Prediction of the Evolution of the Variance in a Barotropic Model

A. BETTI AND A. NAVARRA

IMGA-CNR, Modena, Italy

(Manuscript received 16 August 1993, in final form 23 June 1994)

ABSTRACT

The Schmidt decomposition is applied to the evolution operator of the linearized barotropic equation on a sphere (in the following referred to as the barotropic propagator) to study the evolution of the variance, that is, of the collective evolution of a cloud of trajectories centered around the initial condition. The variance can give reliable information on the tendency that some initial conditions may have to generate large spreads in the subsequent time evolution, especially when many modes with similarly large amplifying rates exist. It appears rather arbitrary, under these circumstances, to pick a particular mode just because it happens to have the largest rate for that particular numerical formulation and resolution setting. It is also shown that the Golden–Thompson generalized inequality and other indicators can be used to estimate the linear variance from the analysis of the initial condition itself, without the need for performing the costly explicit calculation of the propagator.

Numerical experiments performed on a set of initial conditions obtained from a simulation experiment and from observations show that in a barotropic model a spread index based on an indicator of non-self-adjointness, as the Golden–Thompson index, is capable of detecting with good reliability initial conditions with a tendency to produce large spreads.

1. Introduction

Much interest has been raised recently on the issue of local error growth for weather forecasts. Several papers have looked at the problem of analyzing the intrinsic uncertainty caused in forecasts by the growth of small perturbations on the initial condition (Barkmeijer 1992).

The sensitivity to the initial condition has been studied classically by the analysis of the normal modes of the linearized system, but recently it has been pointed out that the normal-modes approach, though very useful as a conceptual and theoretical tool (Simmons et al. 1983), can suffer severe problems when it is applied to a realistic turbulent flow (Lacarra and Talagrand 1988; Farrell 1990; Branstator 1992). These investigations have disclosed that finite time transient growth can be present even in cases when traditional normal-mode analysis would have given a stable flow.

Since the pioneering work of Lorenz (1965), it has been known that the growth of intrinsic instabilities of the initial conditions is a major component of the forecast error. One of the techniques that has been proposed to try to determine the growth of instabilities has been the ensemble forecast. In this approach, conceptually similar to a Monte Carlo technique (Hoffman and Kalnay 1983; Kalnay and Dalcher 1987; Dalcher et al. 1988), a forecast for a nominal date is really composed of a bundle of slightly different forecasts, each perturbed, in various ways, from the nominal initial condition. The argument is that the effect of the perturbations would compensate statistically, and the ensemble would yield more stable information than any particular member. One could visualize the nominal initial condition as surrounded by a roughly spherical cloud of perturbed initial conditions. The time evolution would stretch and distort the sphere in several directions.

The distortion can be measured in several ways. The diameter of the cloud, measured as a root-mean-square deviation, has been a popular choice as an indicator for the average tendency of the forecasts to diverge from each other. Large divergence rates seem to be somewhat correlated with less skillful forecasts. Having the expected divergence corresponding to a nominal initial condition would then be a valuable step toward the goal of attaching confidence limits to the forecast.

Because the initial cloud is distorted along preferred directions in the early stages of the forecasts, several investigators (Mureau et al. 1993; Molteni and Palmer 1992; Toth and Kalnay 1993) have tried to obtain the information on the most sensitive directions in phase space for their ensemble forecast scheme in an attempt to identify a priori a "bad" forecast, that is, a forecast with a stronger than normal sensitivity to initial conditions. In all cases attention has been on the perturbations that would result in a very large amplification at a later time, in an appropriate norm. The hope was that few directions would dominate the growth of the

Corresponding author address: A. Navarra, IMGA-CNR, Via Emilia Est 770, 41100 Modena, Italy.

uncertainty so that it would be sufficient to control the largest, most amplifying, directions to obtain reliable information on the growth of the cloud of uncertainty.

Linear analysis of the initial condition, however, seems to indicate that the situation is more complicated. Linear analysis, either with traditional normal modes or using optimal modes (Farrell 1990), shows that the spectrum of the initial condition is almost never dominated by a small number of modes. The most common situation is that an almost continuous spectrum exists, with a large number of modes contributing to large amplification rates. It is rather arbitrary, then, to pick a few modes on the basis that with the particular numerics chosen, they happen to have large amplification rates. Normal-mode analysis has been found to be very sensitive to the details of the numerical formulation (Anderson 1991). A real perturbation will probably project with equal probability on a large number of modes, and therefore it is the total spectral distribution that is more significant for distinguishing "good" initial conditions from "bad" ones. In this respect it seems that a better measure would be provided by the total variance and its evolution. The time evolution of the variance would represent the time evolution of a cloud of initial states, and it would eliminate the ambiguity of selecting some particular mode by considering all of them.

We propose in this paper a procedure based on the Schmidt decomposition (Navarra 1993; Golub and Van Loan 1989) for studying the variance evolution. In section 2 we discuss the barotropic propagator and how the variance can be expressed in terms of Schmidt modes. A comparison with a set of nonlinear calculations in section 3 will examine how the linear result is applicable to the nonlinear case. In section 4, we illustrate a method for estimating the variance at later times from the initial condition itself, without the need for computing the evolution operator, based on the generalized Golden–Thompson inequality (Bernstein 1988). A set of experiments based on daily initial conditions will be discussed in section 5. Some discussion and conclusions will follow in sections 6 and 7.

2. The barotropic propagator

a. Definition

The work in this paper is based on the nondivergent barotropic equation on a sphere. Though recently attention has been given to the importance of the effect of the divergent part of the basic state (Borges and Hartmann 1992; Hoskins and Sardeshmukh 1987) we feel that for the purposes of this paper it is sufficient to remain in a nondivergent context. The equations are solved using a spectral technique (Bourke 1972) that has been described elsewhere (Navarra 1993), but we are briefly summarizing it for completeness.

The model equation is based on the nondivergent barotropic equation for a two-dimensional flow on a rotating sphere, namely,

$$\frac{\partial \zeta}{\partial t} + \mathbf{v}\cdot\nabla(\zeta + f) = 0, \tag{1}$$

where ζ is the vertical component of the relative vorticity; ∇ is the horizontal gradient operator in spherical coordinates; $\mathbf{v} = (u, v)$ is the horizontal velocity vector; and f is the Coriolis parameter defined as $2\Omega \sin\phi$, where Ω is the earth's angular velocity and ϕ is latitude. In the following, no distinction will be attempted between the operator and its numerical representation. The terms operator and matrix will be taken as denoting the same object.

The vorticity and the velocity field are split into a basic state (ζ_b, $\mathbf{v}_b$) and an anomaly part (ζ', $\mathbf{v}'$). Inserting the split fields into Eq. (1) and dropping primes yields the linearized anomaly equation

$$\frac{\partial \zeta}{\partial t} = \nabla\cdot(\mathbf{v}\zeta_b + \mathbf{v}_b\zeta + f\mathbf{v}). \tag{2}$$

With the help of the diagnostic relations

$$\zeta = \nabla^2\psi$$

$$u = -\frac{\partial \psi}{\partial y}$$

$$v = \frac{\partial \psi}{\partial x},$$

the state of the system can simply be expressed in terms of the scalar field ζ.

The equations are solved in spherical coordinates using an expansion in spherical harmonics (Bourke 1972) with a rhomboidal truncation at wavenumber 15 and 64-bit arithmetic. A real vector can be introduced containing the real and imaginary parts of the complex spherical harmonics coefficients of the vorticity ζ_n^m:

$$\mathbf{x} = (\mathrm{Re}(\zeta_0^2), \cdots, \mathrm{Re}(\zeta_N^M); \mathrm{Im}(\zeta_0^1), \cdots, \mathrm{Im}(\zeta_N^M)).$$

The dimension of the $\mathbf{x}$ vector for the resolution used in this paper is 494 and corresponds to the spectral degrees of freedom of the problem minus the number of components conserved in the absence of external forcing.

Using the compact vector notation, the anomaly equation becomes

$$\dot{\mathbf{x}} = \mathbf{A}_{\bar{\mathbf{x}}}\mathbf{x}, \tag{3}$$

where the matrix $\mathbf{A}_{\bar{\mathbf{x}}}$, which depends on the basic-state representation $\bar{\mathbf{x}}$, is the numerical representation of the linear operator in Eq. (2), and the vector definition of the anomaly is used as follows:

$$\mathbf{x}' \equiv \mathbf{x} - \bar{\mathbf{x}}.$$

The solutions of Eq. (2) can also be written as

$$\mathbf{x}(T) = \mathbf{S}(0, T)\mathbf{x}(0), \tag{4}$$

where $\mathbf{x}(0)$ is the representation of the initial state and $\mathbf{S}(0, T)$ is the linear operator that propagates the state vector from $t = 0$ to $t = T$. Though the same "resolvent" has been used previously in the literature (Farrell 1990; Lacarra and Talagrand 1988; Lorenz 1965) we would like to propose the name "propagator" in analogy with other physical applications where operators of this form appear. The **S** operator propagates the state vectors from one time level to another, and so it seems that the new name describes its special role more accurately.

The calculation of the propagator is therefore mathematically equivalent to the evaluation of the exponential of some matrix, which is not a trivial problem for general matrices with no particular structure. Moler and Van Loan (1978) and Van Loan (1979) have discussed in detail the pitfalls of several popular methods for computing exponentials of matrices. They concluded that when possible, the direct calculation of the column of **S** via numerical integration of the differential equation (2) or (3) is the most accurate method.

The technique is straightforward and is essentially an expansion of the Hoskins and Karoly (1981) approach to nonstationary problems. Equation (4) implies that the application of **S** to the unitary vector $\mathbf{e}_1 = (1, 0, 0, \cdots, 0, 0)$ will result in the first column of **S**. Repeated application of **S** to the unitary vectors will yield all the columns in succession. In practice, one has to integrate the spectral model [Eq. (2) or Eq. (3)] to time T with the unitary initial condition to obtain the numerical representation of **S**. For instance, to obtain the one-day propagator, it is necessary to perform 494 one-day integrations using as initial conditions the unitary vectors.

b. Ensembles and variance

The initial ensemble can be seen as the superposition of a set of perturbations $\{\mathbf{x}(0)_i\}_{i=1}^m (\mathbf{x}(0)_i = (x_i^1, \cdots, x_i^n))$ and the initial condition $\mathbf{z}$ to generate the ensemble initial condition $\{\mathbf{z} + \mathbf{x}(0)_i\}_{i=1}^m$. The perturbations will be chosen to have zero ensemble mean $\langle \mathbf{x}(0) \rangle = 0$, to avoid an obvious drift of the centroid of the ensemble.

The perturbations can be represented as a matrix:

$$\mathbf{X}_0 \equiv [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m],$$

with a perturbation covariance matrix given by

$$\mathbf{C}_0 \equiv \mathbf{X}_0\mathbf{X}_0^{\mathrm{T}},$$

where the T denotes the transpose. The total perturbation variance is then given by

$$\mathrm{Tr}(\mathbf{C}_0) = \sum_{j=1}^{n} \sum_{i=1}^{m} (x_i^j - \langle x \rangle^j)^2,$$

where x_i^j is the j component of the vector $\mathbf{x}_i$. The empirical orthogonal function (EOF) analysis of the statistical ensemble is just the eigenvalue analysis of the covariance matrix. The total variance of the perturbations is then given by the sum of the eigenvalues λ_i of $\mathbf{C}_0$ and the ratio of one eigenvalue λ_k to the sum of the eigenvalues; that is, the ratio,

$$\frac{\lambda_k}{\sum_{i=1}^{n} \lambda_i},$$

is known as the variance explained by the pattern k. In simple terms it represents the probability that pattern k would appear in a random extraction from the ensemble of the anomalies. More generally we can define the covariance matrix at all times as $\mathbf{C}_t = \mathbf{X}_t\mathbf{X}_t^{\mathrm{T}}$, where $\mathbf{X}_t$ is the matrix of the perturbations at time t.

The covariance matrix $\mathbf{C}_t$ contains the information on the divergence of the forecasts. Computing $\mathbf{C}_t$ is in general a demanding task, consisting in practice of performing as many forecasts as the number of columns of the covariance matrix, but some insight can be gained by considering the case of a linear system.

For a linear system whose propagator to time t is $\mathbf{S}_t$, the covariance matrix at time t of an ensemble of initial perturbations (columns of the matrix $\mathbf{X}_0$) can be easily expressed in terms of the propagator:

$$\mathbf{C}_t = \mathbf{X}_t\mathbf{X}_t^{\mathrm{T}} = \mathbf{S}_t\mathbf{X}_0\mathbf{X}_0^{\mathrm{T}}\mathbf{S}_t^{\mathrm{T}} = \mathbf{S}_t\mathbf{C}_0\mathbf{S}_t^{\mathrm{T}}.$$

If $\mathbf{C}_0 = 1$ (i.e., if $\mathbf{X}_0$ is orthogonal or, in other words, if we have perfect statistical independence of the perturbations) then we can write

$$\mathbf{C}_t = \mathbf{S}_t\mathbf{S}_t^{\mathrm{T}},$$

and the ratio of the linear variance at time t with respect to the initial variance is given by

$$R_L(t) = \frac{\mathrm{Tr}(\mathbf{S}_t\mathbf{S}_t^{\mathrm{T}})}{\mathrm{Tr}(\mathbf{C}_0)}.$$

We can use the Schmidt decomposition described in Navarra (1993) to decompose the propagator at time t, $\mathbf{S}_t = \mathbf{U}_t\Sigma_t\mathbf{V}_t^{\mathrm{T}}$, where $\Sigma_t = \mathrm{diag}(\sigma_1, \cdots, \sigma_n)$ is the diagonal matrix of the Schmidt numbers, denoted by σ_i, and $\mathbf{U}_t$, $\mathbf{V}_t$ are the orthogonal matrices of the Schmidt vectors. Inserting the decomposition in the expression for the covariance matrix at time t we get

$$\mathbf{S}\mathbf{S}^{\mathrm{T}} = \mathbf{U}_t\Sigma_t\mathbf{V}_t^{\mathrm{T}}\mathbf{V}_t\Sigma_t\mathbf{U}_t^{\mathrm{T}} = \mathbf{U}_t\Sigma_t^2\mathbf{U}_t^{\mathrm{T}}, \tag{5}$$

so that the ratio of the linear variance is given by

$$R_L(t) = \frac{\sum_{i=1}^{n} \sigma_i^2}{m}. \tag{6}$$

The relation (5) shows that the singular values are identical to the eigenvalues of the covariance matrix, and the Schmidt vectors **U** are the optimal modes of

Farrell (1989) and Lorenz (1965). It is possible then to attempt a different interpretation of the optimal modes based on Eqs. (5) and (6). The optimal modes (or u–Schmidt modes) are the EOF patterns of an ensemble obtained from a random perturbation of the basic state. A Schmidt number σ_i represents an amplification factor for the mode i; the ratio

$$\frac{\sigma_k^2}{\sum_{i=1}^{n} \sigma_i^2}$$

expresses the variance explained by the kth optimal mode.

c. *Probabilistic interpretation of the Schmidt modes*

The relation found between the covariance matrix and the Schmidt modes is suggestive of a possible interpretation of the Schmidt modes in terms of the probability distribution. Assuming Gaussian statistics, the probability distribution corresponding to an ensemble with covariance matrix $\mathbf{X}\mathbf{X}^{\mathrm{T}}$ is written as (Lorenc 1986; Gardiner 1983)

$$P(\mathbf{x}) = P_N \exp\left(-\frac{1}{2}\mathbf{x}^{\mathrm{T}}(\mathbf{X}\mathbf{X}^{\mathrm{T}})^{-1}\mathbf{x}\right),$$

where P_N is a normalization constant. Using the Schmidt-mode decomposition of the covariance matrix (5) in the preceding equation we obtain

$$P(\mathbf{x}) = P_N \exp\left(-\frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{U}\boldsymbol{\Sigma}^{-2}\mathbf{U}^{\mathrm{T}}\mathbf{x}\right). \quad (7)$$

The most probable state is therefore found as the state that satisfies

$$\max(-\ln P(\mathbf{x})) = \min\left(-\frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{U}\boldsymbol{\Sigma}^{-2}\mathbf{U}^{\mathrm{T}}\mathbf{x}\right). \quad (8)$$

We can exploit the orthogonality property of the Schmidt modes to perform a coordinate transformation $\mathbf{y} = U^{\mathrm{T}}\mathbf{x}$; then Eq. (8) becomes simply

$$\max(-\ln P(\mathbf{y})) = \min\left(-\frac{1}{2}\mathbf{y}^{\mathrm{T}}\boldsymbol{\Sigma}^{-2}\mathbf{y}\right),$$

and the most probable state is then $y = (1, 0, \cdots, 0)$, that is, the first Schmidt vector.

A physical interpretation of the ith Schmidt mode is then possible as the mode that tends to emerge over the time period covered by the propagator, with a probability given by $\exp(2\sigma_i^2)^{-1}$. Strictly speaking, to verify this relation numerically, it would be necessary to perform a multidimensional histogram of the ensemble-phase space variables. This is a rather demanding task, but an indirect verification can be found. In a broad sense, the probability expresses the fact that the members of the ensembles tend to project preferentially on the high-probability states rather than on others. This notion can be made more precise by using the probability distribution to calculate the expectation value of the square of the projection on the ith Schmidt mode; that is,

$$E[(\mathbf{x}, \mathbf{u}_i)^2] = \int (\mathbf{x}, \mathbf{u}_i)^2 P(\mathbf{x}) d\mathbf{x}.$$

Substituting the expression for the probability and changing variables as before ($\mathbf{y} = U^{\mathrm{T}}\mathbf{x}$), we obtain

$$E[(\mathbf{x}, \mathbf{u}_i)^2] = P_N \int (U\mathbf{y}, \mathbf{u}_i)^2 \exp\left(-\frac{1}{2}\mathbf{y}^{\mathrm{T}}\boldsymbol{\Sigma}^{-2}\mathbf{y}\right) d\mathbf{y}$$

$$= P_N \int y_i \exp\left(-\frac{1}{2}\mathbf{y}^{\mathrm{T}}\boldsymbol{\Sigma}^{-2}\mathbf{y}\right) d\mathbf{y},$$

where we have used the orthogonality of $\mathbf{U}$ to show that the scalar product is simply y_i, the ith component of $\mathbf{y}$. The integral can then be done by parts to show that

$$E[(\mathbf{x}, \mathbf{u}_i)^2] \propto \sigma_i^2. \quad (9)$$

Equation (9) indicates that state vectors will tend to project on the Schmidt modes proportionally to the size of the square of the Schmidt number σ_i. The most amplifying mode will be favored in the short-term evolution of the system. In the following section this equation will be tested with a set of numerical experiments.

3. Time-dependent calculations

We will test the above ideas on an ensemble of 200-mb streamfunction monthly means taken from an integration of the CCM0 (Community Climate Model, version 0) kindly made available by G. Branstator. The purpose of this ensemble is to test the feasibility of utilizing the variance as an indicator for the initial condition spread under varying conditions but in a simpler context than the utilization of analyzed initial conditions would entail. The model has been kept as simple as possible, with no external forcing of any kind, including the artificial forcing often used to make sure that the barotropic model has the same climatology as the GCM.

The standard deviation of the monthly mean ensemble is shown in Fig. 1a (Fig. 1b will be discussed later). Distinctive maxima are visible, suggesting that the ensemble has enough variability to be able to provide a good test case.

a. *Nonlinear integrations*

For each monthly mean initial condition an ensemble of initial value nonlinear calculations was made. The barotropic model was left free to evolve without the constraint of it having the same climatology as the GCM. The dimension of the ensemble was fixed arbi-

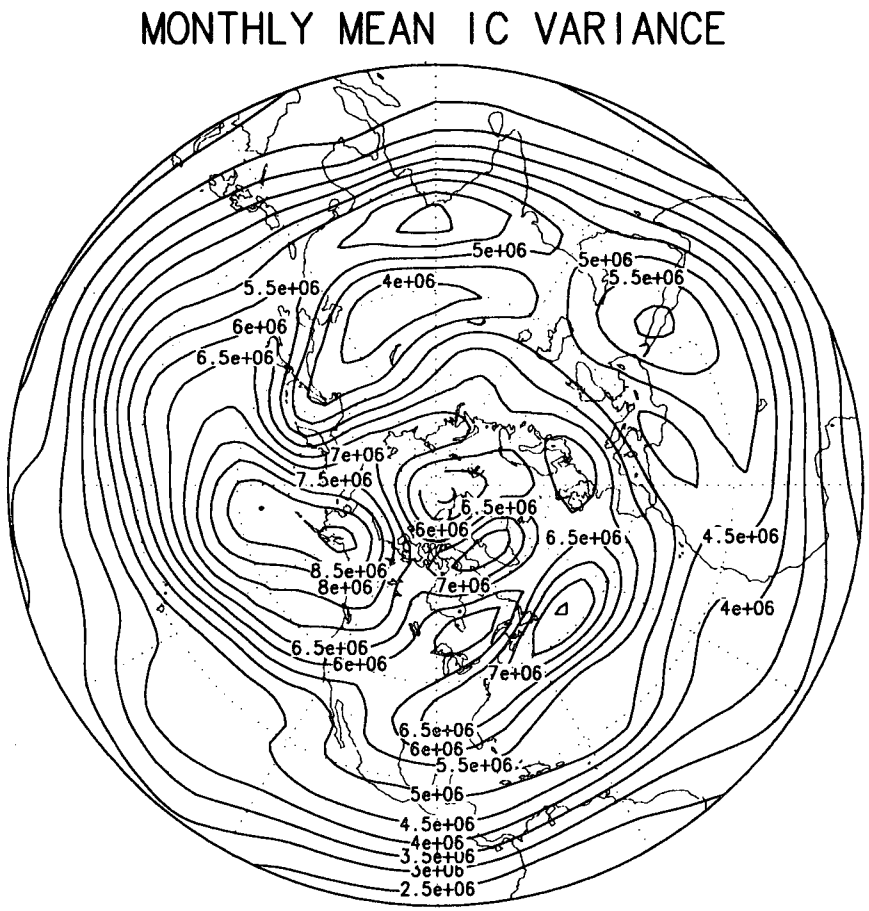


FIG. 1a. Standard deviation of the streamfunction of the CCM ensemble. The contour interval is $5 \times 10\ m^2\ s^{-1}$.

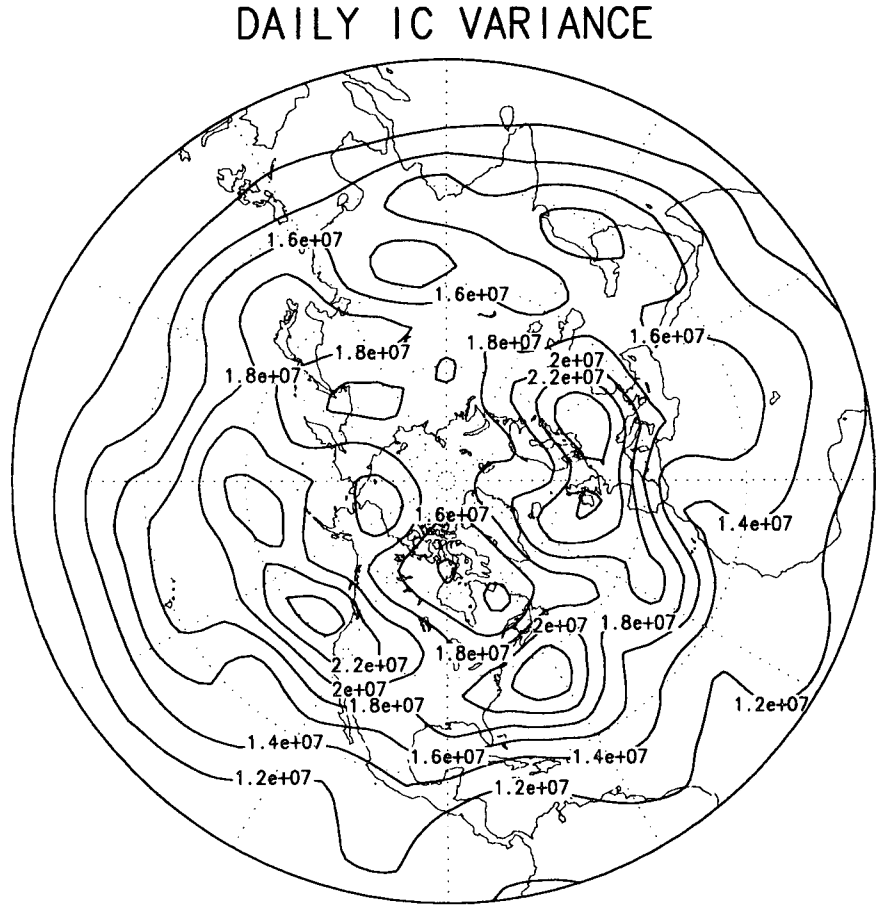


FIG. 1b. Standard deviation of the streamfunction of the initial conditions for January 1987, 1988, and 1989 from ECMWF analyses. The contour interval is $2 \times 10^6\ m\ s^{-1}$.

trarily at 494. The nonlinear calculation was performed essentially to provide a reference variance to compare with the linear variance. The nonlinear ensemble was obtained by perturbing each initial condition with an anomaly corresponding to a perturbation kinetic energy of 20% of the basic state and integrating forward to a specified time (7 days). Ideally, the perturbations must be chosen in such a way so as to sample the space around the initial condition in the most complete way, with the only constraint that the ensemble must have unit variance. Two kinds of perturbations were used: (i) random perturbations with a flat spectrum, and (ii) random orthogonal perturbations. The results show that better results are obtained by using the orthogonal set, and in the following, only results from this method will be shown.

In practice, for each of the 200 initial conditions, 494 seven-day integrations were performed, and an EOF analysis was performed at each day. In this way we can monitor, via the EOF analysis, the deformation of the initial hypersphere into an ellipsoid as the system evolves.

The total nonlinear variance at day 1 as a function of the initial conditions is shown in the bottom panel of Fig. 2. The large oscillation seems to indicate a large variability in the growth of perturbations. Some initial conditions generate more variance and others are more quiescent.

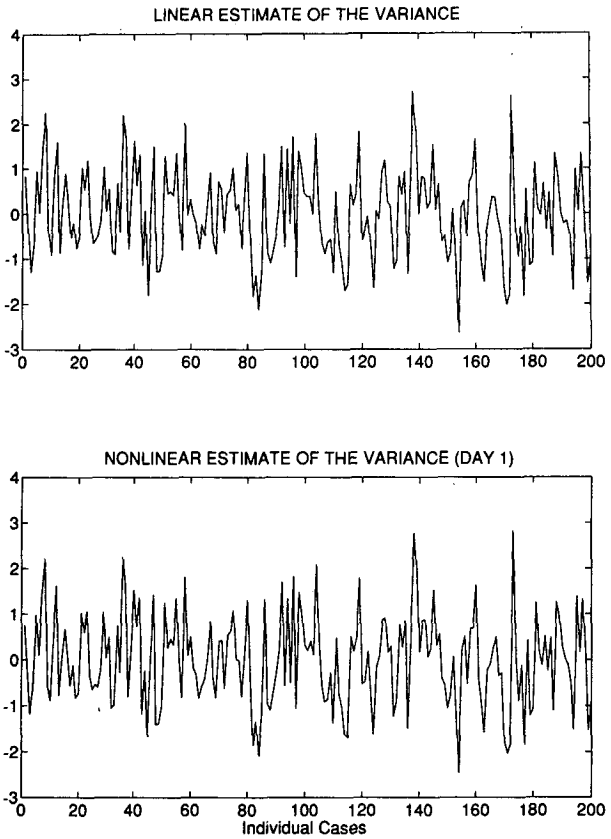


FIG. 2. Total variance of the nonlinear reference set (top), total linear variance from the Schmidt decomposition of the propagator (bottom) at day 1, divided by the initial variance. The ensemble has been normalized by removing the mean and dividing by the standard deviation.

b. Linear integrations

In the linear case, the initial condition was used as a basic state in the linearization (2), and the propagator was computed. A Schmidt decomposition was then performed to obtain the optimal perturbations and the total linear variance.

The results are shown in Fig. 2 (top panel). There is an excellent agreement with the nonlinear variance (bottom panel), and the correlation between the curves is about 0.95. The reliability of the correlation was estimated by computing a shifted correlation coefficient (the correlation coefficient between the two ensembles shifted j positions) and checking that the maximum is achieved for zero shift and that its value is well above the shifted correlations. The shifted correlation coefficient between the ensembles $[f(i), g(i)]$ is defined as

$$c(j) \equiv \sum_{k=0}^{m-1} f(k)g(k+j)$$

for a shift j. The value $c(0)$ is the traditional correlation index; the relative height of the central spike with respect to the "noise" gives a measure of the probability of having a high correlation by chance. In the following, all correlations are understood to have passed such a test.

The absolute values of the variance are difficult to interpret, but in order to have a feeling for their meaning, we have plotted in Fig. 3 the composites of the five states with the largest and smallest variance. An interesting anomaly develops, with well-defined positive and negative centers, arching from the North Pacific to the Atlantic. This picture suggests that large variance states may be connected to a higher than normal amplitude of the stationary waves, whereas small variance states may be connected to rather zonal situations. This picture is only suggestive because the pattern seems to project highly on the dominant EOF of the variability of the IC set.

4. Linear estimators of the variance evolution

In the previous section we have shown that there is some promise for using total variance to select initial conditions that are potentially active, and there are some hints that higher than normal ridges in the North Pacific can be related to high variance. The calculation of the propagator is, however, an expensive calculation, especially when considering a practical application to a forecasting model. It would be desirable to be able to

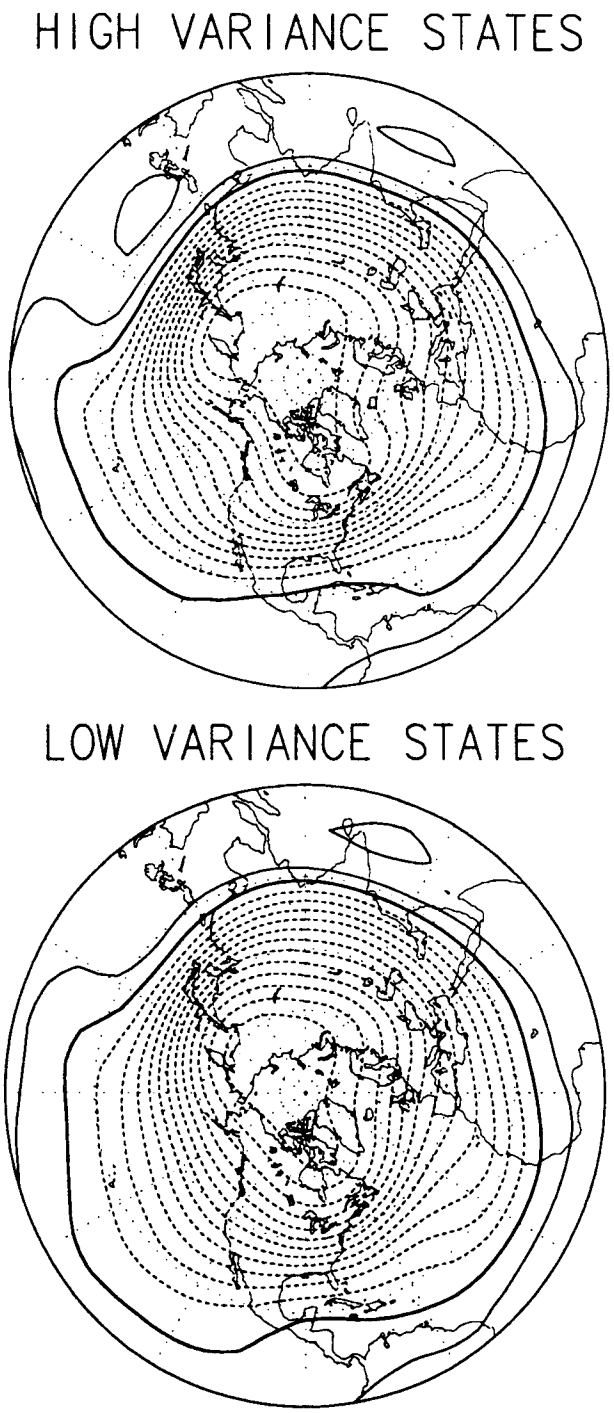


FIG. 3. Northern Hemisphere streamfunction composites of the five initial conditions for the monthly mean case with largest variance and of the five initial conditions with smallest variance. The contour value is irrelevant.

deduce some property of the propagator directly from the initial condition itself. In more precise terms, we would like to deduce the expected variance from the equation linearized around the initial condition (**A**) rather than from the propagator ($e^{\mathbf{A}t}$), because of an obvious computational advantage. We will examine in this section various methods of characterizing the sensitivity to perturbations of the initial condition. The methods will be based on the linearization of the model equation about the nominal initial condition.

Normal-mode analysis has been used by Palmer (1988) to show that states with a large projection on the positive Pacific/North American (PNA) index are more unstable than states with a negative index. A conjecture was then made that highly unstable basic states could lead to a higher than normal divergence of an ensemble forecast. We have tested this idea by calculating the normal-mode analysis of the initial condition ensemble used in this paper. The largest growth rates are plotted in Fig. 4 (top panel). There is little correlation with the reference divergence (Fig. 2, top), and the correlation coefficient between the two curves (Table 1) is not high (0.38). The situation improves if we use the largest amplifying optimal mode at day 1 (Fig. 4, bottom)—the agreement is better (0.63).

Some insight into the dynamical reasons of this behavior can be obtained in Fig. 5 (second panel from the top). In this picture the departure from normality is plotted for the initial condition ensemble. The de-

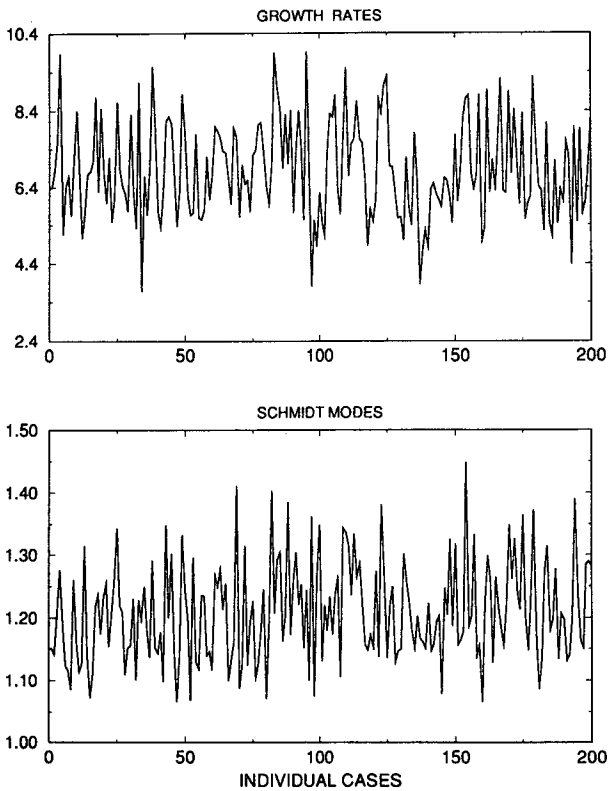


FIG. 4. Normal-mode growth rates (top) in days and optimal mode amplification factors (bottom) per day at day 1.

TABLE 1. Correlation coefficient between several indicators of divergence for the 200 cases in which G–T: Golden–Thompson index, MAXEIG: largest real part of the eigenvalues of **A**, DEP: departure from normality (Navarra 1993); LV: linear variance estimated from the Schmidt decomposition, LOGNORM: logarithmic norm of **A**, MAXSCH: maximum Schmidt number of the propagator, NV: reference nonlinear variance obtained from the Monte Carlo experiments.

Indicator	G–T	MAXEIG	DEP	LV	LOG-NORM	MAXSCH	NV
G–T	1.00	0.39	0.90	0.78	0.93	0.52	0.77
MAXEIG		1.00	0.36	0.37	0.42	0.27	0.38
DEP			1.00	0.77	0.88	0.48	0.75
LV				1.00	0.77	0.57	0.97
LOGNORM					1.00	0.58	0.78
MAXSCH						1.00	0.63
NV							1.00

parture from normality can be defined as (Golub and Van Loan 1989; Navarra 1993)

$$d_p = \sum_{i=1}^{n} \sigma_i^2 - \sum_{i=1}^{n} |\lambda_i|^2,$$

where λ are the eigenvalues and σ the Schmidt numbers. For self-adjoint or normal matrices the departure is zero, since in this case the Schmidt numbers and the eigenvalues coincide, and therefore the optimal modes and the eigenvectors coincide forming an orthonormal set. A highly non-self-adjoint matrix will have a large departure and eigenvectors that form arbitrary angles; a weakly non-self-adjoint matrix will have a small departure and quasiorthogonal eigenvectors. Thus, this quantity measures how much the linearization matrix is not self-adjoint or, in other words, how much the eigenvectors are oblique to each other. It appears that there is a qualitative agreement between highly non-self-adjoint states and high variance states, indicating that strongly oblique eigenvectors are favorable to a large divergence of the ensemble.

The encouraging result of Fig. 5 prompts us to look for an integral indicator that takes into account the entire spectrum of modes, rather than a few selected ones. The good agreement between the top and bottom panels of Fig. 2 may encourage us to adopt a linear estimation of the variance as long as the entire spectrum is used. Unfortunately, the calculation of the linear variance

$$\mathrm{Tr}(\mathbf{S}_t\mathbf{S}_t^{\mathrm{T}}) = \mathrm{Tr}[e^{\mathbf{A}t}e^{\mathbf{A}^{\mathrm{T}}t}]$$

is still expensive, but a recent result by Bernstein (1988) may come to the rescue. The basic result is that the inequality, a modified form of the Golden–Thompson inequality,

$$\mathrm{Tr}(\mathbf{S}_t\mathbf{S}_t^{\mathrm{T}}) = \mathrm{Tr}[e^{\mathbf{A}t}e^{\mathbf{A}^{\mathrm{T}}t}] \leq \mathrm{Tr}[e^{(\mathbf{A}+\mathbf{A}^{\mathrm{T}})t}],$$

shifts the difficulty from the calculation of the exponential of a non-self-adjoint matrix ($\mathbf{A}t$) to a self-adjoint one ($(\mathbf{A} + \mathbf{A}^{\mathrm{T}})/2$), a more secure numerical problem. Because of the self-adjointness, the trace of the exponential can be computed from the eigenvalue spectrum of $\mathbf{A} + \mathbf{A}^{\mathrm{T}}$. The results (Fig. 5, third panel from the top) show that the Golden–Thompson index (hereafter referred to as the G–T index) gives a good agreement with the reference variance, as it is confirmed by the correlation coefficient (Table 1) that gives the higher value for the G–T index. Here and in the following the G–T index is defined by the right-hand side of Eq. (9) computed at 12 hours.

The form of the modified Golden–Thompson inequality suggests that it may be sufficient to calculate the contribution to the trace deriving from the largest eigenvalues of $[e^{(\mathbf{A}+\mathbf{A}^{\mathrm{T}})t}]$. Because $(\mathbf{A} + \mathbf{A}^{\mathrm{T}})$ is a symmetric matrix, the largest eigenvalue is identical to its largest singular value, and therefore it coincides with its norm. This particular norm is also known as the logarithmic norm of the matrix **A**. The bottom panel of Fig. 5 shows the logarithmic norm for the ensemble of initial conditions considered in this paper. It correlates very well with the nonlinear variance and with the G–

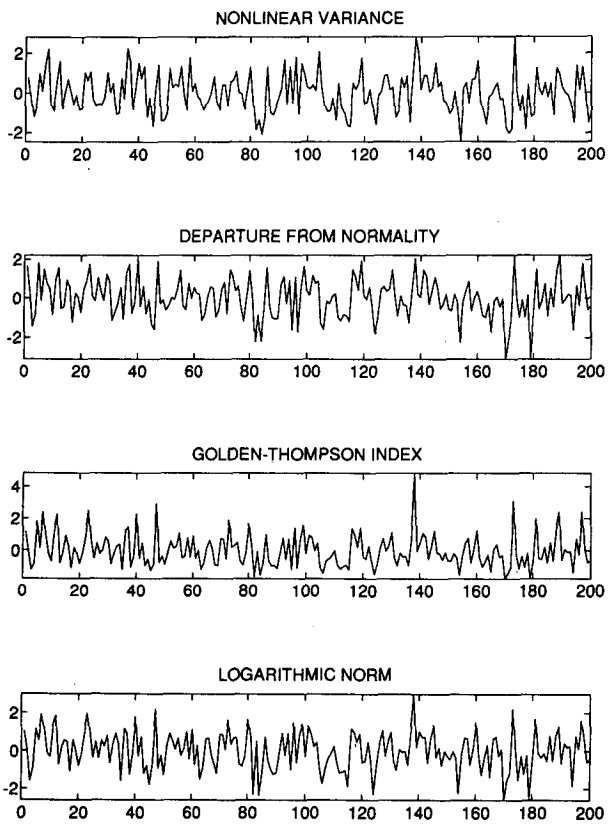


FIG. 5. Linear estimators of the variance at day 1 for the initial condition ensemble. At the top: the nonlinear variance is reproduced for reference; second panel from the top: departure from normality, defined as in Navarra (1993); third panel from the top: the Golden–Thompson index; and in the bottom panel: the logarithmic norm. All figures have been normalized as in Fig. 2.

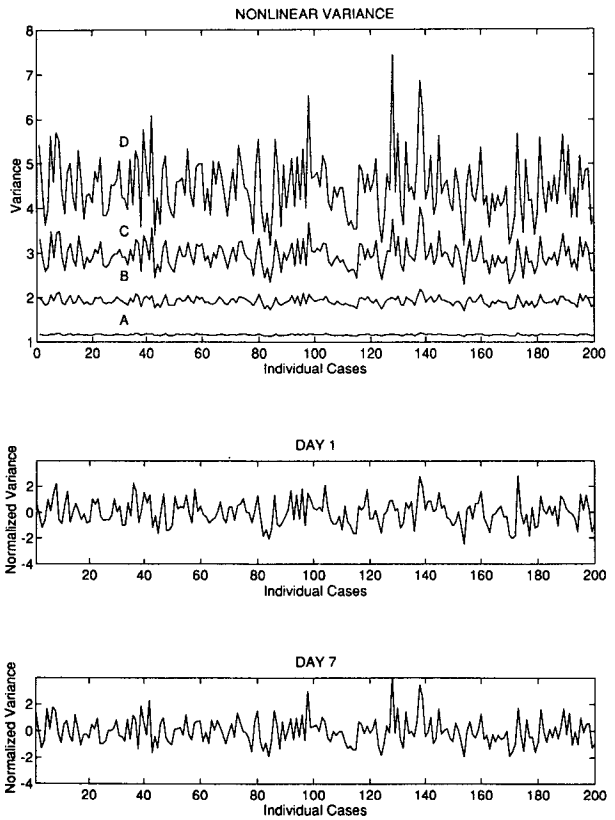


FIG. 6. Evolution of the nonlinear variance in the Monte Carlo experiment. The ratio of the variance to the initial variance is shown in top panel. Curves A, B, C, D refer to day 1, 3, 5, and 7, respectively. In the middle and bottom panel the curves normalized as in Fig. 2 are shown.

T index, indicating that the logarithmic norm, that is, the largest eigenvalue of $(\mathbf{A} + \mathbf{A}^T)$ is the dominant contribution to the trace of the propagator.

The evolution of the variance for the following days is shown in Fig. 6. The absolute total variance increases monotonically (top panel), but significant differences exist between initial conditions. Certain cases diverge at almost twice the speed of others. However, if the mean is removed and the variances are normalized by the standard deviation, then a more consistent picture appears (middle and bottom panels). There is a substantial consistency at different days, implying that large variance initial conditions tend to diverge more rapidly than others at least until day 7. It is to be expected, of course, that in the presence of dissipation and realistic forcing terms, the situation will be altered; but at least in this simple model, nonlinear interactions alone are not sufficient to drastically change the ranking of initial conditions obtained from the variance estimation. A limited number of sensitivity experiments were performed by changing the size of the initial perturbations to 10% of the energy of the climatology, but the time consistency was not affected.

As a consequence, the linear indexes give a surprisingly good estimation of the variance for the following days. Table 2 shows the correlation of several linear indexes with the nonlinear reference variance. Good agreement with the nonlinear variance can be obtained using the G–T index or the departure from normality index, but surprisingly good results are obtained with indexes that are just based on the matrix **A**, as is the logarithmic norm or the maximum singular value. Model indexes as the most unstable mode or the most amplifying Schmidt mode are definitely giving worse results than this class of indexes, indicating that they give poor information on the following evolution of the variance.

To summarize, this section has shown that in an inviscid, unforced, barotropic model the total variance evolves in an orderly way. Initial conditions that give rise to large variance at day 7 can be identified already at day 1, and similarly initial conditions that have small variance at day 7 can be already singled out at day 1.

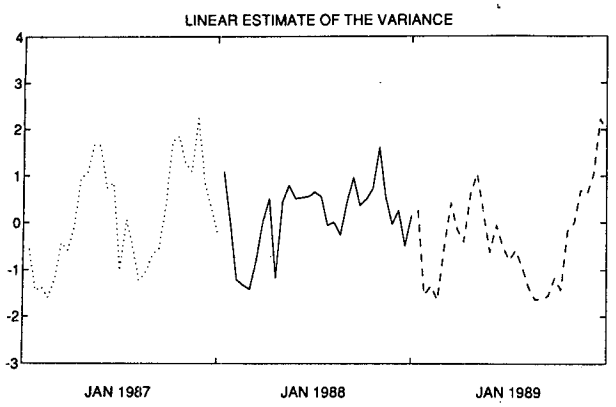
5. Experiments with daily initial conditions

The ensemble experiments performed in the previous sections were repeated with observed initial conditions to verify the robustness of the results with respect to more active initial states. The initial vorticity fields were obtained from daily operational 300-mb analysis from the European Centre for Medium-Range Weather Forecasts (ECMWF) for the months of January 1987, 1988, and 1989 and truncated to the model resolution R15. A total of 90 experiments were performed. As in the previous case a set of nonlinear integrations were performed perturbing the initial state with orthogonal perturbations. These results were compared with linear calculations and the divergence indicators described before. The standard deviation of the ensemble is shown in Fig. 1b. The daily initial conditions have more variability than the monthly means, and they should provide a more severe test for the methods.

Figure 7 (bottom) shows the nonlinear variance estimates for the daily initial conditions dataset. The nor-

TABLE 2. Correlation coefficient between several indicators of divergence for the 200 cases and the nonlinear divergence at days 1 and 7 in which G–T: Golden–Thompson index, MAXEIG: largest real part of the eigenvalues of **A**, DEP: departure from normality (Navarra 1993), LOGNORM: logarithmic norm of **A**, MAXSCH: maximum Schmidt number of the propagator, LV: linear variance from the Schmidt decomposition.

Indicator	Day 1	Day 7
LV	0.99	0.75
MAXEIG	0.38	0.36
MAXSCH	0.54	0.55
DEP	0.76	0.70
LOGNORM	0.76	0.69
G–T	0.77	0.70



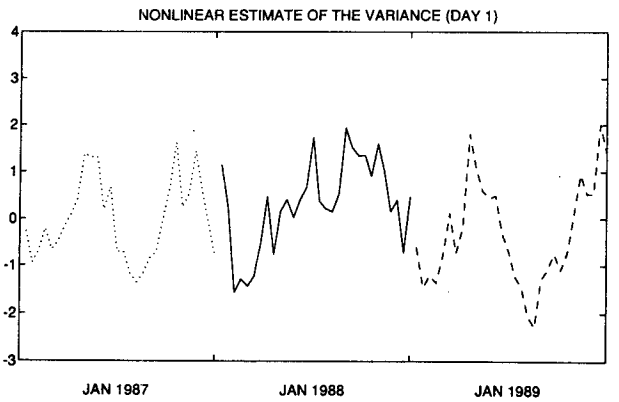


FIG. 7. As in Fig. 2 but for the daily initial condition set.

malized values show behavior similar to the monthly mean case, with fluctuations up to two standard deviations. The linear estimate of the variance evolution obtained from the linear calculation of the propagator is generally in agreement. Better results are obtained for January 1987 and 1989, with some marked disagreement in the January 1988 case. Overall, the correlation between the two curves is 0.86 (Table 3).

The evolution of the nonlinear variance for the daily initial conditions is displayed in Fig. 8. The absolute evolution of the variance for the 90 cases for day 1 to day 7 is shown in the top panel. The daily initial conditions are more active, and they generate variance that is an order of magnitude larger than in the previous monthly mean case. There is good correspondence between the curves, implying that in this simple model the initial conditions with a strong initial divergence will continue to diverge in the following evolution. The normalized variance at day 1 and 7 (middle and bottom panels) are quite similar; as can be seen in Table 3 the correlation is 0.85.

Inspection of the linear indexes for the daily initial conditions case (Fig. 9) shows that also in this case we get good correlations with the linear and nonlinear estimates of the variance at day 1. The indexes that are

TABLE 3. As in Table 1 but for the case of observed initial conditions. The Januarys of 1987, 1988, and 1989 are considered together here.

Indicator	G–T	MAXEIG	DEP	LV	LOG-NORM	MAXSCH	NV
G–T	1.00	0.31	0.94	0.53	0.88	0.38	0.64
MAXEIG		1.00	0.30	0.70	0.24	0.60	0.52
DEP			1.00	0.56	0.85	0.32	0.69
LV				1.00	0.54	0.70	0.86
LOGNORM					1.00	0.39	0.65
MAXSCH						1.00	0.54
NV							1.00

related to various measures of the non-self-adjoint character of the linearized equations around the initial condition, as the DEP, G–T, and LOGNORM indexes, have the highest correlation with the nonlinear variance (Table 3). The situation persists up to day 7 (Table 4), even if with an obvious weakening of the correlations.

6. Linear and nonlinear regimes

In order to gain some understanding of the mechanism at work in the evolution of the variance, we have

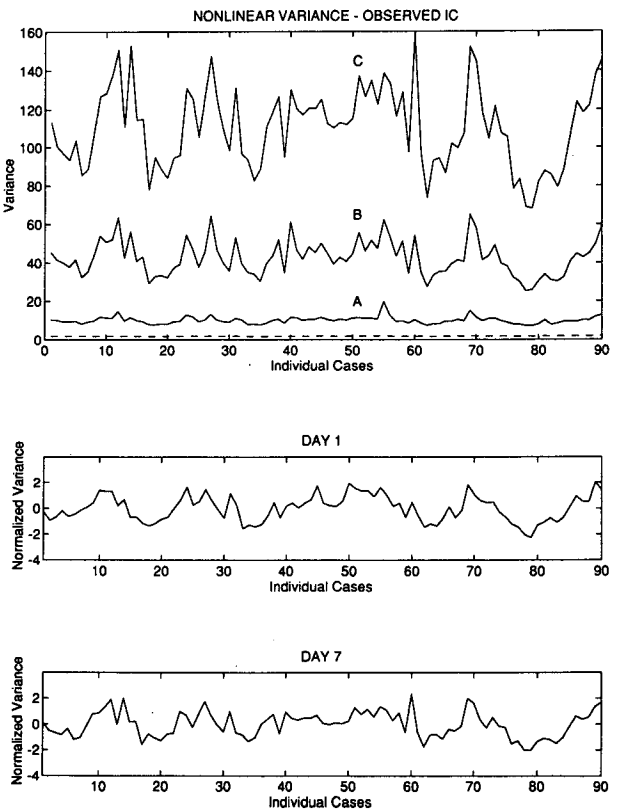


FIG. 8. As in Fig. 6 but for the daily initial condition set. In this case the dashed curve refers to day 1, whereas the curves labeled A, B, and C refer to day 3, 5, and 7, respectively.

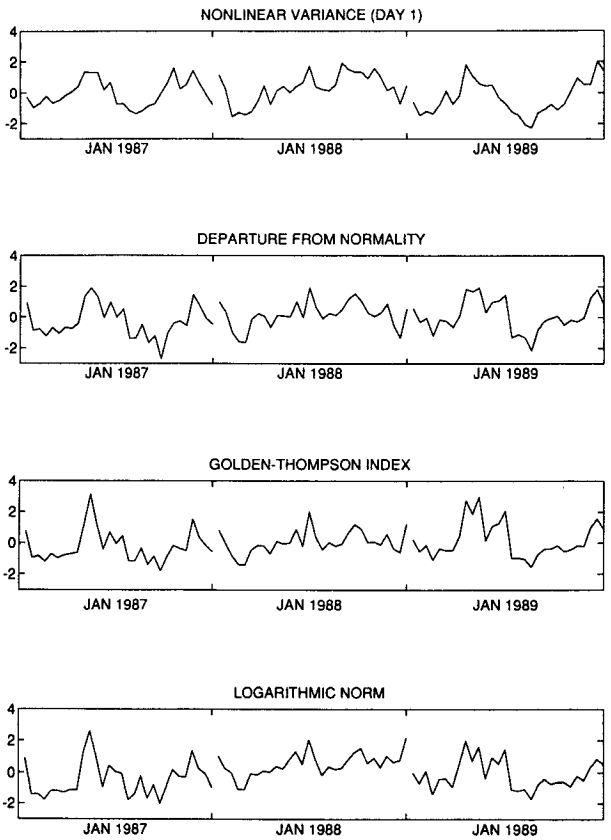


FIG. 9. As in Fig. 5 but for the daily initial condition set.

analyzed in detail the ensemble constructed around initial condition number 138 (for the monthly mean case) that gave rise to a large variance. The top panel of Fig. 10 shows the time evolution of the kinetic energy for the members of the ensemble. In this particular case, the initial energy was 10% of the basic-state energy, corresponding to two normalized units in the picture. Although a slight trend is visible, especially at day 7, the energy of the ensemble members fluctuates rather erratically. If instead optimal modes are chosen as initial conditions (in this case modes optimized for day 3 were used), then the behavior is dramatically changed; a much faster energy growth is obtained (bottom panel), and it is possible to track the ordering of the modes up to day 7. The low-order, fast-growing modes continue to dominate the evolution in the following days. More interesting is the behavior at the other end of the spectrum, that is, the strongly damped modes beyond number 400. After an initial energy decrease (presumably when the system is in the linear regime) these modes start to increase again. The overall picture seems to indicate that an upper limit for the linear regime is around day 5.

In order to sharpen this estimate and to make sure that the results obtained for the estimation of the high-

TABLE 4. As in Table 2 but for the case of observed initial conditions.

Indicator	Day 1	Day 7
LV	0.86	0.80
MAXEIG	0.52	0.47
MAXSCH	0.54	0.54
DEP	0.69	0.70
LOGNORM	0.65	0.72
G–T	0.64	0.68

variance states obtained in the preceding section refer to the nonlinear regime, we can proceed to test the linear hypothesis in the system. The expectation value for the projection on the Schmidt modes, given by Eq. (9), should exactly correspond to the linear Schmidt numbers of the propagator at different days.

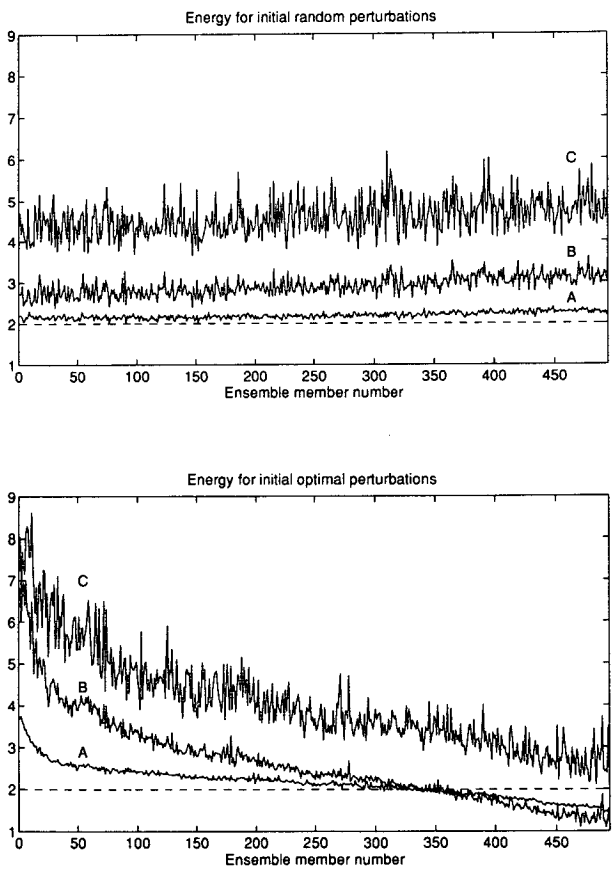


FIG. 10. Energy evolution for different kinds of perturbation sets for a particular initial condition (monthly mean case 138). (Top) The energy evolution for the members of the Monte Carlo simulation for the case in which random orthogonal perturbations were used; (bottom) Schmidt vectors optimized for day 3 are used. In both cases the initial energy of the perturbation was set to 2 arbitrary units (about 10% the energy of the basic state). The curves in the plots labeled A, B, and C refer to day 1, 3, and 7. The initial time values are indicated by the dashed lines.

Equation (9) shows that the ensemble mean of the squared projection on each Schmidt vector should be proportional to σ_i^2, the corresponding Schmidt number, if the statistics of the ensemble is given by the covariance matrix $\mathbf{SS}^T$. Figure 11 shows the ensemble mean of the projection onto the Schmidt vectors at days 1, 3, and 7 by the members of the nonlinear ensemble constructed from random perturbations of the initial condition 138. Good agreement can be noted at day 1 and some correspondence exists at day 3, but by day 7 there is no similarity indicating that we have entered a fully nonlinear regime. The linear estimation of the statistics breaks down after about day 3–4, but within this limit the Schmidt vectors give an effective representation of the most probable states.

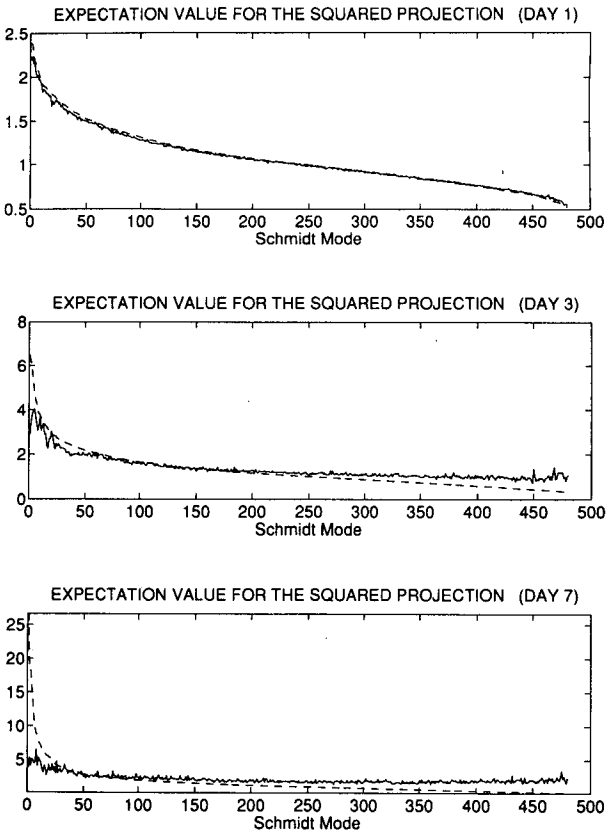


FIG. 11. Average square projection on each Schmidt mode (solid line) compared to the linear estimate (dashed line) for day 1 (top), 3 (middle), and 7 (bottom). The ensemble constructed by perturbing IC 138 is used here.

7. Conclusions

The experimental work in this paper shows that some caution must be exercised in assuming that the dominant linear mode is the key indicator of a sensitive initial condition. Among the possible indicators considered, the most unstable normal mode (largest growth rate) is weakly connected with the variance of the Monte Carlo ensemble, whereas the first Schmidt mode fares relatively better. In both cases, however, taking into account only a few modes does not provide an indicator of the Monte Carlo variance at later times. A much better result is obtained if the total linear estimate is compared with the total ensemble. It is also shown that the Schmidt decomposition of the propagator gives a link with the probability distribution, allowing an interpretation of the first Schmidt mode as the most probable mode in the linear evolution of a Monte Carlo ensemble.

The expensive calculation of the total linear variance can be avoided by resorting to the Golden–Thompson (G–T) inequality that yields an upper bound on the growth of the total variance from the linearization matrix $\mathbf{A}$ and its adjoint $\mathbf{A}^T$. Numerical results, in the context of a simple barotropic model, seem to indicate that the G–T index can identify rather accurately initial conditions that are more likely to generate large variances at later times than others. Various other indexes based on simple linear analysis of $\mathbf{A}$ or $\mathbf{A} + \mathbf{A}^T$ also seem to provide good predictors for large variance states. In general, highly non-self-adjoint linearizations lead to high variance development in the following time evolution, as the good results obtained with the departure from normality (DEP) and the logarithmic norm (LOGNORM) indicate. In practice, the DEP index is very promising because it does not require the calculation of the propagator or the availability of an adjoint model. The DEP index can be calculated entirely from the linearization of the equation around the initial condition by performing an eigenanalysis and a singular value decomposition of $\mathbf{A}$. The index can be calibrated by introducing a grading of the initial conditions based on their deviations from the ensemble-mean variance. For the monthly mean case, the average value of the DEP index for initial conditions with a positive departure greater than 1.5 standard deviations from the mean (corresponding to the most active initial conditions generating a large spread) was 1.5×10^{-5}; for the least active initial conditions departing less than the -1.5 standard deviations, the value was 1.2×10^{-5}. In the daily initial conditions case the difference between the most and the least active cases increases to 4.7×10^{-5} and 1.8×10^{-5}, respectively. It is interesting to note that the DEP index reflects the overall increase in spread from the monthly mean case to the daily case.

Though the results have been obtained with a simple model, they are encouraging enough to prompt the extension of this study to the baroclinic case. The ability to estimate the total divergence of the orbits from the initial condition alone would greatly help progress toward producing practical long-range forecasts. Further investigations are necessary, especially in connection with the problem of regime transitions. In the present study, only the overall distribution of the variance was

considered. If the evolving system will select a subset of directions preferentially, rather than expanding uniformly in all directions, a more sophisticated analysis of the variance will be needed. The preferred directions will show up as bundles of trajectories, or clusters. Each bundle will then correspond to a specific dynamical regime, namely, a region of phase space that the system will visit with higher probability. The present analysis cannot distinguish between initial conditions that generate high variance that is uniformly spread from that which is distributed in two or more separate and distant centers. Investigation of this situation is the subject of an ongoing study.

Acknowledgments. We would like to thank G. Branstator from NCAR for kindly providing us with the initial condition dataset. The work in this paper was supported by the EEC contract "Climate of the XXI Century," within the EPOCH program, and by the EEC Contract EV5V-CT92-0125.

REFERENCES

Anderson, J. L., 1991: The robustness of barotropic unstable modes in a zonally varying atmosphere. *J. Atmos. Sci.,* **48,** 2393–2410.

Barkmeijer, J., 1992: Local error growth in a barotropic model. *Tellus,* **44A,** 314–323.

Borges, M. D., and D. L. Hartmann, 1992: Barotropic instability and optimal perturbations of observed nonzonal flows. *J. Atmos. Sci.,* **49,** 335–354.

Bourke, W., 1972: An efficient, one level primitive equation spectral model. *Mon. Wea. Rev.,* **100,** 683–689.

Bernstein, D. S., 1988: Inequalities for the trace of matrix exponential. *SIAM J. Matrix Anal. Appl.,* **9,** 156–158.

Branstator, G. W., 1992: Barotropic instability and optimal perturbations of observed nonzonal flows. *J. Atmos. Sci.,* **49,** 335–354.

Dalcher, A., E. Kalnay, and R. Hoffman, 1988: Medium-range lagged average forecasts. *Mon. Wea. Rev.,* **116,** 402–416.

Farrell, B. F., 1989: Optimal excitation of baroclinic waves. *J. Atmos. Sci.,* **46,** 1193–1206.

——, 1990: Small error dynamics and the predictability of atmospheric flows. *J. Atmos. Sci.,* **47,** 2409–2416.

Gardiner, C. W., 1983: *Handbook of Statistical Methods.* Springer Verlag, 441 pp.

Golub, G. H., and C. F. Van Loan, 1989: *Matrix Computation.* Johns Hopkins University Press, 642 pp.

Hoffmann, R. N., and E. Kalnay, 1983: Lagged average forecasting an alternative to Monte Carlo forecasting. *Tellus,* **35A,** 100–118.

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 1985–1986. *Quart. J. Roy. Meteor. Soc.,* **113,** 759–778.

Kalnay, E., and A. Dalcher, 1987: Forecasting forecast skill. *Mon. Wea. Rev.,* **115,** 349–356.

Lacarra, J., and O. Talagrand, 1988: Short-range evolution of small perturbations in a barotropic model. *Tellus,* **40A,** 81–95.

Lorenc, A. C., 1986: Methods for numerical weather predictions. *Quart. J. Roy. Meteor. Soc.,* **112,** 1177–1194.

Lorenz, E., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus,* **17,** 321–333.

Moler, C., and C. F. Van Loan, 1978: Nineteen dubious ways to compute the exponential of a matrix. *SIAM Rev.,* **20,** 801–837.

Molteni, F., and T. N. Palmer, 1992: Predictability and finite-time instability of the Northern winter circulation. *Quart. J. Roy. Meteor. Soc.,* **119,** 269–298.

Mureau, R., F. Molteni, and T. N. Palmer, 1993: Ensemble predictions using dynamically conditioned perturbations. *Quart. J. Roy. Meteor. Soc.,* **119,** 299–323.

Navarra, A., 1993: A new system of orthonormal modes for linearized meteorological application. *J. Atmos. Sci.,* **50,** 2569–2583.

Palmer, T. N., 1988: Medium and extended range predictability and stability of the PNA mode. *Quart. J. Roy. Meteor. Soc.,* **114,** 691–713.

Simmons, A. J., J. M. Wallace, and G. W. Branstator, 1983: Barotropic wave propagation and instability, and atmospheric teleconnection patterns. *J. Atmos. Sci.,* **40,** 1363–1392.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of the perturbations. *Bull. Amer. Meteor. Soc.,* **12,** 2317–2330.

Van Loan, C., 1979: The sensitivity of the matrix exponential. *SIAM J. Numer. Anal.,* **14,** 971–981.