$ has no significant features in the Southern Hemisphere, apart from a wavenumber 2 pattern over Antarctica that is practically identical to the most neutral vector of the zonal mean basic state. In $\mathbf{E}_1$ and $\mathbf{E}_5$, tropical features are particularly evident at 200 hPa, with maximum amplitude in the eastern Pacific. This is consistent with the existence of upper-level westerlies and strong meridional PV gradient in this area, allowing midlatitude stationary waves to extend meridionally into the tropics (see Webster and Holton 1982).

In the northern midlatitudes, all three vectors show an equivalent barotropic structure. The eddy component of $\mathbf{E}_1$ and $\mathbf{E}_5$ is hemispheric in extent, whereas in $\mathbf{E}'_1$ the eddies are mostly confined to the Pacific–North American sector. These eddy patterns can be better seen in Figs. 6a, 6b, and 6c, respectively, which show the eddy fields of 500-hPa height computed from the corresponding streamfunction fields by solving the linear balance equation. In both $\mathbf{E}_1$ and $\mathbf{E}_5$, a wave pattern covering the Pacific and the American continent appears to be reflected in the tropical Atlantic, and generate another wave pattern over the central Atlantic, Europe, and northern Asia. It is likely that the nonuniform structure of the subtropical PV gradient in the climatological basic state is responsible for such a complex wave pattern. In the basic state of case 4, the subtropical PV gradient is a very weak function of longitude (due to our smoothing of the tropical stationary eddies), and the tropical belt tends to absorb midlatitude planetary waves. This wave absorption can be noticed in $\mathbf{E}'_1$ as well as in a number of other neutral vectors of case 4.

We can ask ourselves whether the neutral vectors resemble observed patterns of wintertime low-frequency variability found by statistical techniques. Figures 6d and 6e show the fourth and the sixth empirical orthogonal function (EOF) of the 500-hPa geopotential eddies computed by Molteni et al. (1988) from a 32-winter sample of 5-day mean fields. These two EOFs have a clear resemblance to $\mathbf{E}_1$ and $\mathbf{E}_5$, respectively. On the other hand, localized variability patterns not dissimilar from $\mathbf{E}'_1$ have been found by Horel (1981), Barnston and Livezey (1987), and others by means of rotated principal component analysis.

Finally, one can note that eigenvector $\mathbf{E}_5$ has a significant projection on the well-known Pacific–North American (PNA) teleconnection pattern defined by Wallace and Gutzler (1981). A composite PNA anomaly, computed by Molteni et al. (1990) from the above-mentioned 32-winter sample, is shown here in Fig. 6f. Although this map has (by definition) most of its variance in the PNA sector, like $\mathbf{E}'_1$, a secondary wave pattern also exists over the Atlantic Ocean and

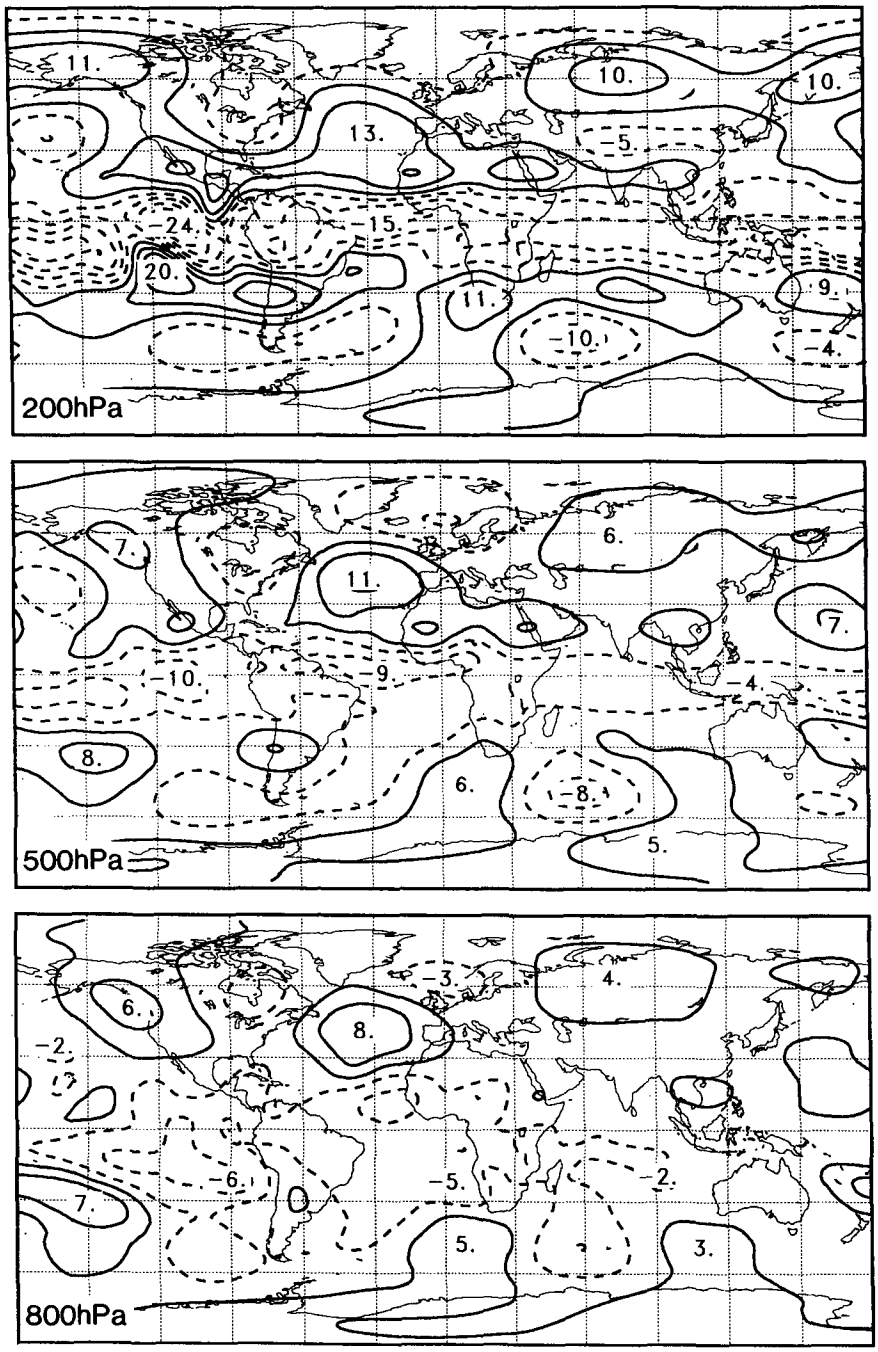


FIG. 3. Streamfunction ($10^6$ m$^2$ s$^{-1}$) at 200 hPa (top), 500 hPa (center), and 800 hPa (bottom) of the neutral vector $\mathbf{E}_1$, with (rms wind) norm of 5 m s$^{-1}$.

Eurasia, which is in phase with the Atlantic–Eurasian part of $\mathbf{E}_5$. Our analysis suggests that the tropical PV structure over the Central American and Atlantic sector may play an important role in determining whether a PNA-like anomaly remains localized or becomes embedded in hemispheric-scale regimes like those found by Molteni et al. (1990).

## 5. Observed low-frequency variability along neutral vectors

In order to assess the importance of the neutral vectors discussed in section 4 for atmospheric low-frequency variability, statistics of the projections of observed atmospheric anomalies onto them will be ex-

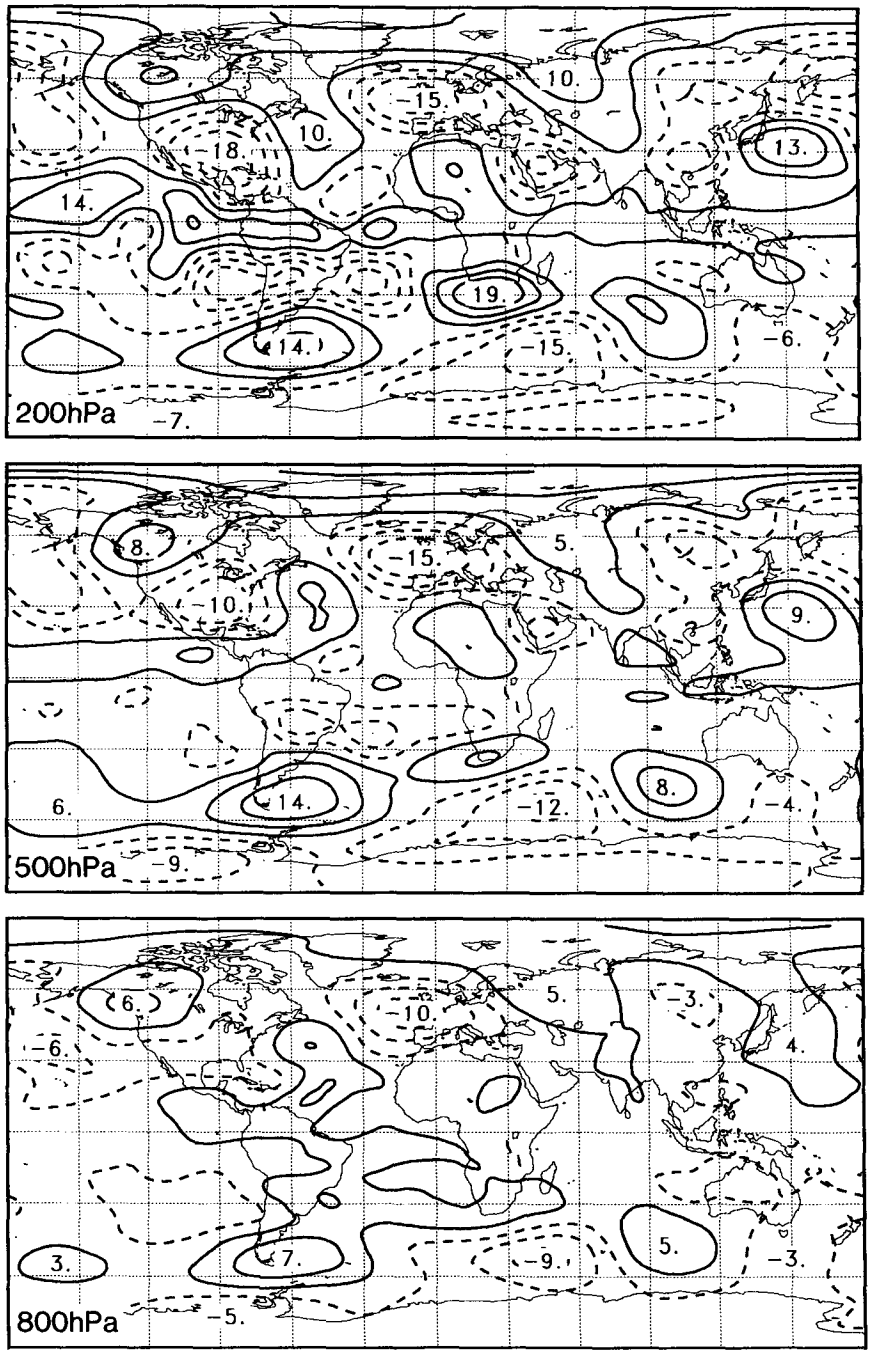


FIG. 4. As in Fig. 3 but for the neutral vector $\mathbf{E}_5$.

amined here. In particular, we want to investigate whether these projections can reveal the signature of multiple regimes on a hemispheric scale. Given the dynamical properties of neutral vectors discussed in section 2, bimodality along neutral vectors would support the hypothesis that different regimes correspond to multiple quasi-stationary solutions.

Although planetary-scale regimes may involve variations in the zonal mean flow, it has been shown that their most evident signature is multimodality in the PDF of indices of amplitude and phase of large-scale eddies (see Sutera 1986; Hansen and Sutera 1987; Molteni et al. 1988, 1990). Therefore, as in Molteni et al. (1988, 1990), we shall use a dataset composed of observed 5-day mean eddy fields of 500-hPa height for the winter season (December to February), from winter 1952/53 to 1983/84. This 32-winter sample is derived from analyses of the U.S. National Meteoro-

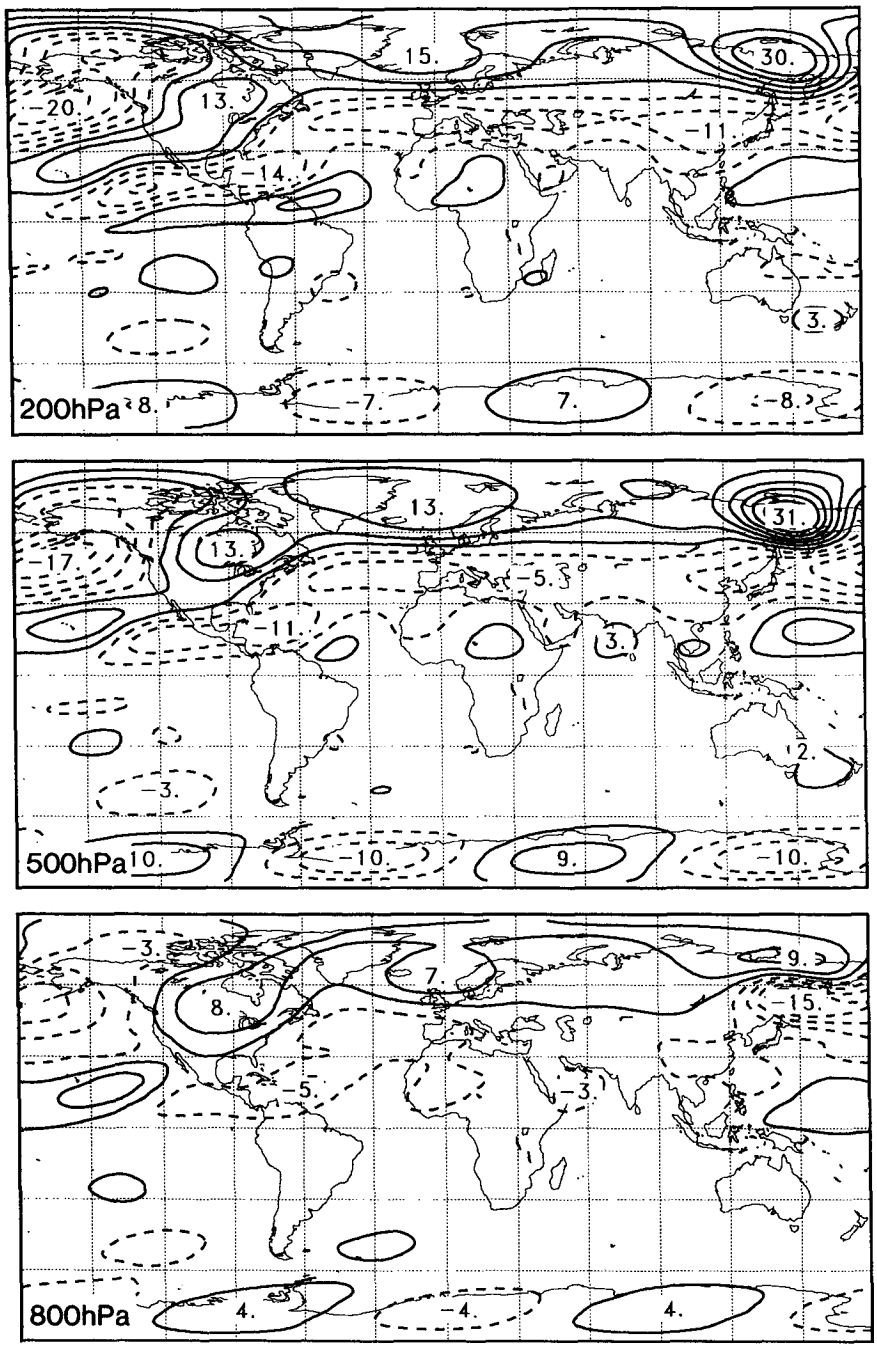


FIG. 5. As in Fig. 3 but for the neutral vector $E'_1$.

logical Center and of ECMWF, and covers the Northern Hemisphere north of 20°N (see Molteni et al. 1988 for more details). Here, we shall consider the eddy anomalies obtained by subtracting a (running) monthly climatology from the original data. In addition, the 5-day mean removes most of the signal from high-frequency baroclinic eddies, so that these fields are representative of intraseasonal and interannual low-frequency variations in planetary waves.

As partially anticipated in the previous section, the 500-hPa streamfunction of the ten most neutral eigenvectors of cases 1 and 4 has been converted to geopotential height by solving a linear balance equation. The zonal means have then been subtracted, and the eddy height fields have been renormalized so that their rms amplitude from 20° to 90°N is equal to one. The 576 observed eddy anomalies have been projected onto each of the eddy neutral vectors, and PDF for such

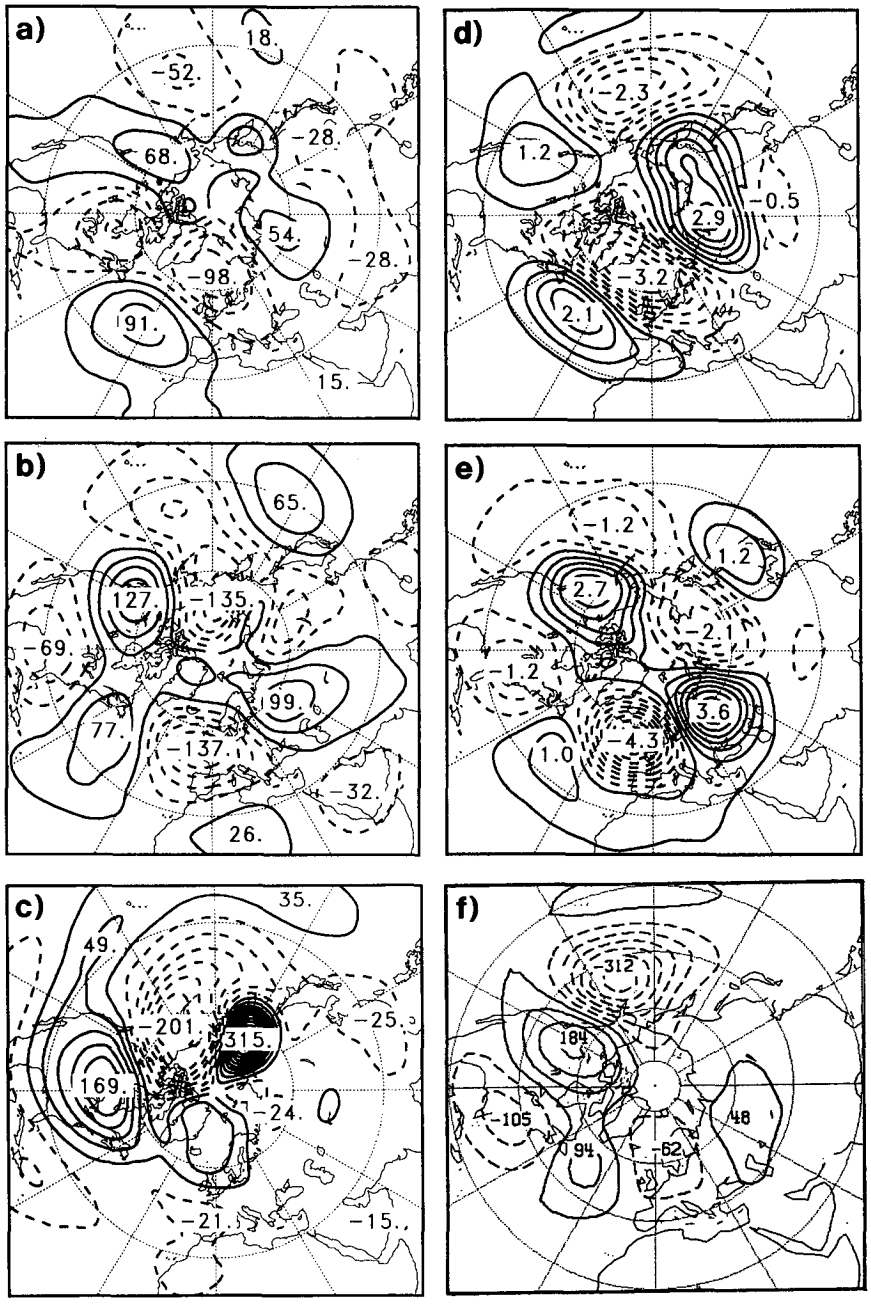


FIG. 6. (a) Eddy geopotential height at 500 hPa of the neutral vector $\mathbf{E}_1$; (b) as in (a) but for $\mathbf{E}_5$; (c) as in (a) but for $\mathbf{E}'_1$; (d) EOF 4 of the geopotential eddies at 500 hPa, from Molteni et al. (1988); (e) as in (d) but for EOF 6; and (f) composite PNA pattern, computed from 5-day mean fields by Molteni et al. (1990).

projections have been computed using a Gaussian kernel estimator (see appendix B). In the estimator, the smoothing parameter has been set to such a value that, given the size of our sample, any bimodality in a PDF is statistically significant at the 95% confidence level.

The PDFs for the projection onto eigenvectors $\mathbf{E}_1$, $\mathbf{E}_5$, and $\mathbf{E}'_1$ (that is, on the renormalized patterns shown in Figs. 6a, 6b, and 6c) are shown in Figs. 7a, 7b, and 7c. The standard deviations of the three projections are 24.6, 22.5, and 24.1 m, respectively, which are comparable with the standard deviations of some of the leading principal components (i.e., projections on EOFs) of the eddy fields (see Molteni et al. 1988). The first neutral vector $\mathbf{E}_1$ of the real climatology is indeed the one that possesses the largest standard deviation among the 20 neutral vectors considered; however, its

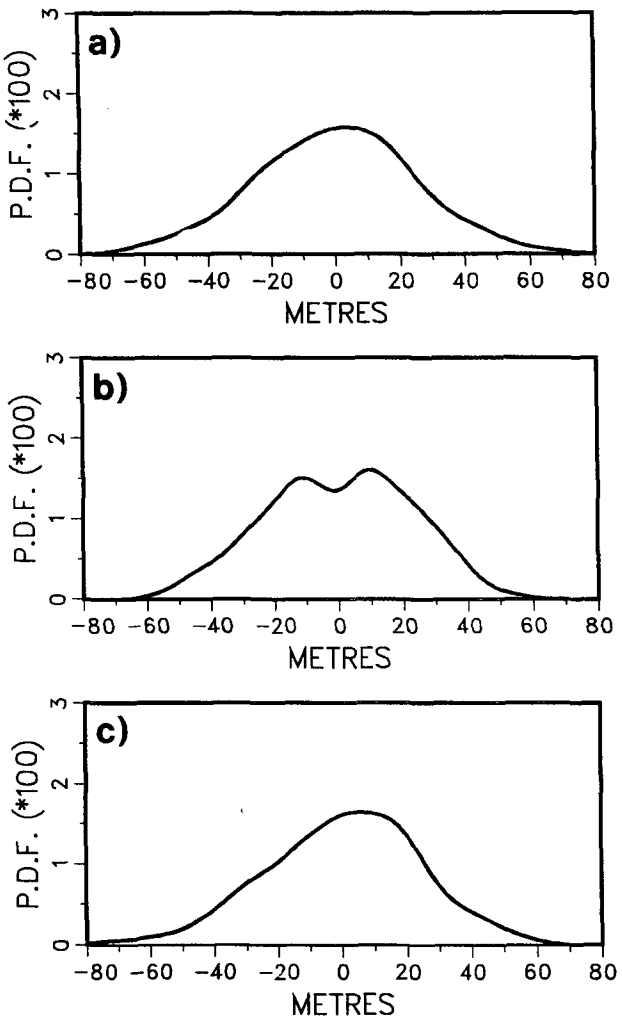


FIG. 7. PDF of the projections of observed 5-day mean eddy fields of 500-hPa height (between 20° and 90°N) onto neutral vectors (a) $\mathbf{E}_1$, (b) $\mathbf{E}_5$, and (c) $\mathbf{E}'_1$.

PDF is unimodal, as well as the PDF of the projection on $\mathbf{E}'_1$. On the other hand, the projection on $\mathbf{E}_5$ shows a highly significant bimodality, suggesting the existence of two regimes aligned along this axis.

Among those considered, $\mathbf{E}_5$ is the only neutral vector along which bimodality has been found. As mentioned previously, neutrality is a necessary but not sufficient property of the vector linking stationary states with opposite anomalies. So, the existence of many neutral vectors with unimodal PDF is of no importance, provided at least one vector shows a significant bimodality. The fact that no regimes can be identified using the neutral vectors of case 4 (computed from an artificially smoothed basic state) is probably not surprising. As far as the neutral vectors of the real climatology are concerned, one might still look for multiple regimes in linear spaces generated by more than one neutral vector (we shall use this method to investigate model regimes in the following section). Although such multimodality would equally be physically significant, we prefer at this stage to consider only projections on individual axes selected a priori, in order to maximize the statistical significance of the results.

It is interesting to note that, on the hemispheric domain, $\mathbf{E}_5$ is almost orthogonal to the wintertime stationary wave pattern (the spatial correlation between the two eddy fields is 0.08). Therefore, bimodality along $\mathbf{E}_5$ means bimodality in the phase, rather than in the total amplitude, of planetary waves, in partial analogy with the bimodality in the Lorenz model considered in section 3. Although bimodality along $\mathbf{E}_5$ cannot explain the bimodality in planetary wave amplitude discussed by Sutera (1986) and Hansen and Sutera (1987), it reveals a phenomenon of comparable statistical significance and physical relevance, and can indeed explain (as it will be shown in section 6) two important atmospheric regimes found by Mo and Ghil (1988) and Molteni et al. (1990).

In order to support our interpretation of the two probability maxima (the two "modes") in the projection on $\mathbf{E}_5$ as evidence of the existence of two stationary solutions, it must be demonstrated that bimodality does not arise from a periodic process. This can be done by studying the distribution of residence time, which is the time interval in which the projection on $\mathbf{E}_5$ remains in one mode (either greater or lower than the value corresponding to the relative minimum in the PDF). A distribution in which the most frequent value is close to the mean value would be the signature of a periodic process; conversely, an exponential distribution would imply that the probability of transition from one regime to the other is independent of the time the system has already spent in a given regime, and would suggest a stochastic nature for the transition process. Dole and Gordon (1983) found exponential distributions for the durations of persistent anomalies, and Mukougawa (1988) found similar results for the probability of persistence of quasi-stationary states in his 28-variable baroclinic model.

The histogram of residence times for $\mathbf{E}_5$ (in pentads, i.e., 5-day periods) is shown in Fig. 8. The black bars show the actual frequencies, while the white bars behind show the best fit to those frequencies by an exponential distribution. The agreement between the two is very good, therefore excluding any periodicity in the transitions. The average residence time deduced from the histogram is 3.2 pentads, that is, 16 days.

It is interesting to consider an example of a transition between two real atmospheric states with large and opposite projections onto $\mathbf{E}_5$. Figure 9a shows the mean 500-hPa height over the Northern Hemisphere in the pentad 5–9 January 1985; Fig. 9b shows the corresponding map for 2–6 February of the same year, that is, 27 days later. The first map has strong ridges over the west coast of North America, the Atlantic, and west Asia at 60°E, where $\mathbf{E}_5$ has positive maxima; the second

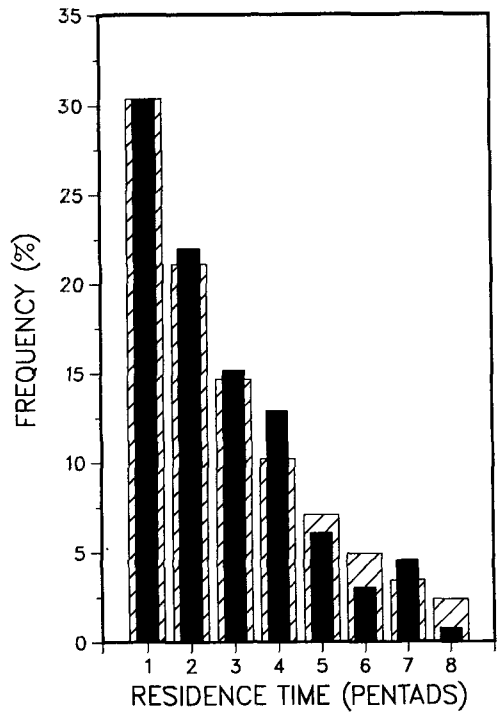


FIG. 8. Black bars: distribution of residence time in either the "positive $E_5$ regime" or the "negative $E_5$ regime" (defined on the basis of the minimum in the PDF in Fig. 7b) for the observed atmospheric fields. Shaded bars: best fit to the observed distribution with an exponential distribution.

one has two ridges over the eastern Pacific and western Europe, corresponding to the lows in Fig. 6b. In fact, the difference map between the two fields (shown in Fig. 9c) is very similar to the eddy component of $E_5$ (Fig. 6b), and also similar to the difference map between the averages of all atmospheric fields in the "positive $E_5$ regime" and of those in the "negative $E_5$ regime," displayed in Fig. 9d. These figures show that in this period a transition occurred between two large-scale hemispheric anomalies aligned along an axis close to $E_5$, although in this particular example the amplitude of the anomalies is larger and the transition time longer than typical values.

We now want to discuss the importance of the variations in PV sources and sinks due to diabatic processes in the maintenance of large-scale anomalies with opposite projections on $E_5$. Figures 9e and 9f show the mean sea level pressure fields in the two 5-day periods considered above. The pressure field corresponding to the positive phase of $E_5$ has a very deep Aleutian low and a strong Siberian anticyclone. These two features almost disappear in the surface field corresponding to the negative phase of $E_5$: the Pacific low, in particular, is much weaker and shifted over the Kamchatka peninsula. Over the North Atlantic, the surface low is shifted westward close to the Canadian coast in the positive $E_5$ phase, with another minimum over the central Mediterranean Sea. As discussed more thoroughly in Marshall and So (1990), a decrease in the intensity of the surface winds is indicative of a decrease in the intensity of diabatic processes. At least over the northeast Pacific and Siberia, the negative phase of $E_5$ appears to be closer to a free-mode state than the positive phase.

In the QG model, the amplitude of the variations of PV sources and sinks in opposite phases of $E_5$ can be easily evaluated. For this neutral vector, we computed the contribution to the linear tendency of streamfunction from the (linearized) PV advection term and the relaxation–dissipation term separately. We found an almost complete compensation between the amplification induced by PV advection and the damping due to diabatic processes, so that the total linear tendency is more than one order of magnitude smaller than each of the two terms.

## 6. Low-frequency variability in the nonlinear QG model

### *a. Simulation of the wintertime climatology*

In the previous section, neutral vectors of the linearized QG model have been used to analyze observed data, in order to find observational evidence of the existence of multiple quasi-stationary solutions. In this and the following section, we shall investigate whether the observed regimes can be simulated in an integration of the full, nonlinear QG model.

As discussed in section 4, in order to integrate the nonlinear model, one must provide an estimate of the PV sources $S_i$ that appear in Eq. (18). This can be done by requiring that, for a given ensemble of observed fields, the average PV tendencies vanish. If the overbar represents the average over the ensemble, and a prime denotes a deviation from the mean, one can write such condition as

$$\overline{\frac{\partial q}{\partial t}} = -J(\bar{\psi}, \bar{q}) - \overline{J(\psi', q')} - D(\bar{\psi}) + S = 0, \quad (20)$$

where for simplicity we have omitted the index of the vertical level. The term

$$J' = -\overline{J(\psi', q')} \quad (21)$$

represents the contribution to the mean tendency of the variability within the ensemble.

The procedure suggested by Roads (1987) to compute $S$ is equivalent to set $\bar{\psi}$ equal to the observed climatological mean and to compute $J'$ using the full range of observed variability. An alternative strategy, frequently used for barotropic models (see, for example, Simmons et al. 1983), is to neglect the term $J'$ completely and assume that the observed climatology (or any other chosen basic state) is an exact stationary solution.

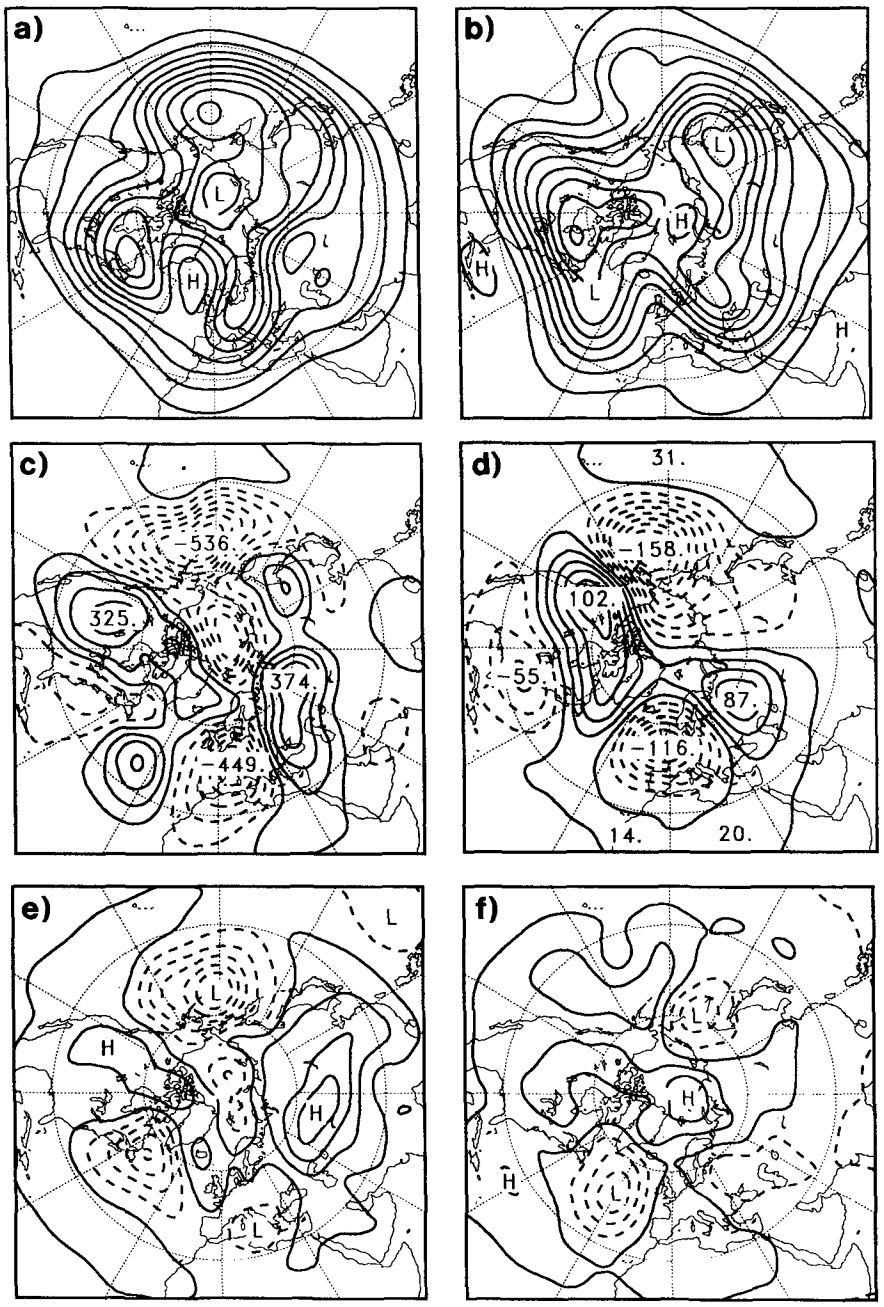


FIG. 9. (a) Mean map of 500-hPa height in the period 5–9 January 1985; (b) as in (a) but for the period 2–6 February 1985; (c) difference between (a) and (b); (d) difference map between the average of the 500-hPa height fields in the "positive $\mathbf{E}_5$ regime" and the average for the "negative $\mathbf{E}_5$ regime"; (e) and (f) as in (a) and (b), respectively, but for mean sea level pressure. (Contour interval is 8 hPa; values below 1013 hPa are dashed.)

A model will actually reproduce a mean state equal to $\bar{\psi}$ only if its internal variability is identical (more precisely, if it produces the same $J'$) to the variability within the ensemble used for the computation of $S$. In the case of a global, multilevel QG model, $J'$ cannot be neglected, but not all forms of observed variability can be reproduced (particularly in tropical areas). Therefore, we shall try to include in $S$ only the contributions of modes of variability that can be simulated in our model. In addition, we want to use the properties of neutral vectors in order to make the PV sources compatible with the existence of more than one regime, as suggested by the bimodality along the neutral vector $\mathbf{E}_5$.

To achieve this, we shall require that a 5-day mean state defined as the observed climatological mean $\psi_c$ plus the vector $\mathbf{E}_5$ (with a 5 m s$^{-1}$ amplitude, as in Fig. 4) and a second 5-day mean state defined as $\psi_c$ minus $\mathbf{E}_5$ (with the same amplitude) are both quasi-stationary solutions of the QG model, in the presence of baroclinic eddies with periods shorter than 5 days. These conditions can be written as

$$-J(\psi_c, q_c) - D(\psi_c) + S + L(\psi_c, \mathbf{E}_5) - J[\mathbf{E}_5, q(\mathbf{E}_5)] + J'_+ \approx 0 \quad (22a)$$

$$-J(\psi_c, q_c) - D(\psi_c) + S - L(\psi_c, \mathbf{E}_5) - J[\mathbf{E}_5, q(\mathbf{E}_5)] + J'_- \approx 0, \quad (22b)$$

where $L$ is the linear operator defined in Eq. (3) and $J'_+$ and $J'_-$ are the contributions from high-frequency baroclinic eddies in the two 5-day mean states. Since $\mathbf{E}_5$ is a neutral vector, the term $L(\psi_c, \mathbf{E}_5)$ has a very small magnitude. To proceed further, we assume that $J'_+$ and $J'_-$ are similar and can be approximated by an average term $J'_c$ (to be computed from a large sample of high-frequency transients). We finally define

$$S = J(\psi_c, q_c) + D(\psi_c) + J[\mathbf{E}_5, q(\mathbf{E}_5)] - J'_c. \quad (23)$$

The term $J'_c$ has been computed from our sample of observed streamfunction fields (January and February 1984 to 1989) as the difference between the average PV tendency of daily fields and the average PV tendency of 5-day means. Therefore, it represents the forcing of the mean PV field by high-frequency transients defined as deviations from 5-day means. The fact the $J'_c$ appears with a negative sign in the definition of $S$ means that we expect the QG model to generate a term close to $J'_c$ by its own high-frequency baroclinic eddies. In this case, the sum of the eddy forcing and the constant source $S$ should approximately compensate the PV tendency due the 5-day mean flow in two states aligned along the $\mathbf{E}_5$ vector; low-frequency variations should therefore occur along this axis.

If our definition of $S$ is appropriate, the QG model should indeed reproduce the observed wintertime mean state and the low-frequency and high-frequency variations of streamfunction (at least in the northern extratropics) in a realistic way, when a long integration is performed with $S$ given by Eq. (23). Such a "perpetual winter" integration has been carried out for 4500 days (equivalent to 50 90-day winters), after a spinup period of 360 days. Figure 10 shows the model-generated mean zonal wind at the three pressure levels, averaged over the 50 winters, and compared with the observed climatological values for January–February. Given the simplicity of the model, the agreement is remarkably good, the main discrepancy being an underestimation of the Southern Hemisphere jet in the lower troposphere and a northward shift of its maximum by about 10 degrees.

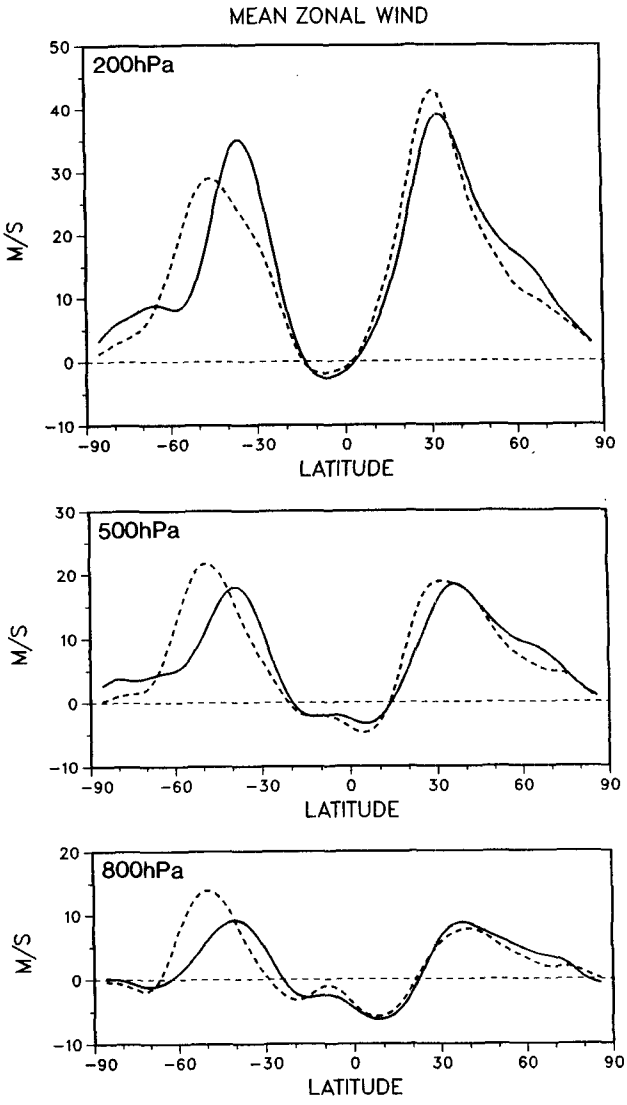


FIG. 10. Mean zonal wind at 200 (top), 500 (center), and 800 hPa (bottom) from a 50-winter nonlinear integration of the QG model (solid lines), and from the observed climatology for January–February (dashed line).

Figure 11 shows maps of the mean and standard deviation of the 500-hPa streamfunction from model and observed fields. The standard deviation is shown separately for 5-day mean fields and deviations from 5-day means, in order to evaluate the low-frequency and high-frequency components of the model variability. Overall, the model mean field has a realistic pattern, but the amplitude of the stationary waves is underestimated. The model distribution of low-frequency variability is very good, with the three maxima in the North Pacific, North Atlantic, and northwestern Siberia (around 60°E) plausibly reproduced. The maxima are about 20%–30% lower than the observed ones, but one must take into account that interannual variations related to surface forcing cannot be simu-

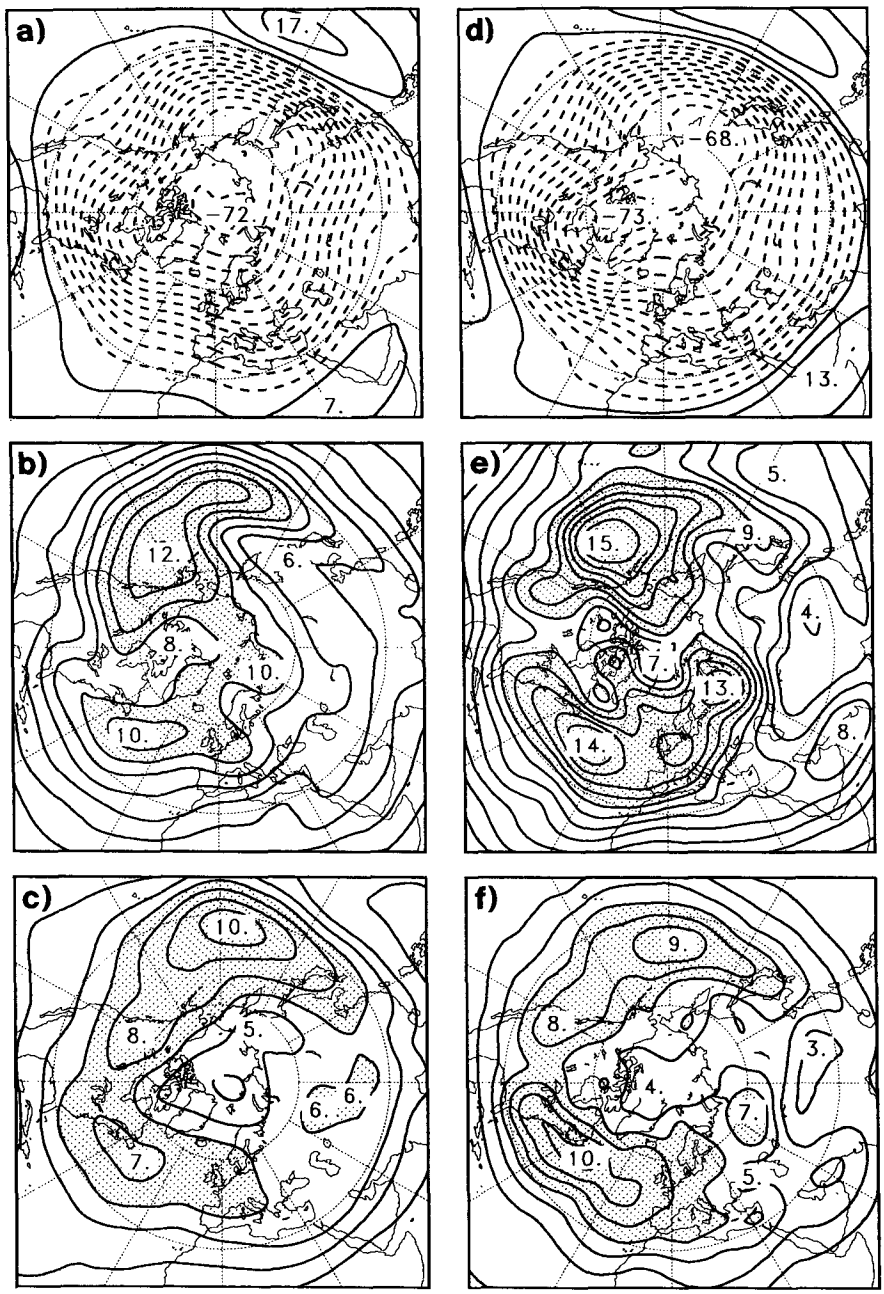


FIG. 11. Statistics of 500-hPa streamfunction ($10^6$ m$^2$ s$^{-1}$) from the 50-winter integration of the QG model and from observed fields in January–February 1984–89. (a) Mean field of the QG model; (b) low-frequency standard deviation in the QG model (from 5-day mean fields); (c) high-frequency standard deviation in the QG model (from deviations from 5-day means); (d), (e), and (f) as in (a), (b), and (c) but computed from observed fields.

lated by this model. Finally, high-frequency variability is of the correct magnitude in the Pacific sector, somewhat underestimated in the Atlantic, but with no clear separation between the two storm tracks over the American continent.

Does the definition of $S$ in Eq. (23) give more satisfactory results than alternative ones? An integration in which $S$ was defined without taking into account the effect of high-frequency eddies [i.e., neglecting the term $J_c'$ in Eq. (23)] showed, as expected, a much poorer zonal mean wind over the Northern Hemisphere and a very weak pattern of stationary waves. A definition based simply on the average tendency for all daily fields gave practically the same mean flow but a

less realistic variability in the Northern Hemisphere, and even larger errors in the mean over the Southern Hemisphere. In summary, among those tested, the PV source defined by Eq. (23) has provided the most realistic climatology overall, both in terms of the mean and the variance.

### b. *Regimes in the QG model*

We have discussed in section 6a how the PV source for the nonlinear QG model has been computed on the assumption of the existence of two quasi-stationary states in the atmospheric flow. Here, we will show that (at least) two regimes indeed exist in the model circulation, similar to those found in the real atmosphere.

As done previously, we looked for the existence of multimodality in planetary waves by projecting model fields from the 50-winter nonlinear integration onto the eddy components of neutral vectors. In this case, the neutral vectors were computed using the time-mean streamfunction of the 50-winter model integration as a basic state. For consistency with the data analysis in section 5, we converted 500-hPa streamfunction fields into geopotential height by solving the linear balance equation, and we constructed a sample of 900 nonoverlapping 5-day means from daily data. These fields were projected onto the eddy components of neutral vectors (also converted to geopotential) in the area 20°–90°N.

Before discussing the results of the projections, we want to describe briefly the structure of some of the neutral modes computed from the model climatology. Figures 12a–d show the 500-hPa eddy height fields for neutral vectors 1, 2, 4, and 6 (again, they are in order of increasing amplitude of the linear time derivative). We shall call these vectors $\mathbf{E}_{1m}$, $\mathbf{E}_{2m}$, . . . etc. The eddy fields can be compared with those shown in Figs. 6a and 6b for vectors $\mathbf{E}_1$ and $\mathbf{E}_5$ from the observed climatology.

The neutral vector $\mathbf{E}_{1m}$ is quite similar to $\mathbf{E}_1$, although it has a much lower amplitude over North America. Vector $\mathbf{E}_{2m}$ is strongly negatively correlated with $\mathbf{E}_{1m}$ over the Northern Hemisphere (but positively correlated over the Southern Hemisphere, the two vectors being orthogonal on the global domain); however, it has less amplitude than $\mathbf{E}_{1m}$ over northeast Asia, and most of its spatial variance is in the Atlantic, European, and west Asian regions. Vector $\mathbf{E}_{4m}$ has a quite similar structure to that of $\mathbf{E}_5$, but over the Atlantic Ocean its highs and lows are more zonally elongated than those in $\mathbf{E}_5$, so that zonal wavenumber 2 is more evident than wavenumber 3 in $\mathbf{E}_{4m}$ (the opposite is true for $\mathbf{E}_5$). A zonal wavenumber 3 pattern appears in $\mathbf{E}_{6m}$, but its structure is clearly orthogonal to both $\mathbf{E}_{4m}$ and $\mathbf{E}_5$. The characteristic time $\lambda^{-1}$ for these neutral vectors is 1880, 602, 350, and 205 days, respectively.

The projections of the model-generated eddy height fields on the first 10 neutral vectors did not reveal bimodality along any individual vector. Quasi-stationary states may be aligned along a linear combination of neutral vectors, however. Therefore, we decided to examine two-dimensional PDFs of the projections of eddy fields onto the plane generated by two neutral vectors. In the choice of the two axes, one can be guided by three criteria:

1) the ordering of the neutral vectors as determined by their characteristic time (i.e., by the magnitude of their time derivative);
2) the similarity with the pattern along which regimes have been found in the real atmosphere; and
3) the variance of the projections onto individual neutral vectors.

The first criterion obviously leads to the selection of $\mathbf{E}_{1m}$; the second one suggests selection of $\mathbf{E}_{4m}$. It turned out that the projection of eddy fields on $\mathbf{E}_{4m}$ had the largest variance among all the neutral vectors considered. On this basis, and given the similarity between the first and the second neutral vector over the Northern Hemisphere, we chose to project the model eddy fields onto the plane generated by neutral vectors $\mathbf{E}_{1m}$ and $\mathbf{E}_{4m}$.

Since these two vectors are not exactly orthogonal when considered as eddy geopotential fields in the region 20°–90°N, they were reorthonormalized on this domain. We chose to keep the pattern of $\mathbf{E}_{4m}$ unchanged; in the following, the projection on the eddy part of $\mathbf{E}_{4m}$ will be the $y$ coordinate of our plane, while the $x$ coordinate corresponds to the eddy component of $\mathbf{E}_{1m}$ that is orthogonal to $\mathbf{E}_{4m}$. The two renormalized axes are shown in Figs. 12e, f.

Estimations of two-dimensional PDFs are obviously more subject to sampling problems than unidimensional PDFs. To avoid noisy estimates in regions of low density, we used an iterative version of the kernel estimator (see Silverman 1986), described in appendix B. In addition, to test the stability of our results, we computed the PDF for various subsamples including 10 or 20 consecutive winters (again, 1 winter is 90 days).

The inspection of the PDFs in these subsamples revealed an unexpected behavior of the model regimes. The PDF for the first 10 winters (Fig. 13a), although clearly different from a multinormal function, is essentially unimodal, whereas all the other 10-winter PDFs (not shown) possess at least three density maxima. A more stable picture is obtained from the PDFs computed from the two subsamples including winters 11–30 and 31–50, which are shown in Figs. 13b and 13c, respectively. Disregarding maxima defined in a very small neighborhood, two major maxima appear in the PDF for winters 11–30, aligned along an axis roughly parallel to $(\mathbf{E}_{1m} + \mathbf{E}_{4m})$. In the PDF for winters 31–50, the two most important maxima are aligned along a similar axis, but their separation is much larger. A third significant maximum appears close to $x = 10$,

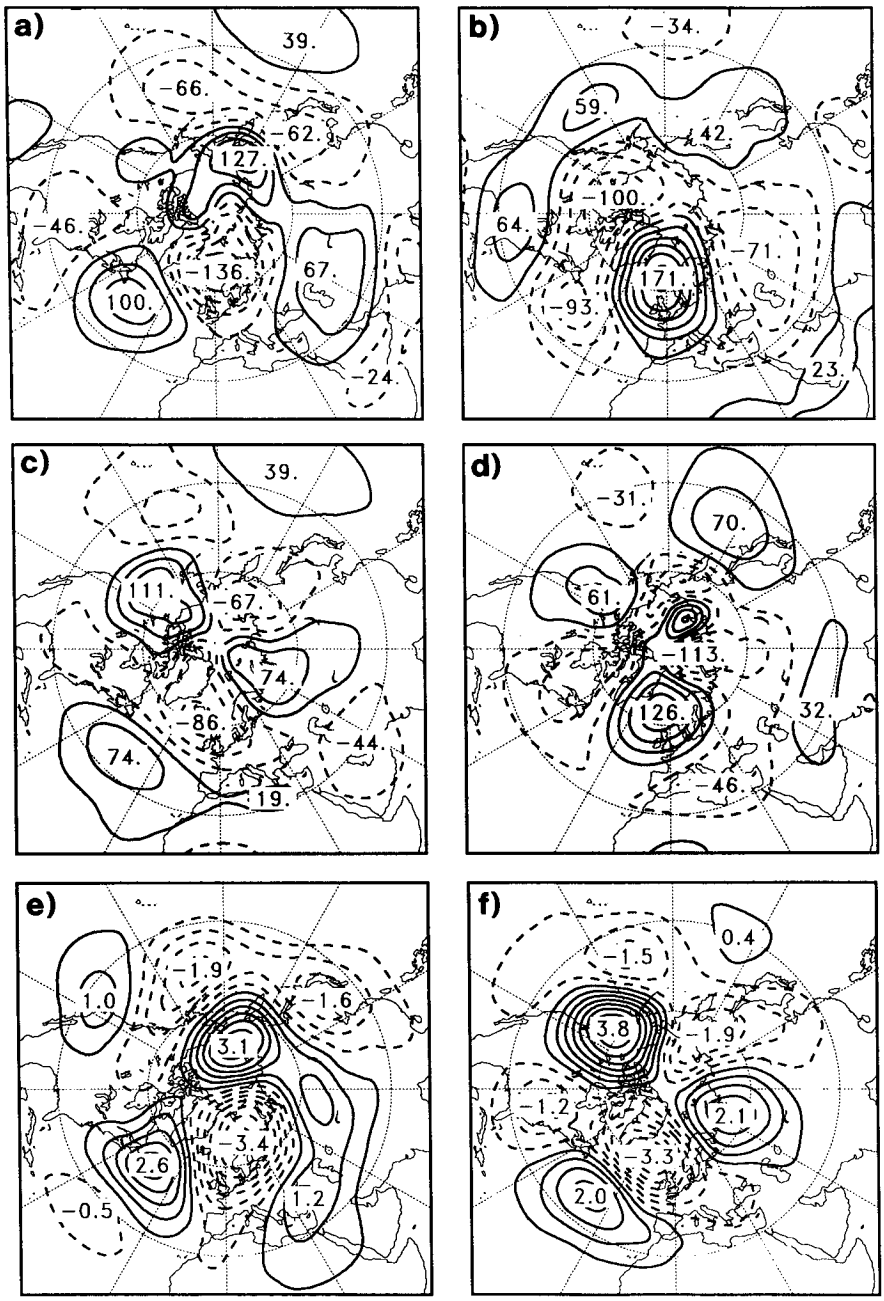


FIG. 12. Eddy geopotential height at 500 hPa corresponding to the neutral vectors (a) $\mathbf{E}_{1m}$, (b) $\mathbf{E}_{2m}$, (c) $\mathbf{E}_{4m}$, and (d) $\mathbf{E}_{6m}$, with a (rms wind) norm of 5 m s$^{-1}$; (e) and (f): as in (a) and (c), respectively, but reorthonormalized in the area 20°–90°N (rms amplitude is 1).

$y = -20$, not far from the position of a less evident maximum in the PDF of the previous 20 winters.

In order to examine whether the apparent trend toward a greater separation of the model regimes persisted, and whether the unimodality in the first 10 years was really an exception, we decided to continue the integration of the QG model for 40 more winters. The time-mean state of the model in these winters was virtually identical to the mean state in the previous 50 winters, so there was no need to recompute the neutral vectors. The new streamfunction fields were again averaged over pentads, converted into geopotential height, and projected onto the same axes used for the analysis of the past 50 winters.

The PDF of the projections of eddy fields in winters 51–90 on the $\mathbf{E}_{1m}$–$\mathbf{E}_{4m}$ plane is shown in Fig. 13d. As

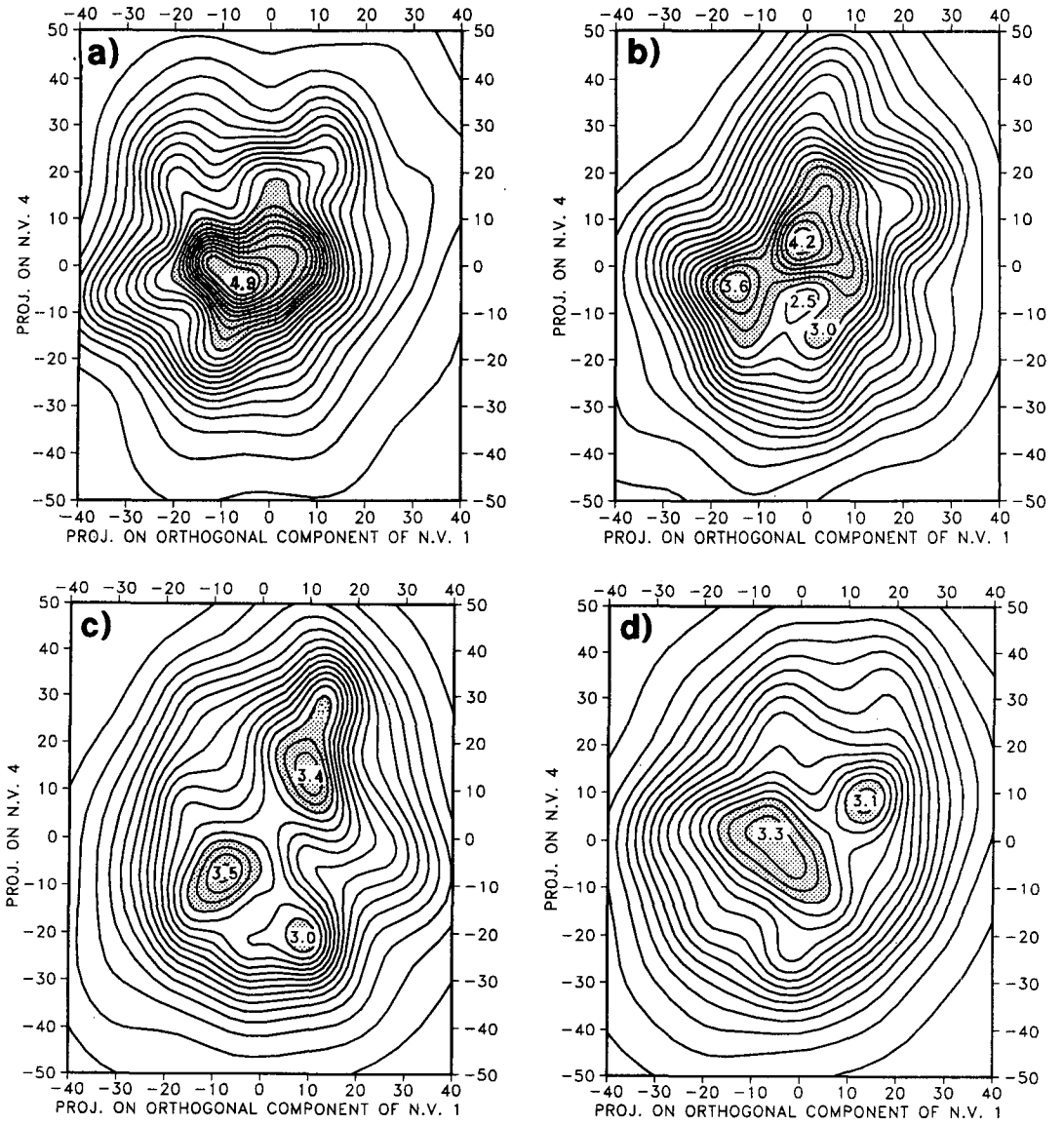


FIG. 13. Two-dimensional PDF of the projection of model-generated eddy fields of 500-hPa height onto the plane generated by neutral vectors $\mathbf{E}_{1m}$ and $\mathbf{E}_{4m}$. The $x$ axis corresponds to the projection onto the component of $\mathbf{E}_{1m}$ that is orthogonal to $\mathbf{E}_{4m}$ in the area 20°–90°N (Fig. 12e), the $y$ axis to the projection on $\mathbf{E}_{4m}$ (Fig. 12f). (a) From data in winters 1–10; (b) winters 11–30; (c) winters 31–50; and (d) winters 51–90 from the extended integration.

in Figs. 13b and 13c, two major regimes are evident along an axis nearly parallel to $(\mathbf{E}_{1m} + \mathbf{E}_{4m})$. Their separation is smaller than in winters 31–50 but larger than in the winters 11–30 subsample. No individual 10-winter subsample from the new data showed a unimodal distribution. Finally, a smooth bimodal PDF can be obtained by merging the data from winters 11–90; this is shown in Fig. 14.

One can conclude that, despite oscillations in the exact position of the maxima, the model possesses two fairly stable regimes that are aligned along a linear combination of its first and fourth neutral vector. Since the time in which the model appeared to have reached a stable mean state, however, it took about ten more winters to stabilize its low-frequency variability, so that regimes became evident in two-dimensional density estimates. With regard to this long adjustment time, we point out that this phenomenon was already noticed in a long perpertual winter integration of a GCM, namely, the CCM0B model developed at the U.S. National Center for Atmospheric Research (Hansen, private communication). In this case, a bimodality in the amplitude of planetary waves that appeared to be statistically significant in a 1200-day integration (see

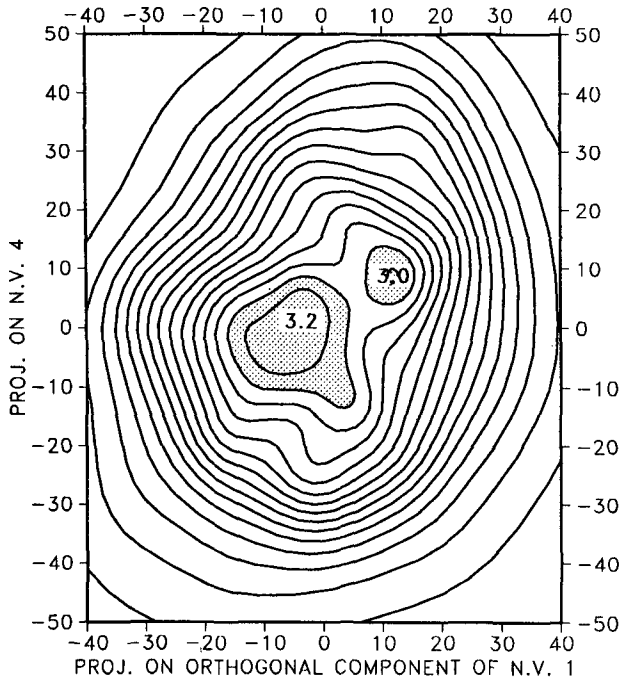


FIG. 14. As in Fig. 13 but from data in winters 11–90.

Hansen and Sutera 1990) could not be found any longer when the integration was extended to a total of 6000 days.

From the results discussed so far, we can deduce that the sample including data from winters 11 to 90 (for a total of 1440 pentads) is sufficiently homogeneous, and assume the PDF in Fig. 14 as a reliable estimate of the climatological PDF of the model. We shall now examine whether the axis along which the QG model has bimodality is similar to the pattern along which bimodality was found in the observed data.

Figure 15a shows the difference between the patterns corresponding to the two density maxima in Fig. 14; this pattern is simply a linear combination of (eddy) neutral vectors $\mathbf{E}_{1m}$ and $\mathbf{E}_{4m}$. In order to compare the difference between the centroids of the two regimes of the QG model with the corresponding map for the real atmospheric regimes described in section 5, we have classified each field in the 80-winter sample in either of the two regimes defined by the density maxima in Fig. 14. To do so, starting from the projection of a given field in the $\mathbf{E}_{1m}$–$\mathbf{E}_{4m}$ plane, one can "climb" the PDF along the direction of its gradient until a density maximum is reached.

The difference between the averages of the (full) height fields classified in the two regimes is shown in Fig. 15b. Comparing this with Fig. 9d (the equivalent map for the observed regimes), one can see that the position of midlatitude highs and lows is in very good agreement. The similarity is definitely weaker north of 60°N, because of the different meridional extent of the various midlatitude features. In particular, the mid-Pacific low is much deeper in the atmospheric than in the model difference map, and extends farther north.

In a number of previous works (Mo and Ghil 1987, 1988; Molteni et al. 1988, 1990; Kimoto 1989), atmospheric regimes have been sought by analyzing PDFs or performing cluster analysis on linear subspaces generated by the leading EOFs of atmospheric fields. We can ask ourselves whether the regimes found along neutral vectors agree with at least some of the clusters found in such EOF-based studies. Figure 16 shows the patterns of clusters 1 and 2 found by Mo and Ghil (1988), and of clusters C'2 and C'5a found by Molteni et al. (1990). These maps confirm the existence of two atmospheric regimes with opposite anomalies, and a hemispheric spatial pattern which is in good agreement with the axes along which the centroids of the observed and the model-generated regimes found in this study are aligned (see Figs. 9d and 15b, respectively). The fact that our regimes appear to be less separated than

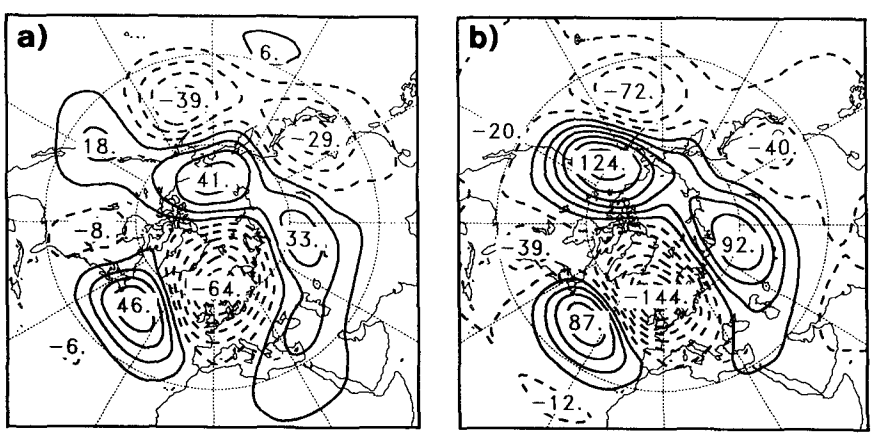


FIG. 15. (a) Difference between the two patterns corresponding to the density maxima in Fig. 14; (b) difference between the centroids of the two clusters defined from the density maxima.

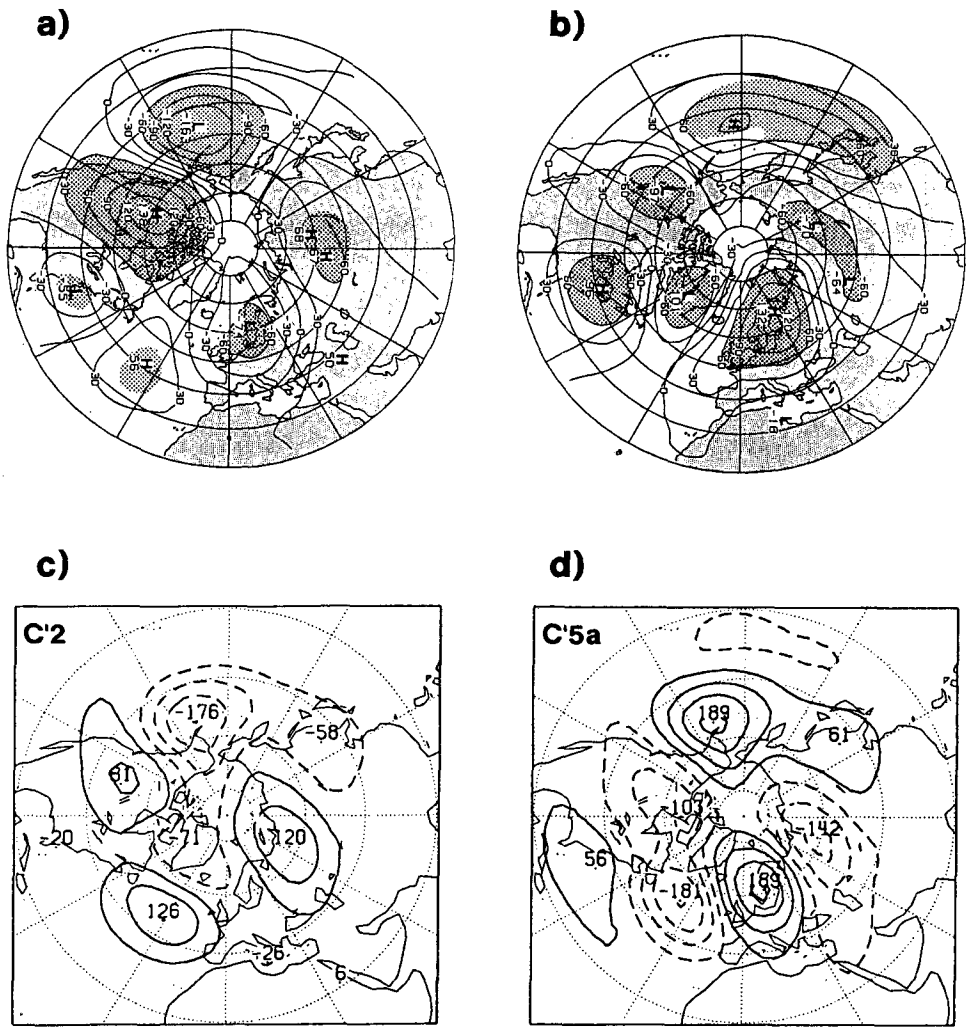


FIG. 16. Anomalies corresponding to the centroids of (a) cluster 1 and (b) cluster 2, found by Mo and Ghil (1988), and of clusters C'2 (c) and C'5a (d) found by Molteni et al. (1990). In (a) and (b), shading indicates statistical significance at 95% confidence level.

the EOF-based regimes is mainly due to the fact that here we are considering a two-regime partition, rather than a multiregime partition, of the data.

Given the simplified nature of the QG model, the agreement between its regimes and the atmospheric regimes is clearly satisfactory, as far as their spatial pattern is concerned; however, the similarity becomes striking when the distribution of residence times is compared. The distribution for the model regimes is shown in Fig. 17. As for the observed regimes defined by $E_5$ (see Fig. 8), the black bars in the histogram show the actual frequencies of different residence times, while the shaded bars in the background give the best fit with an exponential distribution. The exponential distribution, which fits the model data very closely, is practically indistinguishable from the distribution fitted to the observed data, and the mean residence time in the model regimes is 3.2 pentads (16 days), identical to the observed value. Quasigeostrophic dynamics seems therefore able to explain successfully both spatial and temporal aspects of planetary-scale regimes in the wintertime circulation over northern midlatitudes.

## 7. Discussion

In this work, we have shown how a property of stationary solutions in dynamical systems with quadratic nonlinearity (namely, the fact that the difference vector between two such states has a null time derivative in the equations linearized around the middle point) can be exploited to select a linear subspace of phase space in which quasi-stationary states can be sought. In the case of Lorenz's three-variable dynamical system, the most neutral vector of the operator linearized around the time-mean state of the system (i.e., the vector with the smallest amplitude of the linear time derivative)

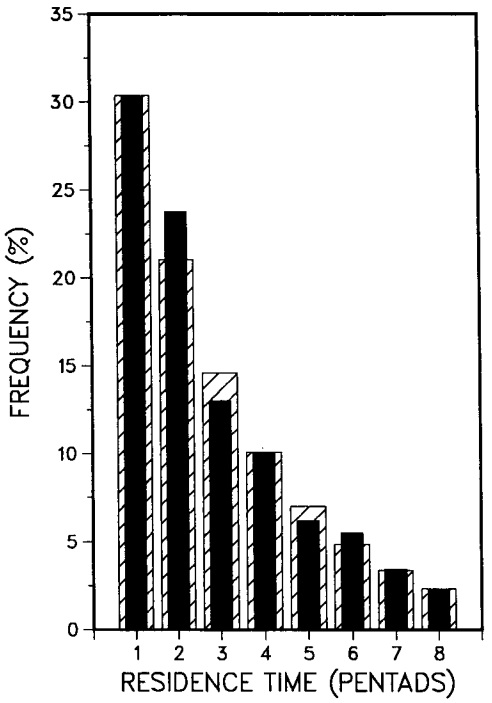


FIG. 17. As in Fig. 8 but for the residence time in the two regimes (clusters) of the QG model.

gives an excellent approximation to the difference vector between the two stationary solutions. In the case of a three-level QG model linearized about the mean state of the real atmosphere, a statistically significant bimodality can be found in the projection of observed eddy fields onto the fifth neutral vector; this vector can be described as a combination of the PNA teleconnection pattern with a second wave pattern covering the Atlantic and Eurasia.

Bimodality found along a neutral vector of a linearized model is *dynamically consistent* with the existence of two stationary solutions in the nonlinear version of the model. The condition for the actual existence of such solutions is the presence of an appropriate source term for the model variables (in our case PV). We have computed a PV source for the QG model that could maintain (under certain hypotheses) two quasi-stationary solutions corresponding to opposite phases of the neutral vector along which bimodality was found in the observed data. This source term proved to be an appropriate one: a long integration of the nonlinear model generated a realistic time-mean field and variance distribution, as well as two regimes with a very similar spatial pattern and almost identical distribution of residence times compared to the observed regimes. Again, the model regimes could be detected by the presence of bimodality in the PDF of the projections of eddy fields onto the plane generated by two neutral vectors computed from the model time-mean state.

Although our classification reveals a coarser structure of phase space than the one proposed in other studies, the important achievement of this work is that, thanks to the dynamical properties of neutral vectors, we have been able to demonstrate that our regimes can be maintained in a realistic QG model in the neighborhood of two quasi-stationary solutions. This result provides a strong dynamical support to the existence of planetary-scale flow regimes: our QG model does not suffer from such serious problems of severe truncation or unrealistic boundary conditions as simpler dynamical models used in earlier studies. On the other hand, the empirical way in which the PV forcing has been determined is not fully satisfactory, and a more physically based definition of the PV sources and sinks is needed to achieve a deeper understanding of the maintenance of multiple regimes. This is particularly important in view of the crucial role forcing plays in selecting which of the neutral vectors corresponds to the axis linking stationary states with opposite anomalies.

The role of transient eddies also deserves further investigation. A possible weak point in our analysis is the assumption that the PV forcing by high-frequency baroclinic eddies is very similar in different quasi-stationary states. It has been shown, in both diagnostic and modeling studies, that this is not the case for the alternation between blocked and zonal regional regimes (Green 1977; Illari and Marshall 1983; Shutts 1986; Haines and Marshall 1987; Hoskins and Sardeshmukh 1987; Vautard and Legras 1988). Given the fact that neutral vectors explain the patterns of at least some of the observed hemispheric regimes but the maxima in the PDF are found for low-amplitude anomalies, it seems plausible that quasi-stationary solutions of the large-scale flow (in the presence of nearly constant eddy forcing) can explain the location of the PDF maxima, and therefore the preferred hemispheric patterns of persistent regimes, but variations in the eddy forcing are probably needed to amplify such anomalies to the stage in which blocking highs develop from the large-scale ridges of the hemispheric regimes.

Finally, we point out that the concept of neutral vectors, and the possibility that some of them are the difference patterns between nonlinear quasi-stationary states, places in a new perspective the results of a number of previous studies based on linearized models. Since the work of Hoskins and Karoly (1981), many investigators have tried to explain patterns of extratropical low-frequency variability in terms of the steady response of a linear, adiabatic model to a prescribed anomaly in diabatic forcing. In a recent and intriguing study, Branstator (1990) showed that some of the EOFs derived from a long integration of the NCAR Community Climate Model can be reproduced by the EOFs of a large sample of fields computed as linear responses to random forcing fields. One could argue, however, that the typical magnitude of the observed anomalies

does not justify a linear approximation, and that the distribution of anomalous diabatic forcing is to a large extent a function of the large-scale flow.

With the possible exception of studies concerned with the response to SST anomalies, one can reinterpret many of these results by saying that the steady linear response of an adiabatic model to an anomalous forcing just "selects" the neutral vectors that can balance this forcing by PV advection. In reality, the change in mid-latitude and subtropical diabatic forcing may be the *result* of the switch between two quasi-stationary states with nearly opposite anomalies; in this case, the nonlinear advection terms can be balanced by a constant source that does not appear explicitly in the linear model. From this point of view, the results of linear studies are not necessarily incompatible with a nonlinear dynamical description of low-frequency variability. With an appropriate interpretation, linear models can still be very useful in the understanding of planetary flow regimes.

*Acknowledgments.* We would like to thank Drs. A. White and W. Verkley for many fruitful discussions, and Dr. A. Trevisan for her suggestion to analyse the Lorenz model. We are also grateful to Profs. E. Lorenz and R. Lindzen, and to Drs. A. Hollingsworth, T. N. Palmer, and J. Tribbia for their useful comments on the manuscript.

## APPENDIX A

### Basic Principles and Dissipative Terms in the QG Model

The formulation of the quasigeostrophic model used in this study is based on standard filtered prognostic equations for vorticity and temperature. Expressing relative vorticity as the Laplacian of streamfunction, and temperature in terms of the vertical derivative of geopotential through the hydrostatic relationship (e.g., Haltiner and Williams 1980), the vorticity and thermodynamic equations can be written as

$$\frac{\partial}{\partial t}\nabla^2\psi = -\mathbf{V}_\psi\cdot\nabla(\nabla^2\psi + f) + f_0\frac{\partial\omega}{\partial p} \qquad \text{(A1)}$$

$$\frac{\partial}{\partial t}\frac{\partial\Phi}{\partial p} = -\mathbf{V}_\psi\cdot\nabla\frac{\partial\Phi}{\partial p} - \sigma\omega, \qquad \text{(A2)}$$

where all symbols have their usual meteorological meaning, and dissipative and forcing terms are omitted for simplicity. In particular, we note that a constant value $f_0$ is used in the stretching term of Eq. (A1), and the static stability parameter $\sigma$ in Eq. (A2) is assumed to be a function of pressure only.

When Eqs. (A1) and (A2) are combined using the relationship:

$$\nabla\Phi = f_0\nabla\psi \qquad \text{(A3)}$$

one obtains the following prognostic equation for QG potential vorticity:

$$\left(\frac{\partial}{\partial t} + \mathbf{V}_\psi\cdot\nabla\right)\left(\nabla^2\psi + f + f_0^2\frac{\partial}{\partial p}\sigma^{-1}\frac{\partial\psi}{\partial p}\right) = 0. \qquad \text{(A4)}$$

It is well known that a constant value $f_0$ must be used in the stretching term of (A1) and in (A3) in order to obtain energetic consistency (see Mak 1991 for a thorough discussion of the quasigeostrophic approximation on the sphere). However, since the *squared* value of $f_0$ appears in Eq. (A4), the PV equation appropriate to the Northern Hemisphere is also appropriate in the Southern Hemisphere if one assumes $(f_0)_{SH} = -(f_0)_{NH}$. Our global QG model can be considered as the result of coupling two hemispheric models at the equator, where continuity of the streamfunction and PV fields is assumed. Although here our primary interest is the Northern Hemisphere flow, the global domain avoids the introduction of boundary conditions at the equator (assumed, for example, in the hemispheric model of Marshall and So 1990), which might affect the structure of planetary-scale waves and regimes. It also permits direct use of observed streamfunction fields for the construction of basic states and forcing terms.

The vertical discretization of Eqs. (A1)–(A4) at three levels is standard, and leads to Eqs. (18a)–(18c) presented in section 4a. In these equations, dissipative and forcing terms have been added to represent the effects of diabatic processes. The dissipative terms, indicated symbolically as $-D_1$, $-D_2$, $-D_3$ at the three pressure levels of 200, 500, and 800 hPa, respectively, include contributions from temperature relaxation, Ekman dissipation, and horizontal diffusion:

$$-D_1 = \mathrm{TR}_{12} - H_1 \qquad \text{(A5)}$$

$$-D_2 = -\mathrm{TR}_{12} + \mathrm{TR}_{23} - H_2 \qquad \text{(A6)}$$

$$-D_3 = -\mathrm{TR}_{23} - \mathrm{EK}_3 - H_3. \qquad \text{(A7)}$$

The term:

$$\mathrm{TR}_{12} = \tau_R^{-1}R_1^{-2}(\psi_1 - \psi_2) \qquad \text{(A8)}$$

represents the effect of temperature relaxation between levels 1 and 2, with a radiative time scale $\tau_R = 25$ days; the corresponding term for temperature relaxation between levels 2 and 3 is

$$\mathrm{TR}_{23} = \tau_R^{-1}R_2^{-2}(\psi_2 - \psi_3). \qquad \text{(A9)}$$

Ekman dissipation is expressed as the vorticity tendency due to a linear drag on the 800-hPa wind:

$$\mathrm{EK}_3 = (a\cos\phi)^{-1}\left\{\frac{\partial}{\partial\lambda}[k(\lambda,\phi,h)v_3] - \frac{\partial}{\partial\phi}[k(\lambda,\phi,h)u_3\cos\phi]\right\}, \qquad \text{(A10)}$$

where $\lambda$ is longitude, $\phi$ latitude, $h$ the orographic height, $a$ the average earth radius, and

$$u_3 = -a^{-1}\frac{\partial \psi_3}{\partial \phi}, \quad v_3 = (a\cos\phi)^{-1}\frac{\partial \psi_3}{\partial \lambda}.$$

The drag coefficient $k$ is dependent on the land–sea mask and on the orographic height:

$$k(\lambda, \phi, h) = \tau_E^{-1}[1 + \alpha_1 \mathrm{LS}(\lambda, \phi) + \alpha_2 \mathrm{FH}(h)], \tag{A11}$$

where $\tau_E = 3$ days, $\alpha_1 = \alpha_2 = 0.5$; $\mathrm{LS}(\lambda, \phi)$ is the fraction of land within a grid box; and

$$\mathrm{FH}(h) = 1 - \exp[-h/(1000\ \mathrm{m})].$$

Since LS and FH vary between 0 and 1, $k$ varies between $(3\ \mathrm{days})^{-1}$ over the oceans, $(2\ \mathrm{days})^{-1}$ over zero-altitude land, and about $(1.5\ \mathrm{day})^{-1}$ over mountains higher than 2000 m. [The gridpoint values of orography and land–sea mask are derived from the spectral T21 representation of the actual high-resolution fields, and represent averages over areas of the order of $(1000\ \mathrm{km})^2$.] If one neglects the spatial variations in the drag coefficient by setting $\alpha_1 = \alpha_2 = 0$, Eq. (A11) reduces to the simple form:

$$\mathrm{EK}_3 = \tau_E^{-1}\nabla^2\psi_3. \tag{A12}$$

Finally, at each pressure level, the time-dependent component $q_i'$ of PV (i.e., PV minus planetary vorticity and orographic component) is subject to a scale-selective horizontal diffusion:

$$H_i = c_H\nabla^8 q_i', \tag{A13}$$

where the coefficient

$$c_H = \tau_H^{-1} a^8 (21 \cdot 22)^{-4} \tag{A14}$$

is such that spherical harmonics of total wavenumber 21 are damped with time scale $\tau_H = 2$ days.

## APPENDIX B

## The Kernel Estimator of PDFs

Let $\{x_i,\ i = 1, \cdots, N\}$ be a sample of observed values of a variable $x$. The kernel estimate of the PDF of $x$ can be written as

$$\mathrm{PDF}(x) = N^{-1}\sum_1^N{}_i K(x - x_i, r), \tag{B1}$$

where $K(x', r)$ is a kernel function with the following properties:

$$K(x', r) \geq 0$$

$$\int_D K(x', r)dx' = 1$$

$$\int_D K(x', r)x'dx' = 0;$$

$x' = x - x_i$ and $D$ is the domain over which $K$ is greater than zero. The parameter $r$ determines the "radius of influence" of each observed value $x_i$: the larger is $r$, the smoother is the PDF.

In practice, the optimal value of $r$ depends on the number of available data. When one looks for multimodality in a PDF, $r$ should be large enough that if the kernel estimator is applied to a sample of $N$ data extracted at random from a population with a (a priori known) unimodal distribution, the probability of obtaining a multimodal estimate is lower than a given threshold (usually 5% or 10%). This probability, as a function of $r$, can be evaluated by Monte Carlo test (see Silverman 1981).

In this paper, estimates of PDFs in one dimension have been obtained using a Gaussian function as kernel:

$$K(x', r) = (\sqrt{2\pi}r)^{-1}\exp\left[-\frac{1}{2}\left(\frac{x'}{r}\right)^2\right], \tag{B2}$$

where $r = 0.25\ \sigma$, and $\sigma$ is the standard deviation of the sample $\{x_i\}$.

Equation (B1) can be easily generalized to more than one dimension. In two dimensions, given a sample $\{(x_i, y_i),\ i = 1, \cdots, N\}$, one can write

$$\mathrm{PDF}(x, y) = N^{-1}\sum_1^N{}_i K(x - x_i, y - y_i, r_x, r_y) \tag{B3}$$

and assume

$$K(x', y') = (2\pi r_x r_y)^{-1} \times \exp\left[-\frac{1}{2}\left(\frac{x'}{r_x}\right)^2 - \frac{1}{2}\left(\frac{y'}{r_y}\right)^2\right]. \tag{B4}$$

However, for samples of the size used in observational studies of atmospheric low-frequency variability (typically a few hundred independent data), the use of a fixed radius of influence already creates problems in two dimensions: values appropriate to estimate the PDF close to the mean point give a noisy estimate for large deviations from the mean. This problem can be overcome by assuming an influence radius depending on the data density itself (see Silverman 1986). A "first-guess" PDF, computed with a fixed influence radius, can be used to estimate the density $d_i$ at each point $(x_i, y_i)$. The final estimate of the PDF can be obtained as

$$\mathrm{PDF}(x, y) = N^{-1}\sum_1^N{}_i K(x - x_i, y - y_i, f_i r_x, f_i r_y), \tag{B5}$$

where $f_i$ is a factor proportional to $d_i^{-1/2}$. In this way, the area influenced by each observed value $(x_i, y_i)$ is inversely proportional to the local density, and therefore, in each point $(x, y)$ of the plane the number of observations that contribute significantly to the final

density estimate is approximately constant. This iterative technique has been used to compute the two-dimensional PDF shown in Figs. 13 and 14.

## REFERENCES


Barnston, A. G., and R. E. Livezey, 1987: Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.*, **115,** 1083–1126.

Baur, F., 1947: *Musterbeispiele Europaeischer Grosswetterlagen.* Dieterich, 35 pp.

Branstator, G., 1990: Low-frequency patterns induced by stationary waves. *J. Atmos. Sci.*, **47,** 629–648.

——, and J. D. Opsteegh, 1989: Free solution of the barotropic vorticity equation. *J. Atmos. Sci.*, **46,** 1799–1814.

Charney, J. G., and J. G. De Vore, 1979: Multiple flow equilibria and blocking. *J. Atmos. Sci.*, **36,** 1205–1216.

Dole, R. M., and N. D. Gordon, 1983: Persistent anomalies of the extratropical Northern Hemisphere wintertime circulation: Geographical distribution and regional persistence characteristics. *Mon. Wea. Rev.*, **111,** 1567–1586.

Green, J. S. A., 1977: The weather during July 1976: Some dynamical considerations of the drought. *Weather,* **32,** 120–128.

Haines, K., and J. C. Marshall, 1987: Eddy-forced coherent structures as a prototype of atmospheric blocking. *Quart. J. Roy. Meteor. Soc.*, **113,** 681–704.

Hansen, A. R., and A. Sutera, 1986: On the probability density distribution of large-scale atmospheric wave amplitude. *J. Atmos. Sci.*, **43,** 3250–3265.

——, and ——, 1990: Weather regimes in a general circulation model. *J. Atmos. Sci.*, **47,** 380–391.

Horel, J. D., 1981: A rotated principal component analysis of the interannual variability of the Northern Hemisphere 500 mb height field. *Mon. Wea. Rev.*, **109,** 2080–2092.

Hoskins, B. J., and D. J. Karoly, 1981: The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J. Atmos. Sci.*, **38,** 1179–1196.

——, and P. D. Sardeshmukh, 1987: A diagnostic study of the dynamics of the Northern Hemisphere winter of 1985–86. *Quart. J. Roy. Meteor. Soc.*, **113,** 759–778.

Illari, L., and J. C. Marshall, 1983: On the interpretation of eddy fluxes during a blocking episode. *J. Atmos. Sci.*, **40,** 2232–2242.

Kimoto, M., 1989: Multiple flow regimes in the northern hemisphere winter. Ph.D. thesis, University of California, Los Angeles. [Available from the University of California, Los Angeles, Department of Atmospheric Sciences, 405 Hilgard Ave., Los Angeles, CA 90024]

Legras, B., and M. Ghil, 1985: Persistent anomalies, blocking and variations in atmospheric predictability. *J. Atmos. Sci.*, **42,** 433–471.

——, T. Desponts, and B. Piguet, 1988: Cluster analysis and weather regimes. *Proc. of the ECMWF Seminar on the Nature and Prediction of Extratropical Weather Systems,* 123–149.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20,** 130–141.

McIntyre, M. E., 1892: How well do we understand the dynamics of stratospheric warmings? *J. Meteor. Soc. Japan,* **60,** 37–65.

Mak, M., 1991: Influences of the Earth's sphericity in the quasi-geostrophic theory. *J. Meteor. Soc. Japan,* **69,** 497–511.

Marshall, J., and D. W. K. So, 1990: Thermal equilibration of planetary waves. *J. Atmos. Sci.*, **47,** 963–978.

Mitchell, H., and J. Derome, 1983: Blocking-like solutions of the potential vorticity equation: Their stability at equilibrium and growth at resonance. *J. Atmos. Sci.*, **40,** 2522–2536.

Mo, K. C., and M. Ghil, 1987: Statistics and dynamics of persistent anomalies. *J. Atmos. Sci.*, **44,** 877–901.

——, and ——, 1988: Cluster analysis of multiple planetary flow regimes. *J. Geophys. Res.*, **93D,** 10 927–10 952.

Molteni, F., A. Sutera, and N. Tronci, 1988: EOFs of the geopotential eddies at 500 mb in winter and their probability density distributions. *J. Atmos. Sci.*, **45,** 3063–3080.

——, S. Tibaldi, and T. N. Palmer, 1990: Regimes in the wintertime circulation over northern extratropics. I: Observational evidence. *Quart. J. Roy. Meteor. Soc.*, **116,** 31–67.

Moritz, R. E., and A. Sutera, 1981: The predictability problem: Effects of stochastic perturbations in multiequilibrium systems. *Advances in Geophysics,* No. 23, Academic Press, 345–383.

Mukougawa, H., 1988: A dynamical model of "quasi-stationary" states in large-scale atmospheric motions. *J. Atmos. Sci.*, **45,** 2868–2888.

Namias, J., 1950: The index cycle and its role in the general circulation. *J. Meteor.*, **7,** 130–139.

Navarra, A., 1993: A new set of orthonormal modes for linearized meteorological problems. Submitted to *J. Atmos. Sci.*

Rex, D. R., 1950: Blocking action in the middle troposphere and its effect upon regional climate. II: The climatology of blocking actions. *Tellus,* **2,** 275–302.

Roads, J. O., 1987: Predictability in the extended range. *J. Atmos. Sci.*, **44,** 3495–3527.

Schneider, E. K., 1988: A formulation for diagnostic anomaly models. *Pure Appl. Geophys.*, **126,** 137–140.

Shutts, G. J., 1986: A case study of eddy forcing during an Atlantic blocking episode. *Advances in Geophysics,* No. 29, Academic Press, 135–161.

Silverman, B. W., 1981: Using kernel-density estimators to investigate multimodality. *J. Roy. Stat. Soc.*, **B43,** 97–99.

——, 1986: *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, 175 pp.

Sutera, A., 1986: Probability density distribution of large scale atmospheric flow. *Advances in Geophysics,* No. 29, Academic Press, 319–338.

Vautard, R., 1990: Multiple weather regimes over the North Atlantic. Analysis of precursors and successors. *Mon. Wea. Rev.*, **118,** 2056–2081.

——, and B. Legras, 1988: On the source of midlatitude low-frequency variability. Part II: Nonlinear equilibration of weather regimes. *J. Atmos. Sci.*, **45,** 2845–2867.

——, ——, and M. Deque, 1988: On the source of midlatitude low-frequency variability. Part I: A statistical approach to persistence. *J. Atmos. Sci.*, **45,** 2811–2843.

Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **109,** 784–812.

Webster, P. J., and J. R. Holton, 1982: Cross-equatorial response to middle-latitude forcing in a zonally varying basic state. *J. Atmos. Sci.*, **39,** 722–733.