MESOSCALE PREDICTABILITY: AN ASSESSMENT THROUGH ADJOINT METHODS

Martin Ehrendorfer*,+, Joseph J. Tribbia*, Ronald M. Errico*

*National Center for Atmospheric Research
Boulder, Colorado, USA

+Institute for Meteorology and Geophysics
University of Vienna, Vienna, Austria

1. Introduction and motivation

The study of atmospheric predictability is primarily concerned with the question of how rapidly and by what physical processes perturbations of small amplitude are amplifying in the atmosphere (Lorenz 1982, 1990). Experience with extremely simplified atmospheric models (e.g., Lorenz 1963), as well as with complex numerical weather prediction (NWP) models has indicated that two initially slightly different states – each evolving according to the same physical laws – may over time develop into states no more similar than two randomly chosen observed states of the atmosphere. This inherent error growth is not an artifact of NWP models, but is a consequence of the nonlinearity and instability of atmospheric dynamics (Leith 1978, Tribbia 1996). It is this internal error–growth mechanism that has been the principal subject of predictability studies (e.g., Leith 1983, Shukla 1985, Thompson 1985, Boer 1994). As a consequence, such error growth, in conjunction with unavoidable inaccuracies in the specification of the initial model state, necessarily leads to limited predictability of the atmosphere.

1.1. Mesoscale predictability

In the present paper the emphasis is on investigating in some detail the predictability of atmospheric flows within a regional mesoscale primitive–equation model. The predictability of mesoscale circulations (considered here on horizontal scales of order a hundred kilometers) has, and still is, been a subject of considerable controversy. Based on a simplified model that included a wide range of different spatial scales, Lorenz (1969) demonstrated that smaller–scale phenomena and, especially, mesoscale phenomena should be less predictable (i.e., have shorter error–doubling times) than synoptic– or planetary–scale circulations (see also, Tennekes 1978). Consequently, as small–scale errors introduce errors

in the larger scales through an inverse cascade of error to larger scales (e.g., Tribbia and Anthes 1987, Leith 1971), he concluded that the time over which prediction of instantaneous weather patterns is possible is intrinsically limited by a finite range of predictability, currently conservatively estimated to be about ten days. This predictability limit must be distinguished from the considerably longer limit for low–frequency, planetary–scale anomalies (see, e.g., Cane 1992, Kumar et al. 1996, Palmer and Anderson 1994).

In contrast to the results of Lorenz (1969), Anthes et al. (1985) found from numerical integrations of a primitive–equation regional model that initial perturbations of considerable size only grew very slowly (if at all) over integration periods of several days. This behavior led to the hypothesis that mesoscale circulations are inherently more predictable, as they are controlled (possibly to a large degree) by the specification of the lateral boundary conditions, the influence of the large–scale flow, and orographic forcing.

In view of these seemingly contradictory results, Errico and Baumhefner (1987) undertook a detailed investigation of the reasons responsible for this behavior. They showed that three major constraining influences (i.e., strong numerical dissipation acting to damp spatially uncorrelated perturbations, effective reduction of initial variance through dynamically uncorrelated perturbations projecting onto gravity waves, lack of boundary perturbations) dominated the results of Anthes et al. (1985), while errors in quasi–geostrophic components of the flow were growing with short doubling times. In further studies (e.g., Van Tuyl and Errico 1989, Vukićević and Errico 1990, Vukićević 1991) it was pointed out that results depended considerably on the structure of the initial perturbations (specifically their correlation structure). In addition, one might make the comment that the sources of possible dynamical instabilities should be expected to remain largely unchanged when atmospheric circulations are considered at mesoscale resolution, until scales of convective instability are resolved.

The motivation for the present study has been to investigate the question of mesoscale predictability from the viewpoint of optimally–growing perturbations, as the spectrum of these perturbations might allow one to draw some more general conclusions about the chance of finding error growth in a regional mesoscale model. The study by Ehrendorfer and Errico (1995), that formes the basis for much of this presentation, investigated the question of how likely it was to encounter a *growing* initial perturbation in a regional model similar to the one that had been used in the above–mentioned predictability studies. The framework used here to assess this likelihood, as well as relevant results are presented in the following sections.

1.2. Atmospheric predictability: amplification of perturbations

Consider the expansion of a (random) initial perturbation in terms of a complete basis set Y_i as:

$$\mathbf{r} = \mathbf{Y}\mathbf{c} = \sum_{i=1}^{m} c_i \mathbf{Y}_i, \tag{1.2.1}$$

where m denotes the dimension of $\mathbf{r}$. The size of $\mathbf{r}$ may be measured in terms of a norm, described in finite dimension through the matrix $\mathbf{M}$:

$$\| \mathbf{r} \|^2 \equiv \mathbf{r}^T \mathbf{M} \mathbf{r}, \tag{1.2.2}$$

where the superscript ^{T} denotes a transpose. In atmospheric predictability experiments carried out with NWP models, the time–evolution of the initial perturbation $\mathbf{r}$ is usually determined as the difference between two nonlinear integrations of the relevant NWP model (i.e., the difference of perturbed nonlinear and control integration). The time–evolution of this difference can, to a high degree of accuracy (depending on the initial size of $\mathbf{r}$ and the forecast time) be approximated by the tangent–linear model (TLM) corresponding to the relevant NWP model and a control integration (e.g., Lacarra and Talagrand 1988, Errico et al. 1993, Buizza 1995). Under this approximation, and denoting the resolvent of the TLM by the operator $\mathbf{L}$, the magnitude of the time–evolved perturbation $\mathbf{r}$ becomes:

$$\| \mathbf{L}\mathbf{r} \|^2 \equiv (\mathbf{L}\mathbf{r})^T \mathbf{M} (\mathbf{L}\mathbf{r}). \tag{1.2.3}$$

From the foregoing, the amplification a^2 of $\mathbf{r}$ over the time period considered (this time period is implicitly contained in $\mathbf{L}$) when measured using the same norm both at the initial and the final times is given by:

$$a^2 \equiv \frac{\| \mathbf{L}\mathbf{r} \|^2}{\| \mathbf{r} \|^2} = \frac{(\mathbf{L}\mathbf{r})^T \mathbf{M} (\mathbf{L}\mathbf{r})}{\mathbf{r}^T \mathbf{M} \mathbf{r}} = \frac{(\mathbf{L}\mathbf{Y}\mathbf{c})^T \mathbf{M} (\mathbf{L}\mathbf{Y}\mathbf{c})}{(\mathbf{Y}\mathbf{c})^T \mathbf{M} \mathbf{Y}\mathbf{c}} = \frac{\mathbf{c}^T (\mathbf{Y}^T \mathbf{L}^T \mathbf{M} \mathbf{L} \mathbf{Y}) \mathbf{c}}{\mathbf{c}^T (\mathbf{Y}^T \mathbf{M} \mathbf{Y}) \mathbf{c}}. \tag{1.2.4}$$

This expression for the amplification a^2 simplifies to:

$$a^2 = \frac{\mathbf{c}^T \mathbf{\Lambda} \mathbf{c}}{\mathbf{c}^T \mathbf{c}} = \Big(\sum_{i=1}^{m} c_i^2 \lambda_i \Big) \Big/ \Big(\sum_{i=1}^{m} c_i^2 \Big) \tag{1.2.5}$$

for a set $\mathbf{Y}$ with:

$$\mathbf{Y}^T \mathbf{L}^T \mathbf{M} \mathbf{L} \mathbf{Y} = \mathbf{\Lambda} \qquad \text{and} \qquad \mathbf{Y}^T \mathbf{M} \mathbf{Y} = \mathbf{I}. \tag{1.2.6}$$

The set of singular vectors (SVs) as defined below (see section 2.2) possesses the properties (1.2.6) and may therefore be used advantageously for the expansion (1.2.1). Consequently,

given (part of) the matrix $\mathbf{\Lambda}$ (i.e., the amplifications of the leading SVs) which is similar to $\mathbf{L}^T\mathbf{ML}$ (through the SVs $\mathbf{Y}$) the (expected) amplification of any given perturbation $\mathbf{r}$ may be computed from (1.2.5), if the expansion coefficients c_i (or their statistical properties) are known. This approach forms the basis for assessing the expected amplification of random perturbations in the present work (see section 3.2).

2. Background

2.1. Non–modal finite–time instability

In the consideration of the stability properties of linear(ized) dynamical systems, written generically in the form:

$$\frac{d\mathbf{w}}{dt} = \mathbf{Hw}, \tag{2.1.1}$$

where $\mathbf{w}$ denotes the model state vector, and $\mathbf{H}$ describes the linearized dynamics, it is advantageous to distinguish between systems characterized through *normal* and *nonnormal* operators $\mathbf{H}$. An operator $\mathbf{H}$ is said to be normal, if it commutes with its adjoint $\mathbf{H}^\dagger$:

$$\mathbf{HH}^\dagger = \mathbf{H}^\dagger\mathbf{H}, \tag{2.1.2}$$

where the adjoint of $\mathbf{H}$ is defined as the unique operator $\mathbf{H}^\dagger$ satisfying the following relationship for a given inner product:

$$(\mathbf{u}, \mathbf{Hv}) = (\mathbf{H}^\dagger\mathbf{u}, \mathbf{v}) \tag{2.1.3}$$

(here $\mathbf{u}$ and $\mathbf{v}$ are two arbitrary elements of the Hilbert space considered). Obviously, symmetric (self–adjoint) operators/matrices, as well as unitary (orthonormal) matrices are normal (for the Euclidean inner product). The class of normal operators is important with regard to its spectral properties. Normal operators/matrices possess a complete and *orthogonal* set of eigenvectors (achievable even in degenerate situations). In fact, it is true that a matrix $\mathbf{H}$ is normal if and only if it is unitarily similar to a diagonal matrix. For nonnormal matrices an *orthogonal* set of eigenvectors does not exist (in general), and the set of eigenvectors may (or may not) be complete.

Clearly, through the above definition, normality of an operator (and the associated orthogonality of its eigenvectors) depend on the inner product considered, since the form of the adjoint operator $\mathbf{H}^\dagger$ will depend on this specification. It should therefore be pointed out that, for given $\mathbf{H}$, it is always possible to consider an inner product such that normality of $\mathbf{H}$ will result; or, in other words, there is always an inner product in terms of which

the eigenvectors of **H** are orthogonal. In fact, this very special inner product depends in its construction explicitly on the eigenstructure of **H** (see, Farrell and Ioannou 1993a, b). However, it should also be pointed out that, in general, such very specific norms are not physically meaningful; in fact, for physically meaningful norms (e.g., an energy or enstrophy norm), **H** is in meteorologically relevant applications, in general, nonnormal.

The implication of the above considerations for the study of the stability of atmospheric flows relates to the fact that the governing equations, when linearized about complex basic states lead (in general) to nonnormal dynamical systems. As a consequence, the non–orthogonality of the accompanying eigenvectors will allow for *finite–time* growth that may be quite different from growth described by exponentially growing shape–preserving normal–mode solutions (because of interacting non–orthogonal eigenvectors). In fact, through the process of interacting non–orthogonal eigenvectors, growth is possible, over a finite time at least, even for systems found to be stable in a conventional normal–mode analysis.

Such growth over finite times has first been studied in a meteorological context by Lorenz (1965), and has received considerable attention recently as being potentially important for explaining growth of synoptic–scale disturbances (e.g., Farrell 1988, 1989, 1990). In a tangent–linear context the *optimally*–growing structure can be found by solving a symmetric eigenvalue problem (see below). The property of these *optimal perturbations* (used synonymously with the term SV) of amplifying most rapidly over a finite–time interval makes them a highly useful concept for studying various questions, including atmospheric predictability and growth arising from instabilities (e.g., Mukougawa et al. 1991, Molteni and Palmer 1993, Yoden and Nomura 1993, Palmer 1993, Buizza and Palmer 1995, Nicolis et al. 1995, Trevisan and Legnani 1995). Further, expansions of the form (1.2.1) in terms of the normal modes of linearized evolution equations (such as (2.1.1)) were, for example, considered by Farrell and Moore (1992), Borges and Hartmann (1992), and Borges and Sardeshmukh (1995) in the study of finite–time growth in nonnormal systems. It is of interest to note in this context that the adjoint modes (i.e., the eigenfunctions of the adjoint operator) are used to determine the operators for projecting fields onto particular exponentially–growing normal modes. In addition, optimal perturbations allow the investigation of the stability properties of *non–autonomous* linear systems (for which the normal–mode concept is not immediately applicable), that result when linearization is performed about time–varying basic states (as is usually the case for TLMs).

Many further applications exist that are related to the concept of optimal perturbations. For example, Penland and Sardeshmukh (1995) investigated optimal growth occuring through the constructive interference of decaying normal modes in a low–order

dynamical model describing tropical sea surface temperature anomalies. The potential applicability of SVs for the problem of predicting forecast skill (e.g., the efficient generation of initial perturbations) as well as within the area of data assimilation has been studied within various classes of models by, for example, Buizza et al. (1993), Mureau et al. (1993), Thépaut et al. (1993, 1996), Molteni et al. (1996), Järvinen et al. (1996), Rabier et al. (1996), Houtekamer (1995), Houtekamer and Derome (1995), and Ehrendorfer and Tribbia (1995, 1996). Additional reference in the context of predicting forecast skill is made here to the work by Epstein (1969), Thompson (1986), and Toth and Kalnay (1993).

2.2. Definition of optimal perturbations

Optimal perturbations are defined here as the sequence of solutions to the following maximization problem.

$$\text{Maximize}: \qquad \tilde{J}(\mathbf{x}_0) = (\mathbf{P}\mathbf{L}\mathbf{x}_0)^T \mathbf{C} (\mathbf{P}\mathbf{L}\mathbf{x}_0) \tag{2.2.1a}$$

$$\text{subject to}: \qquad (\mathbf{P}\mathbf{x}_0)^T \mathbf{C} (\mathbf{P}\mathbf{x}_0) = 1 \quad \text{and} \quad \mathbf{x}_0 = \mathbf{Q}\mathbf{y}_0 \tag{2.2.1b}$$

where:

$$\mathbf{C} \equiv \mathbf{A}^T \mathbf{A} > 0 \tag{2.2.2}$$

is a symmetric, positive–definite norm–defining matrix. The operator $\mathbf{L}$ denotes (as before) the resolvent of the TLM suitable in any given context to propagate an n–dimensional initial perturbation $\mathbf{x}_0$ over a specified time interval (e.g., the solution of system (1.2.1)). The matrices $\mathbf{A}$, $\mathbf{P}$, and $\mathbf{Q}$ defining the choice of the norm considered here will be specified in section 3.1. The first constraint in (2.2.1b) ensures that $\tilde{J}(\mathbf{x}_0)$ is unity initially (i.e., for $\mathbf{L}$ the identity); the second constraint is required for constraining the components of $\mathbf{x}_0$ not constrained by the first one; that is, for a well–posed maximization problem it is necessary that $\mathbf{x}_0$ be specified in terms of an m–dimensional initial perturbation $\mathbf{y}_0$ ($m \leq n$), if $\mathbf{P}^T\mathbf{C}\mathbf{P}$ is of rank m. Clearly, an optimization problem different from the above may be stated in a way such that $\tilde{J}$ does not degenerate to the initial constraint for $\mathbf{L}$ the identity; this may be desirable in certain applications, but, for the reasons discussed below, is not done here.

Using the second constraint in (2.2.1b), the maximization problem stated above may be equivalently formulated in terms of $\mathbf{y}_0$ as:

$$\text{maximize}: \qquad J(\mathbf{y}_0) = (\mathbf{P}\mathbf{L}\mathbf{Q}\mathbf{y}_0)^T \mathbf{C} (\mathbf{P}\mathbf{L}\mathbf{Q}\mathbf{y}_0) \tag{2.2.3a}$$

$$\text{subject to}: \qquad (\mathbf{P}\mathbf{Q}\mathbf{y}_0)^T \mathbf{C} (\mathbf{P}\mathbf{Q}\mathbf{y}_0) = 1. \tag{2.2.3b}$$

This problem therefore asks for the initial perturbation $\mathbf{y}_0$ giving a maximum value for the quadratic function $J(\mathbf{y}_0)$, when the initial value of J is constrained by (2.2.3b). Due to the special form of the problem (2.2.3), specifically its quadratic form, it may be solved efficiently through the following eigenproblem (as may be seen by defining and subsequently differentiating a Lagrangian):

$$\left(\mathbf{B}^T\mathbf{B}\right)^{-1} \underbrace{\left(\mathbf{Q}^T\mathbf{L}^T\mathbf{P}^T\mathbf{C}\mathbf{P}\mathbf{L}\mathbf{Q}\right)}_{\equiv\tilde{\mathbf{S}}} \mathbf{y}_0 = \lambda\mathbf{y}_0 \quad \text{s.t.} \quad \mathbf{y}_0^T\mathbf{B}^T\mathbf{B}\mathbf{y}_0 = 1 \tag{2.2.4}$$

where:

$$\mathbf{B} \equiv \mathbf{A}\mathbf{P}\mathbf{Q}. \tag{2.2.5}$$

The set of eigenvectors of the eigenproblem (2.2.4) is referred to as *optimal perturbations*, or *singular vectors* (SVs). Any given SV, corresponding to an eigenvalue λ, when used in (2.2.3a), will lead to $J = \lambda$. Therefore, and due to the constraint (2.2.3b), the eigenvalues directly reveal the amplification a^2 [see (1.2.4)] of a given SV. The function J and the initial constraint in (2.2.3) correspond closely, in the sense that (2.2.3a) reduces to (2.2.3b) at the initial time (i.e., for $\mathbf{L} = \mathbf{I}$). Such correspondence may not appear, for example, if no projections are used initially, but if projections are used at the final time (in other words, different norms are used at initial and final times); in this case, an eigenvalue λ does no longer (necessarily) correspond to amplifications. Note also that the set of SVs $\mathbf{y}_0$ satisfies a relationship of the form (1.2.6), as is immediately seen by operating on the first part of (2.2.4) by $\mathbf{y}_0^T\mathbf{B}^T\mathbf{B}$. Also, if $\mathbf{y}_0$ is a solution to (2.2.3), then $\mathbf{Q}\mathbf{y}_0$ is clearly a solution to (2.2.1).

In order to point out the relationship between the optimal perturbations $\mathbf{y}_0$ and the *right* and *left* singular vectors of the modified (by the norm–defining matrices) resolvent $\tilde{\mathbf{L}}$ one may define:

$$\mathbf{z}_0 \equiv \mathbf{B}\mathbf{y}_0 \tag{2.2.6}$$

which allows (2.2.4) to be rewritten in standard (symmetric) form:

$$\underbrace{\left(\mathbf{B}^{-1}\right)^T\tilde{\mathbf{S}}\mathbf{B}^{-1}}_{\equiv\tilde{\mathbf{L}}^T\tilde{\mathbf{L}}}\mathbf{z}_0 = \lambda\mathbf{z}_0 \quad \text{s.t.} \quad \mathbf{z}_0^T\mathbf{z}_0 = 1 \tag{2.2.7}$$

where:

$$\tilde{\mathbf{L}} \equiv \mathbf{A}\mathbf{P}\mathbf{L}\mathbf{Q}\mathbf{B}^{-1}. \tag{2.2.8}$$

Introducing the *singular value decomposition (SVD)* (e.g., Golub and Van Loan 1989) of $\tilde{\mathbf{L}}$ as:

$$\tilde{\mathbf{L}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \qquad \text{with}: \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}, \quad \mathbf{V}^T\mathbf{V} = \mathbf{I}, \tag{2.2.9}$$

one obtains:

$$\tilde{\mathbf{L}}^T\tilde{\mathbf{L}} = \mathbf{V}\mathbf{\Sigma}^2\mathbf{V}^T \tag{2.2.10}$$

which, when compared with (2.2.7) in the form:

$$\tilde{\mathbf{L}}^T\tilde{\mathbf{L}} = \mathbf{Z}_0\mathbf{\Lambda}\mathbf{Z}_0^T, \tag{2.2.11}$$

shows that the matrix $\mathbf{Z}_0$ contains the *right singular vectors* (i.e., $\mathbf{V}$) of $\tilde{\mathbf{L}}$, since $\mathbf{U}$ and $\mathbf{V}$ are (by definition) the matrices of left and right singular vectors of $\tilde{\mathbf{L}}$, respectively. The left singular vectors may be obtained as multiples of the time–evolved right singular vectors (see, e.g., Buizza et al. 1995, Ehrendorfer and Tribbia 1996). The positive square–root of $\mathbf{\Lambda}$ is the matrix of *singular values* $\mathbf{\Sigma}$. Thus, the optimal perturbations, or SVs, $\mathbf{y}_0$, defined in the context of (2.2.3, 4) are obtained from the right singular vectors $\mathbf{Z}_0$ of $\tilde{\mathbf{L}}$ through (2.2.6). Clearly, one may also define the singular value decomposition of $\mathbf{L}$ itself, and then establish the norm–dependent relationship of that SVD to the SVD of $\tilde{\mathbf{L}}$ discussed above.

2.3. The mesoscale model MAMS1

The model used in the study discussed here is the dry–adiabatic (with the exception of section 4) version of the Mesoscale Adjoint Modeling System version 1 (Errico et al. 1994). This model is in important aspects closely related to the model used for the predictability experiments described by Anthes et al. (1985) (see also, Anthes et al. 1987). Results are presented here for two cases, namely a rapid cyclogenesis off the East Coast of North America, and a lee cyclogenesis over the Mediterranean (see also section 3.2). A nonlinear integration (including a nonlinear normal–mode initialization) over 24 hours defines the basic states that are used to define the TLM (and its corresponding adjoint). The horizontal grid spacing in the numerical integrations is 120 km; the domain considered includes all of the contiguous United States and parts of the Atlantic ocean (for case 1), and parts of the Eastern Atlantic and western Europe (for case 2). In the TLM integrations, no perturbations are considered at the boundaries of the domain; however, as these boundaries are far removed from the storms in the interior of the domain, their influence on the growth of perturbations is considered to be small.

2.4. The Lanczos algorithm

The Lanczos algorithm (Lanczos 1950; see also, e.g., Strang 1986, Grimes et al. 1994) is a semi–direct method for finding an approximation to the eigenstructure of a symmetric matrix $\mathbf{G}$, in the situation in which the elements of $\mathbf{G}$ are not explicitly known and/or $\mathbf{G}$

is too large to be kept in computer memory. Both of these restrictions usually apply when dealing with primitive–equation NWP models. For the problem of computing optimal perturbations, the matrix $\mathbf{G}$ is identified with the matrix $\tilde{\mathbf{L}}^T\tilde{\mathbf{L}}$ [see (2.2.7)]; from the eigenvectors of $\tilde{\mathbf{L}}^T\tilde{\mathbf{L}}$, the SVs are immediately obtained from (2.2.6).

The Lanczos–algorithm is based on the fact that for any symmetric matrix $\mathbf{G}$ there exists an orthonormal sequence $\mathbf{q}_1, \mathbf{q}_2, \ldots$ such that:

$$\mathbf{G}\mathbf{q}_j = b_{j-1}\mathbf{q}_{j-1} + a_j\mathbf{q}_j + b_j\mathbf{q}_{j+1}. \tag{2.4.1}$$

Through the orthogonal matrix $\mathbf{Q}$ (different from the $\mathbf{Q}$ used in other sections of this paper), $\mathbf{G}$ is similar to the tridiagonal matrix $\mathbf{T}$ (described by the sequence a_j, b_j; see below):

$$\mathbf{Q}^T\mathbf{G}\mathbf{Q} = \mathbf{T}, \quad \Leftrightarrow \quad \mathbf{G}\mathbf{Q} = \mathbf{Q}\mathbf{T}. \tag{2.4.2}$$

By rewriting (2.4.1) and (2.4.2) as:

$$b_j\mathbf{q}_{j+1} = \mathbf{G}\mathbf{q}_j - a_j\mathbf{q}_j - b_{j-1}\mathbf{q}_{j-1}, \tag{2.4.3}$$

$$(\,\mathbf{G}\mathbf{q}_1 \quad \mathbf{G}\mathbf{q}_2 \quad \ldots \quad \mathbf{G}\mathbf{q}_n\,) = (\,\mathbf{q}_1 \quad \mathbf{q}_2 \quad \ldots \quad \mathbf{q}_n\,) \begin{pmatrix} a_1 & b_1 & 0 & \ldots & 0 \\ b_1 & a_2 & b_2 & \ldots & 0 \\ \ldots & b_2 & a_3 & b_3 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & b_{n-1} & a_n \end{pmatrix}, \tag{2.4.4}$$

it is seen that the following *Lanczos iteration* can be used for the computation of the a, b, and $\mathbf{q}$ sequences:

$$\text{start}: \quad j = 0 \quad \mathbf{q}_0 = \mathbf{0} \quad b_0 = 1 \quad \mathbf{s}_0 = \mathbf{q}_1 \tag{2.4.5}$$

$$\text{iterate}: \quad \begin{cases} \text{(i)} \quad \mathbf{q}_{j+1} = \mathbf{s}_j/b_j \qquad \text{(ii)} \quad j = j+1 \\ \text{(iii)} \quad \mathbf{x} = \mathbf{G}\mathbf{q}_j \qquad \text{(iv)} \quad a_j = \mathbf{q}_j^T\mathbf{x} \\ \text{(v)} \quad \mathbf{s}_j = \mathbf{x} - a_j\mathbf{q}_j - b_{j-1}\mathbf{q}_{j-1} \qquad \text{(vi)} \quad b_j = \sqrt{\mathbf{s}_j^T\mathbf{s}_j} \end{cases} \tag{2.4.6}$$

It should be pointed out that each new vector $\mathbf{q}_{j+1}$ is orthogonal to all previous vectors $\mathbf{q}$. The destruction of this orthogonality by numerical roundoff errors poses one of the main difficulties in practical implementations of the Lanczos iteration. Note that within the Lanczos iteration, the only information about $\mathbf{G}$ enters through the matrix–vector product $\mathbf{G}\mathbf{q}_j$, which provides an adequate way for dealing with the two restrictions mentioned above. As a semi–direct method, the Lanczos algorithm rapidly provides, through the approximation of $\mathbf{T}$, good estimates for the leading portion of the eigenspectrum of $\mathbf{G}$. Experience indicates that k Lanczos iterations provide reliable (i.e., converged) information

about the first $k/3$ eigenvalues/vectors of $\mathbf{G}$. Since the matrices $\mathbf{G}$ and $\mathbf{T}$ are similar, the eigenvectors of $\mathbf{G}$ are obtained by operating with $\mathbf{Q}$ on the eigenvectors of $\mathbf{T}$. In the work presented here, the Lanczos iteration was performed with code written by B.N. Parlett, and provided through the Numerical Algorithms Group (NAG).

3. SV–Spectra for a dry version of MAMS1

The computation of singular vectors requires the specification of a norm (see (2.2.3)). In the work presented here, attention is restricted to results obtained with a rotational–modes norm (R–norm). In the model version relevant for the computations discussed here, the R–norm proved to have a number of advantages (e.g., geostrophic–adjustment related processes that might contribute to growth are efficiently excluded) when compared to the total energy norm that has been used in other studies (e.g., Buizza et al. 1993, Molteni et al. 1996). A brief description of this norm, as well as of results relevant for predictability, are contained in the following two subsections.

3.1. A rotational–modes norm (R–norm)

The main goal in the design of the R–norm was to allow only SVs that were initially geostrophically balanced and to omit the potential influence of gravity waves on measuring growth. This aim is achieved by considering the normal–mode representation of the model variables. Specifically, the operators $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{C}$ introduced in section 2.2 take on the following forms (for a more complete description, refer to Ehrendorfer and Errico 1995).

The $(m \times n)$ operator $\mathbf{P}$ represents the projection of the physical fields u, v, T, and p_s (horizontal wind, temperature, and surface pressure) onto the amplitudes of the rotational modes $\tilde{\Pi}_{\tilde{m},\tilde{n},\tilde{k}}$ in the normal-mode space of the model (horizontal wavenumbers $\tilde{m}, \tilde{n}$):

$$\begin{pmatrix} u \\ v \\ T \\ p_s \end{pmatrix} \xrightarrow{\mathbf{P}} \tilde{\Pi}_{\tilde{m},\tilde{n},\tilde{k}} \equiv \tilde{\zeta}_{\tilde{m},\tilde{n},\tilde{k}} - \frac{f_0}{g H_{\tilde{k}}} \tilde{\varphi}_{\tilde{m},\tilde{n},\tilde{k}}, \tag{3.1.1}$$

where $\tilde{\zeta}$ is the normal–mode representation of the vertical component of relative vorticity, f_0 is the Coriolis parameter at the center of the grid, g is gravity, and $H_{\tilde{k}}$ is the equivalent depth of the $\tilde{k}$th vertical mode; $\tilde{\varphi}$ is the normal–mode representation of the linearized model form of the pseudo–geopotential φ:

$$\varphi = \phi + R\bar{T} \ln \frac{p_s}{\bar{p}}, \tag{3.1.2}$$

where ϕ is hydrostatic geopotential (compare also, Errico 1991, Errico et al. 1994). As a consequence, $\mathbf{y}_0$, introduced in (2.2.1), consists of spectral components $\tilde{\Pi}_{\tilde{m},\tilde{n},\tilde{k}}$ and has dimension m.

The $(n \times m)$ operator $\mathbf{Q}$ can be regarded as an inverse to $\mathbf{P}$ (as the relation $\mathbf{PQ} = \mathbf{I}$ holds) and describes the construction of balanced fields u, v, T, and p_s from $\tilde{\Pi}_{\tilde{m},\tilde{n},\tilde{k}}$. In this construction, the constraints of geostrophy and non–divergence are enforced in the following form:

$$f_0\zeta - \nabla^2\varphi = 0, \qquad \delta = 0, \tag{3.1.3}$$

to get divergence-free, geostrophically balanced u, v, and φ. With an additional closure assumption (i.e., minimization of the sum of squares of the temperature gradients from level to level), the T and p_s fields are determined from φ. The property $\mathbf{PQ} = \mathbf{I}$ implies that the projection (through $\mathbf{P}$) of balanced fields $\mathbf{Qy}_0$ results again in the rotational normal–mode coefficients $\mathbf{y}_0$ (or, $\tilde{\Pi}_{\tilde{m},\tilde{n},\tilde{k}}$).

The $(m \times m)$ matrix of weights $\mathbf{C}$ in normal–mode space is taken as a diagonal matrix, but with wave–number dependent entries:

$$c_{\tilde{m},\tilde{n},\tilde{k}} \propto \frac{1}{\kappa^2_{\tilde{m},\tilde{n}}}, \tag{3.1.4}$$

leading to larger weights for larger horizontal scales. In addition to the projections, expressed through $\mathbf{P}$ and $\mathbf{Q}$, the normal–mode representations have been truncated in the vertical and also horizontally. The truncation is such that small horizontal and vertical scales were explicitly excluded (these scales should not contribute to growth in the model, as they are strongly damped in the model formulation). Through this procedure the resulting state vector is of size $m = 4830$. This number, in turn, defines the size of the $(m \times m)$ eigenproblem (2.2.4) (or, (2.2.7)) to be solved for the determination of the SVs.

3.2. R–norm spectra and implications for predictability

The leading portions of singular vector spectra were computed for the R–norm for two different cases, or basic states (case 1 is an explosive cyclogenesis over the Atlantic off the East Coast of North America, case 2 is a rapid lee cyclogenesis over the Mediterranean). In both cases the optimization interval has been taken to be $T = 24$ hours which ensures that the effect of the lateral boundaries is small. The R–norm spectra are shown in Fig. 1 in terms of the eigenvalues λ_i (see eq. (2.2.4)) as crosses (case 1) and dots (case 2).

The largest eigenvalues are 29.7 ($=5.45^2$, case 1) and 39.8 ($=6.31^2$, case 2), indicating that the first SVs amplify – in terms of the R–norm – by about a factor of six. Assuming

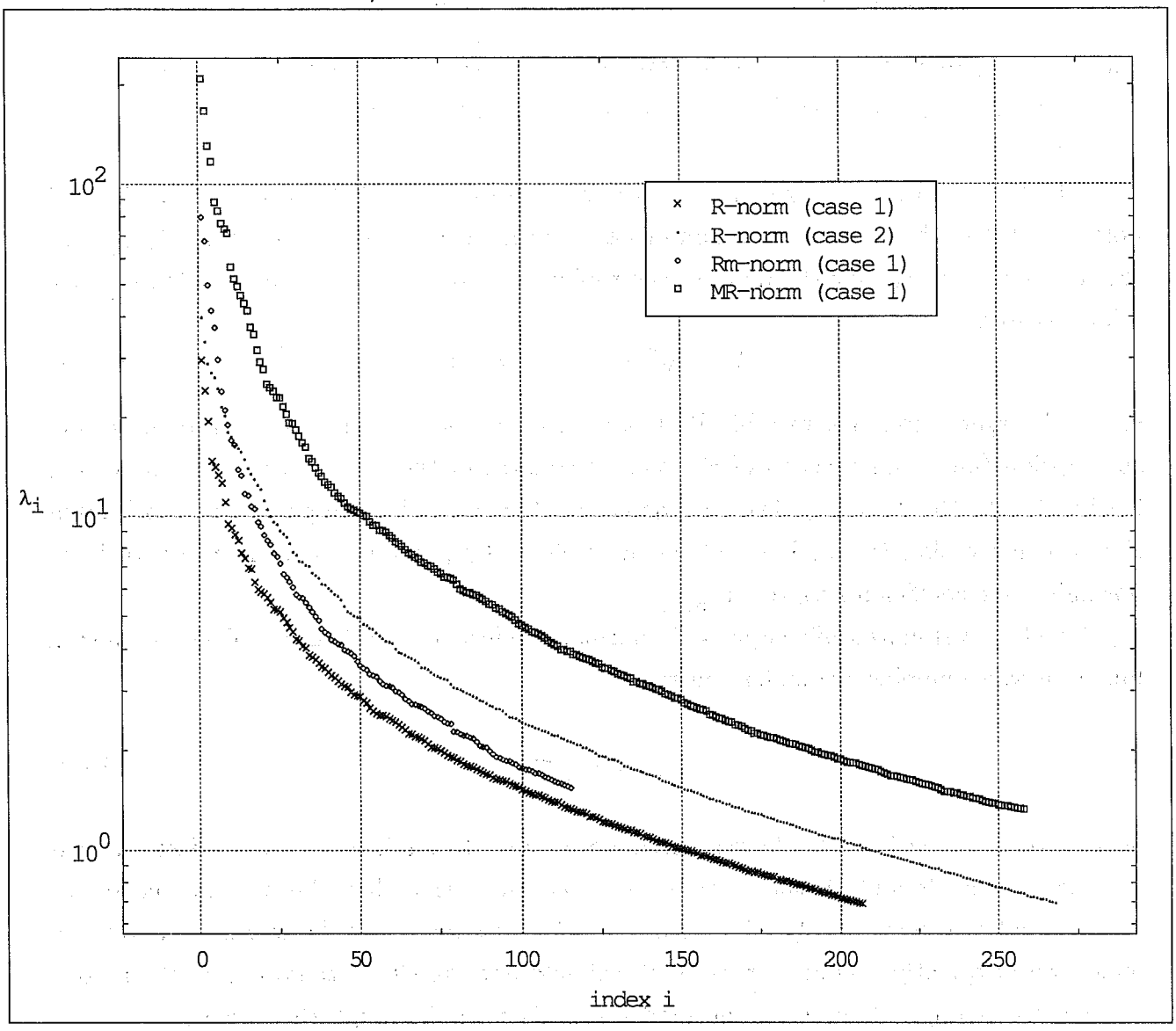

Fig. 1. Spectra of eigenvalues λ_i of $\tilde{\mathbf{L}}^T\tilde{\mathbf{L}}$ (see (2.2.7)) for dry experiments with the R–norm (crosses for case 1, dots for case 2), as well as for the two norms considered in the moist experiments for case 1 (diamonds for Rm–norm, squares for MR–norm). Note that the square–roots of the λ_i correspond to increases in length (as measured by the relevant norm) of the corresponding SV. Only the converged parts of the spectra are shown.

---

exponential growth, these amplifications may be related to an e–folding time τ ($\tau = T/\ln\lambda$; recall that λ is, in the situation in which the same norms are used at initial and final times, as done here, the ratio of the respective values of J). The e–folding times are 7.1 and 6.5 hours, respectively, translating into error–doubling times (denoted in Table 1 as τ_2) of 4.9 and 4.5 hours, respectively. These doubling times are significantly smaller than the error–

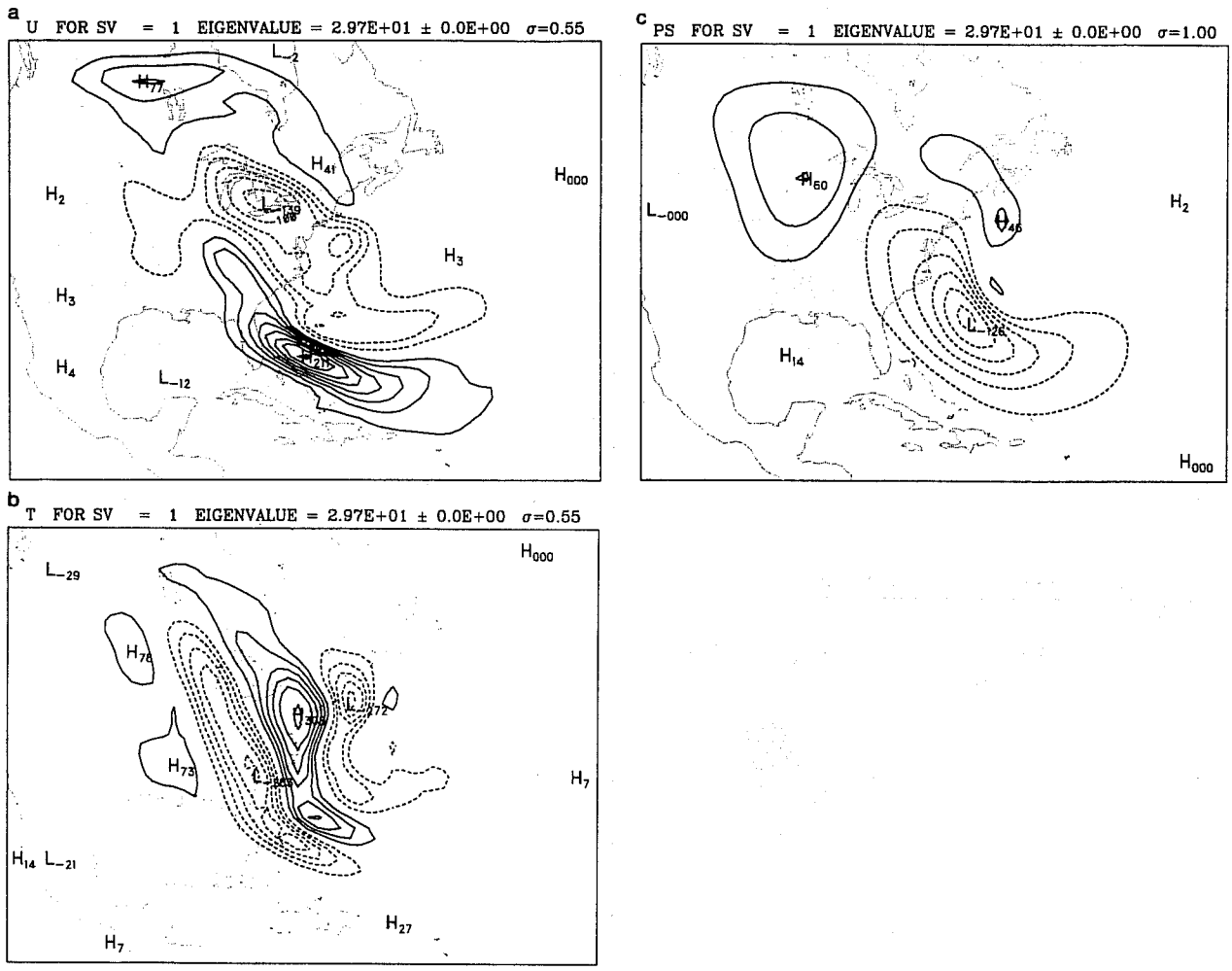

Fig. 2. First R–norm SV (at initial time) for case 1. Selected fields, that is, (a) zonal wind, and (b) temperature perturbations at $\sigma = 0.55$ (approximately 550 hPa), and (c) the surface pressure perturbation, are shown before carrying out the scaling for the tangent–linear integration.

---

doubling times between one and two days reported recently by Simmons et al. (1995) for the global forecast model at ECMWF.

Selected spatial patterns, that is, zonal wind and temperature perturbations at $\sigma = 0.55$ (approximately 550 hPa) and the surface pressure perturbation, of the first R–norm SV of case 1 are shown in Fig. 2 for the initial time. These fields are part of the first R–norm SV $\mathbf{y}_0$, as defined in (2.2.4), transformed into physical space according to (2.2.1b). Consequently, they satisfy the relevant normalization to one. The first SV is of considerable horizontal extent, but its structures are at the same time clearly connected to the upper–level trough at 500 hPa. The vertical structure (not shown) of this SV is strongly baroclinic. Fig. 3 shows the corresponding tangent–linearly time–evolved patterns at the end of the

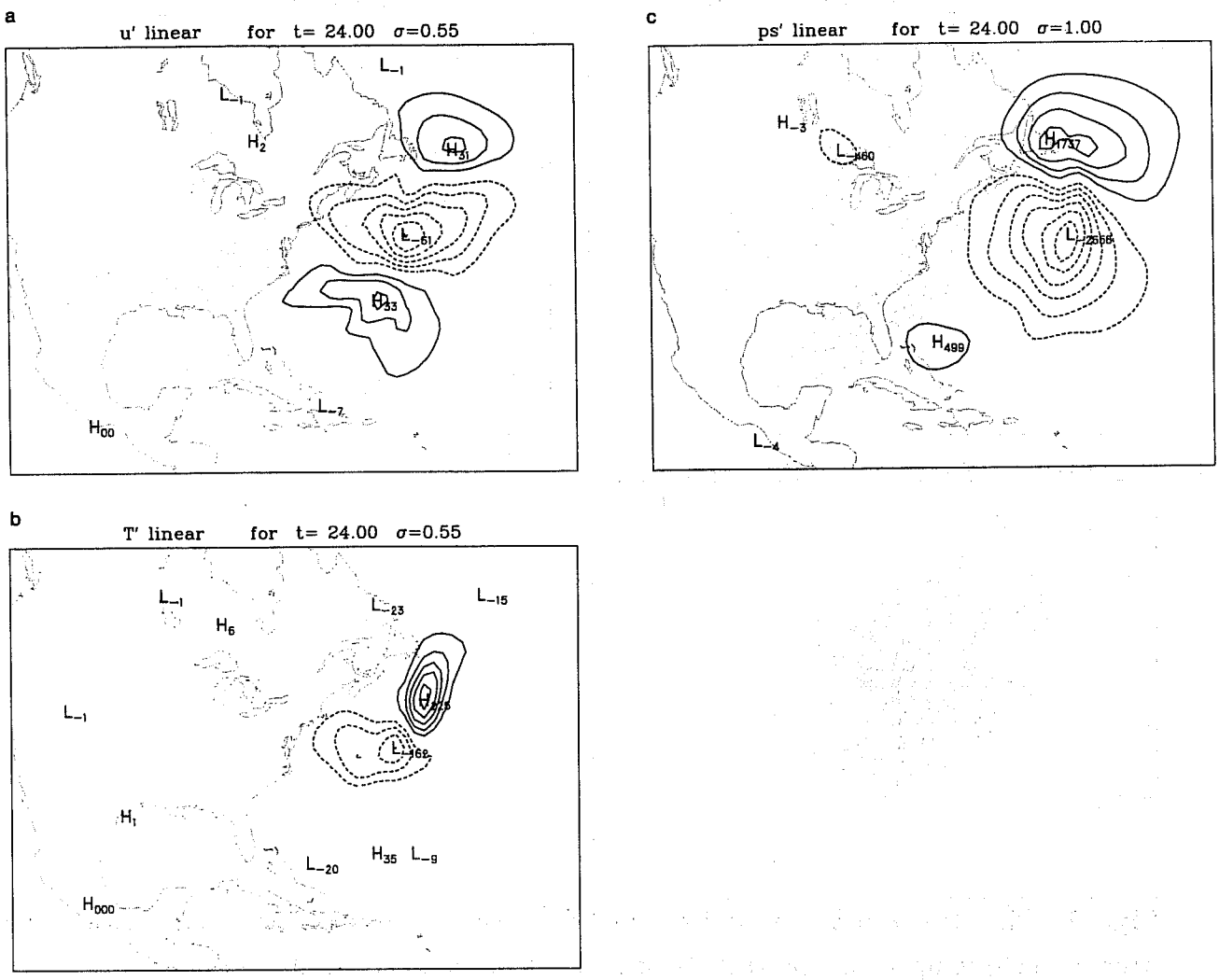

Fig. 3. Time–evolved first R–norm SV (at final time T=24 hours) for case 1, for the same selected fields as in Fig. 2. Note that this SV has been scaled at initial time by a factor of 0.7923 before the tangent–linear integration.

---

optimization interval. Before time–evolving the SV, it was scaled by a factor of 0.7923, in order to make its magnitude more comparable to current analysis errors. After this rescaling, the maximum initial perturbations in u, T, and p_s were 1.67 m s^{-1}, 0.24 K, and 1.0 hPa, respectively. At the end of the tangent–linear time evolution, the corresponding maxima are 6.1 m s^{-1}, 2.2 K, and 2.7 hPa, indicating that most of the growth occurs in the wind and temperature fields, resulting in the overall amplification of $\lambda = 5.45^2$ (see above). Also, the spatial structures of the time–evolved perturbation have followed the movement of the trough at 500 hPa.

At this point, the amplification of a perturbation expanded in terms of a given set of SVs $\mathbf{Y}$ will be derived using the result (1.2.5). Since the SVs possess the property (1.2.6)

(see section 2.2), (1.2.5) may be immediately used here in the form:

$$a^2 = \Big(\sum_{i=1}^{m} c_i^2 \lambda_i\Big) / \Big(\sum_{i=1}^{m} c_i^2\Big), \tag{3.2.1}$$

where the c_i are the expansion coefficients of a perturbation $\mathbf{r}$ in terms of the R–norm SVs (see eq. (1.2.1)). Since the SVs are normalized at the initial time to unity (see (2.2.4)), it is reasonable to normalize $\mathbf{r}$ similarly:

$$\| \mathbf{r} \|^2 = \mathbf{r}^T \mathbf{B}^T \mathbf{B} \mathbf{r} = 1, \tag{3.2.2}$$

which, together with (2.2.4) and (1.2.1) implies:

$$\sum_{i=1}^{m} c_i^2 = 1. \tag{3.2.3}$$

Note that any other initial normalization for $\mathbf{r}$ may be chosen without consequence for the following, as it will cancel in the expression for a^2. If $\mathbf{r}$ is randomly drawn from a certain ensemble with as yet unspecified distribution, the corresponding expectation operator, $< \ldots >$, can be applied to (3.2.1) to yield the following expression for expected growth (taking into account (3.2.3)):

$$< a^2 > = \sum_{i=1}^{m} \lambda_i < c_i^2 >. \tag{3.2.4}$$

From this expression, further insight with regard to expected amplifications may be obtained by making the assumption that the perturbations $\mathbf{r}$ have a white spectrum in terms of the SVs, which means that the $< c_i^2 >$ are independent of index i. As such, the $< c_i^2 >$ take on a constant value, say c^2. This assumption, together with (3.2.3), implies that $c^2 = m^{-1}$, or, in terms of expected growth:

$$< a^2 > = \frac{1}{m} \sum_{i=1}^{m} \lambda_i. \tag{3.2.5}$$

Thus, for normalized random perturbations with a spectrum that is white in terms of the SVs, the expected growth is the mean of the amplifications of the individual SVs (see also, Farrell 1990). Thus, assuming a white noise spectrum, this result may be used to assess whether it is likely to observe growth of such perturbations, or not, once the SV spectrum is available.

Table 1. Upper bounds μ^2 on expected values of growth. Shown is norm and case considered, the dimension m of the eigenproblem, the largest eigenvalue obtained ($\lambda_{\max}$) and the associated error–doubling time τ_2, the number k of accurately computed eigenvalues, as well as the last converged eigenvalue λ_k, the sum of the eigenvalues up to index k, and the upper bound μ^2 as defined in (3.2.6), together with its square–root μ.

norm	m	$\lambda_{\max}$	τ_2	k	λ_k	$\sum_{i=1}^{k}\lambda_i$	μ^2	μ
R/1	4830	29.74	4.9h	207	0.693	557.8	0.779	0.882
R/2	4830	39.78	4.5h	268	0.695	995.7	0.863	0.929
Rm/1	4830	79.67	3.8h	115	1.546	812.8	1.677	1.295
MR/1	29770	208.63	3.1h	258	1.332	2670.87	1.410	1.188

In the present high–dimensional contexts, only the leading portions of the SV–spectra have been computed to limit computational costs (see Fig. 1). Therefore, only an upper bound on $< a^2 >$ (as given by (3.2.5)), denoted below as μ^2, may be derived from that information, assuming that the amplifications of higher SVs are not larger than the amplification λ_k of the last SV computed (index k):

$$< a^2 >= \frac{1}{m}\sum_{i=1}^{m}\lambda_i = \frac{1}{m}\Big[\sum_{i=1}^{k}\lambda_i + \underbrace{\sum_{i=k+1}^{m}\lambda_i}_{\leq (m-k)\lambda_k}\Big] \leq \frac{1}{m}\Big[\sum_{i=1}^{k}\lambda_i + \lambda_k(m-k)\Big] \equiv \mu^2. \qquad (3.2.6)$$

The values of μ^2 derived from the spectra shown in Fig. 1, are listed in Table 1. For the experiments with the dry–adiabatic model version, these bounds are considerably smaller than one, indicating that growth of a random perturbation (assumed white in terms of the SVs) is not likely for these basic states; in fact, decay, as indicated by $\mu^2 < 1$ should be expected. It appears that the unstable subspace of the model is small (on the order of 150 to 200 growing SVs; see Fig. 1); in addition, the associated growth rates are too small to achieve growth of perturbations introduced without favoring projections on growing SVs. It should be emphasized, however, that this property of the model phase space cannot, without further investigation be used to vindicate the hypothesis put forward by Anthes et al. (1985) that mesoscale phenomena are inherently more predictable, partly also because the timescale considered here may be more responsible for the small unstable subspace rather than the fact that the spatial scales considered are mesoscale.

One may, at this point, question the degree of validity of the hypothesis that the perturbations considered are white in terms of the set of SVs. It is, however, arguable that the perturbations used in previous predictability studies would be white to a good approximation in terms of the R–norm SVs, since they were specified randomly with no prescribed correlation structures. Nevertheless, this hypothesis should clearly undergo more detailed investigation; for example, ongoing research concentrates on the question of how current analysis errors (as typical perturbation patterns) would project on the SVs spanning the unstable model subspace.

4. SVs in the presence of moist convection: preliminary results

Given the results obtained with the dry–adiabatic version of MAMS1, a few preliminary results relevant for finite–time nonmodal optimal growth are presented here that were obtained with a moist version of MAMS1. In this moist version, specific humidity is a prognostic variable; in addition, moist convection is parameterized using a stability–dependent mass flux representation of moist convective processes. This representation is used in the NCAR Community Climate Model (see, Hack 1994), but has been slightly modified in the present context (see, Errico et al. 1994).

One of the difficulties encountered when moist processes such as convection are considered in investigating optimal growth in terms of SVs is the construction of an appropriate tangent–linear model. Various aspects of this problem related to the above–mentioned scheme, as well as to other schemes (e.g., Moorthi and Suarez 1992) are presently under investigation (see also below). In the computations described here, the linearization of the scheme was computed numerically (instead of *coding* a tangent–linear model) by differencing a perturbed and unperturbed calculation of the convective parameterization module (see, Errico et al. 1994).

Another problem encountered when dealing with moist processes is associated with the choice of an appropriate norm to account for moisture. In the dry situation, reference to quantities conserved under (idealized) conditions (e.g., energy) allows for a physical justification for the choice of a specific norm. Such justification is no longer necessarily available when moist processes are taken into consideration.

In the present context, two norms were considered. In the first one, denoted as Rm–norm, the moist TLM is used, but specific humidity is not measured at the final time, and set to zero at the initial time; the dry fields are measured as within the R–norm. In the second one, denoted as moist R–norm (MR–norm), moisture is accounted for in the

following way:

$$J = J_d + \frac{1}{2}w^2 \sum_{i,j,k} (\Delta\sigma)_k q_{i,j,k}^2 \equiv J_d + J_m, \tag{4.1}$$

where J represents the measurement of the magnitude of the state vector (see 2.2.3a). The two terms J_d (dry contribution) and J_m (moist contribution) represent the magnitudes of the dry fields (horizontal wind components, temperature, and surface pressure) as measured through the R–norm (see section 3.1), and of the specific humidity field q in physical space over the full three–dimensional model domain, respectively. The vertical grid spacing $(\Delta\sigma)_k$ is included in J_m. In addition, the weighting factor w (taken independent of horizontal and vertical position) is included to allow varying the relative contributions of J_d and J_m to J.

In the experiments described here, w takes on the value of $10^{2.5}$. This particular weight was chosen after some experimentation, in order to have about equal contributions of dry and moist fields to J at the initial time. This rationale seems to be appropriate since otherwise the results asymptotically approach the same as when initially perturbing only dry fields but measuring only moisture at the forecast's end (as w is made larger) or vice–versa (as w is made smaller). It should be emphasized at this point that for both norms (Rm– and MR–norm) both dry fields and q at the final time will depend on both dry and moist initial fields, due to the interactions described by the moist TLM.

As in the dry experiments, the variational problem to be solved results from maximizing J subject to $J = 1$ at the initial time. For the Rm-norm, the relevant eigenproblem is of the same size in the dry situation, since q is not included in the control variable; for the MR–norm, the eigenproblem is of size $m = 29770$. As in the dry situation, the results presented in this section refer to an optimization time interval of T=24 hours; however, only the situation of case 1 with the basic state computed with the moist nonlinear model is considered here.

The portion of the eigenspectrum of the relevant operator $\tilde{\mathbf{L}}^T\tilde{\mathbf{L}}$ that has been computed reliably is included in Fig. 1 as open diamonds (Rm–norm), and as small squares (MR–norm). Relevant additional results are included in Table 1. The largest eigenvalues are 79.67 and 208.63, respectively, corresponding to amplifications of 8.92 and 14.44, respectively (for spatial patterns of selected SVs, reference is made to Errico and Ehrendorfer 1995). Comparison of these results with the R–norm computations using the dry model (crosses and dots in Fig. 1) shows clearly that the inclusion of moisture leads to (dramatically) increased amplification factors (e.g., the largest amplification factor for the dry model when measured with the R–norm was 5.45). Obviously, however, these differences depend considerably on the specification of the weights w. Another important difference

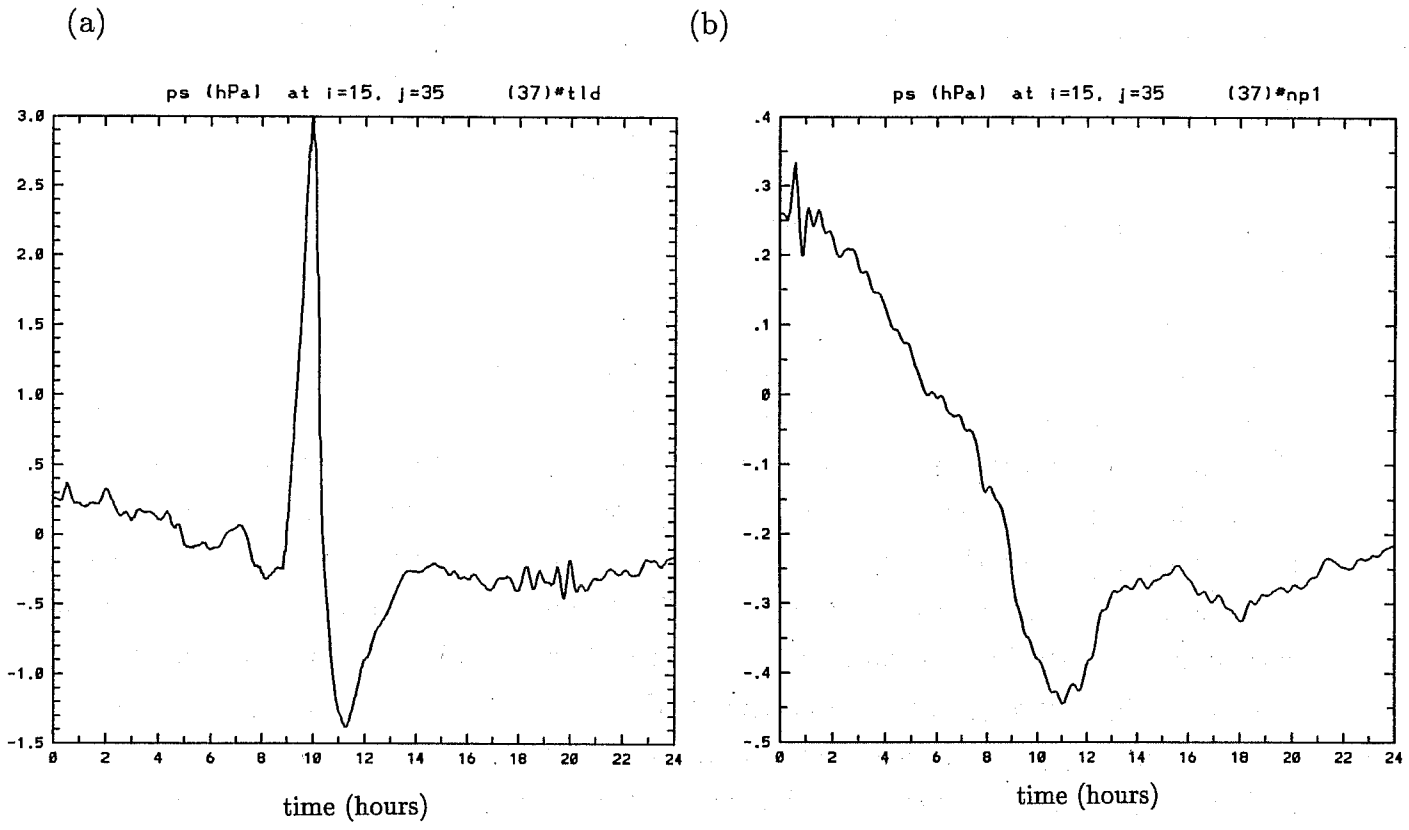

Fig. 4. Tangent–linear (panel a) and nonlinear (panel b) evolution of the surface pressure perturbation at a given grid point for case 1, as obtained from the first SV with the moist model and the MR–norm. The initial perturbation at this grid point is approximately 0.3 hPa.

---

compared to the dry experiments reported in section 3 (see also, Ehrendorfer and Errico 1995) is the fact that the number of growing perturbations has increased due to the inclusion of moisture. From Fig. 1 it can be seen that the number of growing perturbations is now (possibly significantly) larger than 250, whereas in the corresponding dry experiment approximately 150 growing SVs were found.

Under the same assumptions as have been made in section 3.2, upper bounds on the expected amplifications have been computed from the two moist spectra shown in Fig. 1. The relevant numbers for these two experiments are 1.677 (Rm–norm) and 1.410 (MR–norm) (see column 8 in Table 1). These numbers are slightly larger than one, and would in fact indicate that growth should be expected. Nevertheless, it should be pointed out that only small fractions (on the order of 1%) of the entire spectra are available (due to the extremely large cost of computing these spectra) for the assessment of these (highly conservative) upper bounds (eq. (3.2.6)) that consequently must decrease if amplifications

of higher–order SVs were included. Again, in terms of implications for predictability, the growth of random perturbations, white in terms of the SVs defined through the above norms, appears unlikely (in agreement with Anthes et al. 1985). However, as apparent from the short error–doubling times (less than four hours; see Table 1) this property of the phase space does not imply enhanced predictability of the processes represented in this model.

One of the problems encountered in the process of linearization of the moist convective parameterization is shown in terms of an example in Fig. 4, comparing tangent–linear (Fig. 4a) and nonlinear (Fig. 4b) time evolution of the surface pressure perturbation at a given grid point in the model domain. Apparently, the agreement between tangent–linear and nonlinear perturbation evolution for this perturbation size (approximately 0.3 hPa) is not satisfactory (note the large spike present in the tangent–linear time series). Clearly, such disagreement limits the validity of the TLM and, in turn, limits the inferences that can be made on the basis of tangent–linear optimal growth with regard to nonlinear error growth. Also, apparently the computation of SVs allows easy detection of potentially unstable segments in parameterizations. Even though time series at certain points appear to be having problems, the agreement between patterns appears to be reasonably good (for examples relevant for the present scheme, see Errico and Ehrendorfer 1995). Further investigation of this behavior is underway. Among other things, the corresponding behavior of different moist convective parameterizations (e.g., the relaxed Arakawa–Schubert scheme; see Moorthi and Suarez 1992) is investigated. Results from latest relevant experiments approximately will be reported in subsequent papers.

5. Summary and concluding remarks

The predictability of mesoscale circulations has been investigated on the basis of singular vectors for a regional, primitive–equation model. Singular vectors represent a highly useful generalization of the concept of normal modes, since they allow one to address in a general way the question of error growth in linearized dynamical systems that might be nonnormal, that is, possess a set of non–orthogonal eigenvectors for a given inner product. In such systems, even if they are found to be stable in a traditional normal–mode analysis, nonmodal perturbation growth is possible over finite times through the constructive interference of non–orthogonal (possibly decaying) normal modes. In addition, singular vectors also present a useful framework for stability studies in the situation of non–autonomous systems, such as those resulting from linearization about time–dependent basic states.

Amplification spectra of singular vectors were computed for two different basic states for a dry–adiabatic model version; preliminary results obtained using a model version that

included a linearized moist convective parameterization were discussed. Under certain assumptions these spectra allow derivation of bounds for expected amplifications of perturbations inserted into the model. Most notably, in the computation of the expected amplifications it has been assumed that the relevant perturbations have a white spectrum in terms of the SVs. More specifically, it is assumed that the variance of the expansion coefficients with which the SVs enter into the representation of the perturbations is the same for all SVs considered in the expansion. The validity of this assumption in the situation of typical perturbations (e.g., analysis errors) will require more detailed further study, but it should be reasonably valid for the norm considered here in view of the nature of perturbations used in previous mesoscale predictability studies.

One of the main results presented here is that these expected amplifications are smaller than one, or, at the most bounded from above by numbers slightly larger than one. This result indicates that such perturbations should not be expected to grow over the time period of one day considered here. In other words, the unstable subspace of the model identified through the growing SVs is a small fraction of the entire model phase space. It has been emphasized, however, that this property does not support the hypothesis made in previous predictability studies that mesoscale circulations are inherently more predictable than synoptic–scale, or planetary–scale circulations; one of the indications against such enhanced predictability are the short error–doubling times (of a few hours) found for the most rapidly growing SVs computed here.

Preliminary results of SV computations with a model version accounting for moist processes, especially moist convection parameterized in terms of the mass flux scheme described by Hack (1994), showed substantial differences when compared to the dry results. Specifically, for both norms considered (Rm–norm and MR–norm; see section 4), the amplifications of the leading SVs were considerably higher (e.g., by a factor of seven when comparing MR–norm and R–norm; see Table 1). In addition, the number of growing SVs has also been found to be substantially larger. However, these preliminary findings depend on a number of important issues. One issue relates to the norm that should be used to measure growth when moist processes are included. Here, it was decided to investigate growth through a straightforward extension of a rotational–modes amplitudes norm that was developed for measuring growth in a dry–adiabatic model. Another important question concerns the degree of the validity of the tangent–linear approximation – that is made in order to be able to compute the finite–time most unstable structures by (partially) solving an eigenproblem – in the presence of moist processes. Both issues will require further attention in future work, especially in view of the problems encountered with the linearized version of the moist convection scheme.

The study of structures growing optimally over finite times has recently been gaining increased attention for various reasons. As a generalization of normal–mode stability analysis they allow for studying the possibility of growth in nonnormal systems; in such systems the constructive interaction of nonorthogonal modes may lead to growth (over finite times) even if the system has been found to possess only stable normal modes. The relevance of such growth is currently an area of active research (e.g., predictability, cyclogenesis). For the purpose of studying the predictability of atmospheric flows, the relevance of SVs relates to the fact that their growth rates represent bounds on error growth rates maximally achievable in linearized situations. SVs are also heavily relied upon in the development of systems (ensemble prediction systems) designed for assessing the uncertainty of forecasts made with NWP models. In this latter context, the SV–property of most rapid finite–time amplification may allow one to base upon them sampling strategies that are highly efficient for achieving a given objective (such as covariance prediction) in view of extremely high–dimensional model phase spaces. In addition, for certain such objectives, SVs may also prove to be highly useful in data assimilation contexts. Given the potential of SVs for such wide applications, the further study of their properties and relevance, as initiated by Lorenz (1965), will likely continue to see wide interest in various areas.

Acknowledgments. The Lanczos algorithm used for the eigencomputations was written by B.N. Parlett, UC Berkeley, and was kindly provided to us by the Numerical Algorithms Group. The National Center for Atmospheric Research is sponsored by the National Science Foundation. Much of this work was carried out while the first author was visiting NCAR.

References

Anthes, R.A., Y.–H. Kuo, D.P. Baumhefner, R.M. Errico, and T.W. Bettge, 1985: Predictability of mesoscale motions. *Issues in atmospheric and oceanic modeling, Part B, Advances in Geophysics,* **28**, B. Saltzman and S. Manabe, Eds., Academic Press, 159–202.

Anthes, R.A., E.-Y. Hsie, and Y.-H. Kuo, 1987: *Description of the Penn State/NCAR Mesoscale Model Version 4 (MM4).* NCAR Tech. Note, NCAR/TN–282+STR. 66 pp. [Available from the National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307, U.S.A.]

Boer, G.J., 1994: Predictability regimes in atmospheric flow. *Mon. Wea. Rev.,* **122**, 2285–2295.

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

Borges, M.D., and P.D. Sardeshmukh, 1995: Barotropic Rossby wave dynamics of zonally varying upper–level flows during northern winter. *J. Atmos. Sci.,* **52**, 3779–3796.

Buizza, R., 1995: Optimal perturbation time evolution and sensitivity of ensemble prediction to perturbation amplitude. *Quart. J. Roy. Meteor. Soc.,* **121**, 1705–1738.

Buizza, R., J. Tribbia, F. Molteni, and T. Palmer, 1993: Computation of optimal unstable structures for a numerical weather prediction model. *Tellus,* **45A**, 388–407.

Buizza, R., R. Gelaro, F. Molteni, T.N. Palmer, 1995: Predictability studies with high resolution singular vectors. *Quart. J. Roy. Meteor. Soc.,* under review.

Buizza, R., and T.N. Palmer, 1995: The singular–vector structure of the atmospheric global circulation. *J. Atmos. Sci.,* **52**, 1434–1456.

Cane, M.A., 1992: Tropical Pacific ENSO models: ENSO as a mode of the coupled system. *Climate System Modeling,* K. Trenberth, Ed., Cambridge University Press, 583–614.

Ehrendorfer, M., and R.M. Errico, 1995: Mesoscale predictability and the spectrum of optimal perturbations. *J. Atmos. Sci.,* **52**, 3475–3500.

Ehrendorfer, M., and J.J. Tribbia, 1995: Efficient prediction of covariances using singular vectors. *Preprint Volume, Sixth International Meeting on Statistical Climatology,* Galway, Ireland, 135–138.

Ehrendorfer, M., and J.J. Tribbia, 1996: Optimal prediction of forecast error covariances through singular vectors. *J. Atmos. Sci.,* under review.

Epstein, E.S., 1969: Stochastic dynamic prediction. *Tellus,* **21**, 739–759.

Errico, R.M., 1991: Theory and application of nonlinear normal mode initialization. NCAR Tech. Note, NCAR/TN–344+IA. 137 pp. [Available from the National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307, U.S.A.]

Errico, R.M., and D.P. Baumhefner, 1987: Predictability experiments using a high–resolution limited–area model. *Mon. Wea. Rev.,* **115**, 408–504.

Errico, R.M., and M. Ehrendorfer, 1995: Moist singular vectors in a primitive-equation regional model. *Preprint Volume, Tenth Conference on Atmospheric and Oceanic Waves and Stability,* Big Sky, Montana, 235–238.

Errico, R.M., K. Raeder, and T. Vukićević, 1994: *Mesoscale Adjoint Modeling System Version 1.* NCAR Tech. Note, NCAR/TN–410+IA. 214 pp. [Available from the National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307, U.S.A.]

Errico, R.M., T. Vukićević, and K. Raeder, 1993: Examination of the accuracy of a tangent linear model. *Tellus,* **45A**, 462–477.

Farrell, B.F., 1988: Optimal excitation of neutral Rossby waves. *J. Atmos. Sci.* **45**, 163–172.

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

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

Farrell, B.F., and P.J. Ioannou, 1993a: Stochastic forcing of the linearized Navier–Stokes equations. *Phys. Fluids A,* **5**, 2600–2609.

Farrell, B.F., and P.J. Ioannou, 1993b: Stochastic dynamics of baroclinic waves. *J. Atmos. Sci.,* **50**, 4044–4057.

Farrell, B.F., and A.M. Moore, 1992: An adjoint method for obtaining the most rapidly growing perturbation to oceanic flows. *J. Phys. Oceanogr.,* **22**, 338–349.

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

Grimes, R.G., J.G. Lewis, and H.D. Simon, 1994: A shifted block Lanczos algorithm for solving sparse symmetric generalized eigenproblems. *SIAM J. Matrix Analysis and Applications,* **15**, 228–272.

Hack, J.J., 1994: Parameterization of moist convection in the National Center for Atmospheric Research community climate model (CCM2). *J. Geophys. Res.,* **99**, 5551–5568.

Houtekamer, P.L., 1995: The construction of optimal perturbations. *Mon. Wea. Rev.,* **123**, 2888–2898.

Houtekamer, P.L., and J. Derome, 1995: Methods for ensemble prediction. *Mon. Wea. Rev.,* **123**, 2181–2196.

Järvinen, H., J.-N. Thépaut, and P. Courtier, 1996: Quasi–continuous variational data assimilation. *Quart. J. Roy. Meteor. Soc.,* **122**, 515–534.

Kumar, A., M. Hoerling, M. Ji, A. Leetmaa, and P. Sardeshmukh, 1996: Assessing a GCM's suitablity for making seasonal predictions. *J. Climate,* **9**, 115–129.

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

Lanczos, C., 1950: An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Nat. Bureau Standards,* **45**, 255–281.

Leith, C.E., 1971: Atmospheric predictability and two–dimensional turbulence. *J. Atmos. Sci.,* **28**, 145–161.

Leith, C.E., 1978: Objective methods for weather prediction. *Annual Review of Fluid Mechanics,* **10**, 107–128.

Leith, C.E., 1983: Predictability in theory and practice. *Large–scale dynamical processes in the atmosphere.* B.J. Hoskins and R.P. Pearce, Eds., Academic Press, 365–383.

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

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

Lorenz, E.N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus,* **21**, 289–307.

Lorenz, E.N., 1982: Atmospheric predictability experiments with a large numerical model. *Tellus,* **34**, 505–513.

Lorenz, E.N., 1990: Effects of analysis and model errors on routine weather forecasts. *Ten years of medium–range weather forecasting,* European Centre for Medium–Range Weather Forecasts 1989 Seminar Proceedings, 115–128.

Molteni, F., and T.N. Palmer, 1993: Predictability and finite-time instability of the northern winter circulation. *Quart. J. Roy. Meteor. Soc.* **119**, 269–298.

Molteni, F., R. Buizza, T.N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: methodology and validation. *Quart. J. Roy. Meteor. Soc.,* **122**, 73–119.

Moorthi, S., and M.J. Suarez, 1992: Relaxed Arakawa–Schubert: a parameterization of moist convection for general circulation models. *Mon. Wea. Rev.,* **120**, 978–1002.

Mukougawa, H., M. Kimoto, and S. Yoden, 1991: A relationship between local error growth and quasi–stationary states: case study in the Lorenz system. *J. Atmos. Sci.,* **48**, 1231–1237.

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

Nicolis, C., S. Vannitsem, and J.-F. Royer, 1995: Short–range predictability of the atmosphere: mechanisms for superexponential error growth. *Quart. J. Roy. Meteor. Soc.,* **121**, 705–722.

Palmer, T.N., 1993: Extended–range atmospheric prediction and the Lorenz model. *Bull. Amer. Meteor. Soc.,* **74**, 49–65.

Palmer, T.N., and D.L.T. Anderson, 1994: The prospects for seasonal forecasting – a review paper. *Quart. J. Roy. Meteor. Soc.,* **120**, 755-793.

Penland, C., and P. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J. Climate,* **8**, 1999–2024.

Rabier, F., E. Klinker, P. Courtier, and A. Hollingsworth, 1996: Sensitivity of forecast errors to initial conditions. *Quart. J. Roy. Meteor. Soc.,* **122**, 121–150.

Shukla, J., 1985: Predictability. *Issues in atmospheric and oceanic modeling, Part B, Advances in Geophysics,* **28**, B. Saltzman and S. Manabe, Eds., Academic Press, 87–122.

Simmons, A.J., R. Mureau, and T. Petroliagis, 1995: Error growth and estimates of predictability from the ECMWF forecasting system. *Quart. J. Roy. Meteor. Soc.,* **121**, 1739–1771.

Strang, G., 1986: *Introduction to applied mathematics.* Wellesley–Cambridge Press, 758 pp.

Tennekes, H., 1978: Turbulent flow in two and three dimensions. *Bull. Amer. Meteor. Soc.,* **59**, 22–28.

Thépaut, J.-N., R.N. Hoffman, and P. Courtier, 1993: Interactions of dynamics and observations in four–dimensional variational assimilation. *Mon. Wea. Rev.,* **121**, 3393–3414.

Thépaut, J.-N., P. Courtier, G. Belaud, and G. Lemaître, 1996: Dynamical structure functions in a four–dimensional variational assimilation: a case study. *Quart. J. Roy. Meteor. Soc.,* **122**, 535–561.

Thompson, P.D., 1985: A review of the predictability problem. *Predictability of fluid motions,* G. Holloway and B.J. West, Eds., American Institute of Physics, 1–10.

Thompson, P.D., 1986: A simple approximate method of stochastic–dynamic prediction for small initial errors and short range. *Mon. Wea. Rev.,* **114**, 1709–1715.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: the generation of perturbations. *Bull. Amer. Meteor. Soc.,* **74**, 2317–2330.

Trevisan, A., and R. Legnani, 1995: Transient error growth and local predictability: a study in the Lorenz system. *Tellus,* **47A**, 103–117.

Tribbia, J.J., 1996: Weather prediction. *Economic value of weather and climate forecasts,* R.W. Katz and A.H. Murphy, Eds., Cambridge University Press, in press.

Tribbia, J.J., and R.A. Anthes, 1987: Scientific basis of modern weather prediction. *Science,* **237**, 493–499.

Van Tuyl, A.H., and R.M. Errico, 1989: Scale interaction and predictability in a mesoscale model. *Mon. Wea. Rev.,* **117**, 495–517.

Vukićević, T., 1991: Nonlinear and linear evolution of initial forecast errors. *Mon. Wea. Rev.,* **119**, 1602–1611.

Vukićević, T., and R.M. Errico, 1990: The influence of artificial and physical factors upon predictability estimates using a complex limited–area model. *Mon. Wea. Rev.,* **118**, 1460–1482.

Yoden, S., and M. Nomura, 1993: Finite-time Lyapunov stability analysis and its applications to atmospheric predictability. *J. Atmos. Sci.,* **50**, 1531–1543.