An Adjoint Sensitivity Study of Blocking in a Two-Layer Isentropic Model

X. ZOU
Supercomputer Computations Research Institute, The Florida State University, Tallahassee, Florida

A. BARCILON
Geophysical Fluid Dynamics Institute and Department of Meteorology, The Florida State University, Tallahassee, Florida

I. M. NAVON
Department of Mathematics and Supercomputer Computations Research Institute, The Florida State University, Tallahassee, Florida

J. WHITAKER
CIRES, University of Colorado, Boulder, Colorado

D. G. CACUCI*
Department of Nuclear Engineering, University of Illinois, Urbana, Illinois

(Manuscript received 9 December 1992, in final form 1 April 1993)

ABSTRACT

This paper presents a new methodology for adjoint sensitivity analysis, previously developed in general terms by Cacuci, into a form directly applicable to meteorological problems. This technique is illustrated by examining the sensitivity of a blocking index in a two-layer isentropic model. The index represents a response function for the sensitivity analysis that, unlike previous meteorological applications, is an operator and not a functional, and thus, extends the scope of adjoint sensitivity to general operator-type responses depending on time and/or space.

The sensitivity of the blocking index to perturbations introduced into the model atmosphere, as well as to model parameters, is discussed. The methodology of generalized adjoint sensitivity analysis described in this paper constitutes a prototype for further applications in the atmospheric and/or oceanic sciences.

1. Introduction

The use of adjoint functions for sensitivity analysis dates as early as the 1940s, and has evolved from the perturbation theory work of Wigner (1953) or the variational approaches of Levine and Schwinger (1949) and Roussololos (1953). The first use of adjoint functions for sensitivity analysis of a simple model in atmospheric sciences is due to Marchuk (1974), who applied established techniques already used routinely in reactor physics. All of these applications involved linear models and responses that were functionals (as opposed to general operators) of the model's parameters and dependent (i.e., state) variables.

In 1981, a general sensitivity theory for nonlinear systems was formulated by Cacuci (1981a), comprising two alternative formalisms: the "forward sensitivity formalism" and the "adjoint sensitivity formalism." The forward sensitivity formalism is expressed in a general linear vector space and can be used, in a conceptually straightforward way, to assess the effects of few parameter changes on many responses. For problems involving many parameter alternations or a large database and comparatively few responses, the adjoint sensitivity formalism is computationally far more economical, since the sensitivity of one response to all the model parameters and the model state at previous time can be evaluated in terms of a single adjoint solution. In contrast to the forward sensitivity formalism, the form of the adjoint equations remains unchanged regardless of the parameters considered. Furthermore, the adjoint functions provide a quantitative measure of the importance of data or region in phase space in contributing to the response under consideration. This sensitivity theory has been applied by Cacuci and his coworkers to nonlinear problems in several different fields [see Cacuci (1988) for a review], including a

* Current affiliation: Institut fur Reaktorsiecherheit Kernforschungzentrum, Karlsruhe, Germany.

Corresponding author address: Dr. I. Michael Navon, The Florida State University, SCRI, B-186, 415 Science Library, Tallahassee, FL 32306.

comprehensive sensitivity analysis for a radiative–convective model of the climate (Hall et al. 1982).

Cacuci (1981b) has also extended the scope of this theory to general operator-type responses, such as time- and/or space-dependent functions of the model's dependent (state) variables and parameters. The first application of this theory to a time-dependent response in a nonlinear heat transfer problem has been reported by Cacuci et al. (1980). A review of sample applications of this comprehensive sensitivity theory has been presented by Cacuci (1988).

The first application to a nonlinear problem in atmospheric research of this adjoint sensitivity theory was presented by Hall et al. (1982). Subsequently, Hall and Cacuci (1983) have shown that the adjoint functions in a radiative convective model quantify the importance of previous (antecedent) states to the current response functional. In a similar vein, Errico and Vukicevic (1992) indicated that the adjoint fields quantify the antecedent conditions that most affect a specified forecast aspect.

Adjoints of atmospheric and oceanic models are used for variational data assimilation (LeDimet and Talagrand 1986; Thépaut and Courtier 1992; Navon et al. 1992, to mention but a few), for optimal parameter estimation (Smestad and O'Brien 1991; Zou et al. 1992), as well as for the evaluation of optimal growth rates of initial perturbations (Farrell 1990; Tribbia 1991; Barkmeijer 1991).

The variability in predictive skills of numerical weather prediction (NWP) models is strongly related to the occurrence of blocks (Kimoto et al. 1992; Tracton et al. 1989), and it is therefore important to understand the model errors associated with blocking situations. The research reported in this paper focuses on the methodology of adjoint sensitivity analysis and uses the sensitivity of a blocking index to the source of model error as an illustrative example. This novel application to a meteorological problem represents a qualitatively significant generalization of previous applications of the adjoint sensitivity method to geophysical sciences, and will hopefully open the way for additional research.

In this study we use a two-layer, isentropic primitive equation model on a hemisphere with wavenumber 2 bottom topography and simple parameterizations of radiative heating and surface drag. This model is described in detail in appendix A. In section 2, we cast the spectral model equations into operator form and present the mathematical definition of sensitivity, the derivation of the adjoint equation model, and the expressions of sensitivities in terms of these adjoint functions. Three different types of response—functionals, operators depending on time, and operators depending on both time and space—are described. In section 3 we use the adjoint technique to study the sensitivity of a blocking index at a fixed longitude so that index becomes, for simplicity, a function of time only. Summary and conclusions are presented in section 4, which discusses further application of the adjoint method to the sensitivity analysis of more sophisticated NWP models.

2. Sensitivity analysis

The two-layer isentropic model simulates the nonlinear life cycles of baroclinic waves quite realistically, and the occurrence of blocks is observed (Figs. 1–3). The model equations can be written in operator form as [see (A.23)–(A.25) in appendix A for details]

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(t; \mathbf{x}, \alpha), \tag{2.1}$$

where $\mathbf{x}$ is a state vector of dimension $6 \times$ nmdim, where nmdim $= [J(J-1)+2]/2$—which denotes all the spectral coefficients of vorticity ζ, divergence D, and perturbation layer thickness $\Delta\pi'$—J is the total zonal wavenumbers, the operator $\mathbf{F}$ represents all processes that change the model state $\mathbf{x}$, and α is the model parameter vector.

In sensitivity analysis studies, the results of interest that are calculated using the model are usually referred to as the system's response. In this section we follow Cacuci (1981a,b) to outline the mathematical foundation of adjoint sensitivity analysis for responses that are either functionals (i.e., scalar-valued operators) or time- and/or space-dependent operators of the model's dependent variables and parameters.

a. *System response: Functional*

The specific response $R(\mathbf{x}, \alpha)$ considered in this subsection is a functional of $\mathbf{x}$ and α of the form

$$R(\mathbf{x}, \alpha) = \int_{t_0}^{t_a} r(t; \mathbf{x}, \alpha)dt, \tag{2.2}$$

where $r(t; \mathbf{x}, \alpha)$ depends on model variables $\mathbf{x}$ and the parameters α, and the time interval $[t_0, t_a]$ represents the selected time window, where $(t_a - t_0)$ is the time interval of most interest.

The most general definition of the sensitivity of a response to variations in the system parameters is the Gâteau (G) differential (see Cacuci 1981a). The G differential $VR(\mathbf{x}^0, \alpha^0, \mathbf{h}_x, \mathbf{h}_\alpha)$ of $R(\mathbf{x}, \alpha)$ at the nominal values $(\mathbf{x}^0, \alpha^0)$ for increments $(\mathbf{h}_x, \mathbf{h}_\alpha)$ around $(\mathbf{x}^0, \alpha^0)$ is given by

$$VR(\mathbf{x}^0, \alpha^0; \mathbf{h}_x, \mathbf{h}_\alpha) = \int_{t_0}^{t_a} \mathbf{r}'_x \cdot \mathbf{h}_x dt + \int_{t_0}^{t_a} \mathbf{r}'_\alpha \cdot \mathbf{h}_\alpha dt, \tag{2.3}$$

where

$$r'_x \equiv \left[\left(\frac{\partial r}{\partial x_1}, \cdots, \frac{\partial r}{\partial x_P}\right)\right]_{(\mathbf{x}^0, \alpha^0)}, \tag{2.4}$$

$$r'_\alpha \equiv \left[\left(\frac{\partial r}{\partial \alpha_1}, \cdots, \frac{\partial r}{\partial \alpha_N}\right)\right]_{(\mathbf{x}^0, \alpha^0)}; \quad (2.5)$$

N is the dimension of a vector of model parameters; P is the dimension of the model variable $\mathbf{x}$; $P = 6 \times$ nmdim, if $\mathbf{x}$ represents the spectral coefficients of the model variables; and $P = 6 \times \text{NLONS} \times \text{NLATSH}$, if $\mathbf{x}$ represents the model variables on Gaussian grid, where NLONS and NLATSH are the total number of grid points along longitude and latitude.

When $R(\mathbf{x}, \alpha)$ is *continuous* in $\mathbf{x}$ and α, the total variation of R is given by

$$R(\mathbf{x}^0 + \mathbf{h}_x, \alpha^0 + \mathbf{h}_\alpha) - R(\mathbf{x}^0, \alpha^0) = VR(\mathbf{x}^0, \alpha^0; \mathbf{h}_x, \mathbf{h}_\alpha) + O(\|\mathbf{h}_x\|^2) + O(\|\mathbf{h}_\alpha\|^2); \quad (2.6)$$

where, $VR(\mathbf{x}^0, \alpha^0; \mathbf{h}_x, \mathbf{h}_\alpha)$ is linear in $\mathbf{h}_x$ and $\mathbf{h}_\alpha$. If R or its derivatives are discontinuous, the G differential still has meaning as a general functional, i.e. a distribution.

The remaining problem is to find an efficient way of calculating VR as defined in (2.3). In (2.3) the second term is called the "direct effect," which can be calculated directly without the use of the adjoint model. The first term is the "indirect effect" and its calculation requires knowledge of $\mathbf{h}_x(t)$.

Taking the G differential of (2.1), we obtain the linear system

$$\mathbf{L}[\mathbf{x}^0(t), \alpha^0]\mathbf{h}_x(t) = \mathbf{Q}[x^0(t), \alpha^0]\mathbf{h}_\alpha, \quad (2.7a)$$

$$\mathbf{h}_{x|t=t_0} = \mathbf{h}_x(t_0), \quad (2.7b)$$

where $\mathbf{L}$ and $\mathbf{Q}$ are constant matrices defined as

$$\mathbf{L}[\mathbf{x}^0(t), \alpha^0] = \frac{d}{dt}\mathbf{I} - \frac{\partial \mathbf{F}}{\partial \mathbf{x}}, \quad (2.8)$$

$$\mathbf{Q}[\mathbf{x}^0(t), \alpha^0] = \frac{\partial \mathbf{F}}{\partial \alpha}, \quad (2.9)$$

and $\mathbf{I}$ is the respective unit matrix. The system of (2.7) is sometimes referred to as the tangent linear system, and $\mathbf{h}_x(t)$ may be obtained by solving (2.7). However, when the dimension of the initial state vector and the number of parameters are large, the computational cost of calculating this so-called indirect effect is very high. Therefore, we eliminate $\mathbf{h}_x(t)$ from (2.3) by using the adjoint formulation.

The adjoint operator $\mathbf{L}^*$ is defined through the relationship

$$\int_{t_0}^{t_a} \mathbf{h}_x \cdot (\mathbf{L}^*\mathbf{q})dt \equiv \int_{t_0}^{t_a} \mathbf{q} \cdot (\mathbf{L}\mathbf{h}_x)dt - (\mathbf{h}_x \cdot \mathbf{q})_{t_0}^{t_a}, \quad (2.10)$$

where $\mathbf{q}$ is at this stage an arbitrary column vector of dimension P.

Defining the adjoint model as

$$\mathbf{L}^*\mathbf{q} = \mathbf{r}'_\mathbf{x}, \quad (2.11a)$$

$$\mathbf{q}(t_a) = 0, \quad (2.11b)$$

where the inhomogeneous term $\mathbf{r}'_\mathbf{x}$ is given by (2.4), and using (2.11), we write equation (2.10) as

$$\int_{t_0}^{t_a} \mathbf{r}'_\mathbf{x} \cdot \mathbf{h}_x dt = \int_{t_0}^{t_a} \mathbf{q} \cdot (\mathbf{L}\mathbf{h}_x)dt + \mathbf{h}_x(t_0) \cdot \mathbf{q}(t_0). \quad (2.12)$$

Substituting (2.7) into (2.12), we obtain

$$\int_{t_0}^{t_a} \mathbf{r}'_\mathbf{x} \cdot \mathbf{h}_x dt = \int_{t_0}^{t_a} \mathbf{q} \cdot (\mathbf{Q}\mathbf{h}_\alpha)dt + \mathbf{h}_x(t_0) \cdot \mathbf{q}(t_0). \quad (2.13)$$

With the use of (2.13), (2.3) can be written as

$$VR = \int_{t_0}^{t_a} \mathbf{r}'_\alpha \cdot \mathbf{h}_\alpha dt + \int_{t_0}^{t_a} \mathbf{q} \cdot (\mathbf{Q}\mathbf{h}_\alpha)dt + \mathbf{h}_x(t_0)\mathbf{q}(t_0), \quad (2.14)$$

which is the adjoint formulation for sensitivity analysis of a response functional (2.2).

The main advantage of the adjoint formulation is that (2.14) is independent of $\mathbf{h}_x(t)$. Thus, (2.14) replaces the time integration of the model consisting of different equations for $\mathbf{h}_x$ [which would have been required to evaluate (2.3)] with the calculation of a quadrature, an operation much cheaper to perform when the number of the model parameters is large. The adjoint variable $\mathbf{q}(t)$ is the solution of the adjoint equations (2.11), which are *independent* of $\mathbf{h}_x(t)$ and $\mathbf{h}_\alpha$. Therefore, a *single* adjoint model calculation suffices to obtain the sensitivities to variations of all the model parameters. However, the forcing term, $\mathbf{r}'_\mathbf{x}$, in the adjoint model depends on the functional defining the response, so that for each response the adjoint model equations must be integrated anew.

b. System response: Operator that depends on time

In section 2a, we considered a system response mapping the domain $E = E_x \times E_\alpha$ into the underlying space Λ of real scalars. However, in many practical problems, the system response may be time and/or space dependent, and is therefore a more general operator, rather than a functional, whose range is not in Λ but in some other normed linear space E_R. In the following we shall specifically consider a time-dependent operator to be the response R. This application of sensitivity analysis of a time-dependent response is novel in meteorology.

For a general derivation of the sensitivity of a time-dependent response, we note that

$$R(t) = R(\mathbf{x}(t), \alpha), \quad (2.15)$$

where $\mathbf{x}$ represents a vector of dimension P and consists of all the spectral coefficients of the model dependent variables, and α represents a vector of all the model-parameters with dimension M. We are interested in the behavior of $R(t)$ over a finite window of time $t_0 \leq t \leq t_a$.

As before, the sensitivity of $R(t)$ to changes $\mathbf{h}_\alpha$ in the parameters and changes $\mathbf{h}_x$ in the model solution is defined as the Gâteau differential of R at (x^0, α^0):

$$VR(\mathbf{x}^0(t), \alpha^0, \mathbf{h}_x(t), \mathbf{h}_\alpha) = R'_x(\mathbf{x}^0(t), \alpha^0)\mathbf{h}_x(t) + R'_\alpha(\mathbf{x}^0(t), \alpha^0)\mathbf{h}_\alpha. \quad (2.16)$$

The indirect effect [the first term in (2.16)] now is a function of time instead of a scalar, as in (2.3). However, the adjoint operator can be introduced only through a *scalar* in E_x that is an inner product of $\mathbf{h}_x$ with some other vector. Therefore, we restrict E_x, E_d, and E_R to be Hilbert spaces with inner products denoted by $\{,\}$; $[,]$; and $\langle,\rangle$, respectively, and proceed by expanding the term $R'_x(x^0(t), \alpha^0)\mathbf{h}_x(t)$ in terms of a truncated infinite-dimensional orthogonal basis $w_m(t)$, $m = 1, 2\cdots, mm$, such that

$$R'_x(\mathbf{x}^0(t), \alpha^0)\mathbf{h}_x(t) = \sum_{m=1}^{mm} a_m w_m(t), \quad (2.17)$$

$$a_m = \langle R'_x(\mathbf{x}^0(t), \alpha^0)\mathbf{h}_x(t), w_m(t)\rangle. \quad (2.18)$$

Since a_m are functionals of $\mathbf{h}_x(t)$ (i.e., scalar-valued quantities), they may be rewritten as

$$\langle R'_x(\mathbf{x}^0(t), \alpha^0)\mathbf{h}_x(t), w_m(t)\rangle = \mathbf{h}_x(t), M(\mathbf{x}^0(t), \alpha^0)w_m(t), \quad (2.19)$$

where $M = (R'_x)^{\mathrm{T}}$ is the adjoint of R'_x and the inner product defined in the E_x space is

$$\mathbf{h}, \mathbf{g} = \sum_{i=1}^{P} \langle h_i, g_i\rangle \quad (2.20)$$

for any two P-dimensional vectors $\mathbf{h}$ and $\mathbf{g}$.

Introducing the mm-adjoint systems,

$$\mathbf{L}^*(\mathbf{x}^0(t), \alpha^0)\mathbf{q}_m(t) = Mw_m(t), \quad (2.21a)$$

$$\mathbf{q}_m(t_a) = 0, \quad m = 1, \cdots, mm, \quad (2.21b)$$

and using (2.7), (2.21), we obtain

$$\begin{aligned}\mathbf{h}_x(t), M(\mathbf{x}^0(t), \alpha^0)w_m(t) &= [\mathbf{h}_x(t), \mathbf{L}^*(\mathbf{x}^0(t), \alpha^0)\mathbf{q}_m(t)] \\ &= [\mathbf{q}_m(t), \mathbf{L}(\mathbf{x}^0(t), \alpha^0)\mathbf{h}_x(t)] + [\mathbf{h}_x(0), \mathbf{q}_m(0)] \\ &= [\mathbf{q}_m(t), \mathbf{Q}(\mathbf{x}^0(t), \alpha^0)\mathbf{h}_\alpha] + [\mathbf{h}_x(0), \mathbf{q}_m(0)]. \quad (2.22)\end{aligned}$$

Substituting (2.21) into (2.17) for the indirect effect term, the sensitivity (2.16) can now be written as

$$\begin{aligned}VR(\mathbf{x}^0(t), \alpha^0; \mathbf{h}_x(t), \mathbf{h}_\alpha) &= \sum_{m=1}^{mm} [\mathbf{q}_m(t), \mathbf{Q}(\mathbf{x}^0(t), \alpha^0)\mathbf{h}_\alpha]w_m(t) \\ &\quad + [\mathbf{h}_x(0), \mathbf{q}_m(0)]w_m(t) \\ &\quad + \{R'_\alpha(\mathbf{x}^0(t), \alpha^0)\mathbf{h}_\alpha\}. \quad (2.23)\end{aligned}$$

We observe, from (2.21), that the calculation of the sensitivity of a certain response in the adjoint formulation involves the integration of the adjoint model mm times [mm being in practice a small number that depends on the level of truncation in (2.17)]. The computational effort in sensitivity calculations based on the adjoint approach is as follows:

(i) develop the adjoint model, which depends on the nominal values of the initial conditions and parameters;

(ii) calculate the matrix $\mathbf{Q}$, which also depends on the nominal values of the initial conditions and parameters;

(iii) find the forcing term $Mw_m(t)$, which depends on the definition of the response; and

(iv) calculate quadratures $[\mathbf{q}_m, \mathbf{Q}\mathbf{h}_\alpha]mm$ times and perform the summation shown in (2.23).

By comparing (2.16) and (2.23) we observe that $\mathbf{h}_x(t)$ no longer appears; instead, the adjoint variables $\mathbf{q}_m$ appear in the expression (2.23) giving the sensitivity VR. The adjoint variables $\mathbf{q}_m$ depend only on the basic state $\mathbf{x}^0(t)$, α^0 and the response function (for the forcing term), and are independent of changes in the parameters and initial conditions, while in (2.16), $\mathbf{h}_x(t)$ is *the* solution of the linear differential equation (2.7), which depends on the changes in the parameters and the initial conditions. For each parameter change or model-state change at time t_0, the linear differential system (2.7) would have to be solved anew, an expensive task when the dimensions of the parameter and the model-state vectors are large.

c. *System response: Operator that depends on both time and space*

Possibly the most interesting meteorological cases involve operator responses that depend on both time and space. We present below the treatment of such a response: $R(\lambda, \phi, t)$. Expanding $R(\lambda, \phi, t)$ into spherical harmonics, or any other basis $(e_n(\lambda, \phi))$, we obtain

$$R(\lambda, \phi, t) = \sum_{n=1}^{N} R_n(t)e_n(\lambda, \phi), \quad (2.24)$$

where N represents the total number of the expansion coefficients. Taking the G differential of R yields

$$VR(\mathbf{x}^0(t), \alpha^0, \mathbf{h}_x(t), \mathbf{h}_\alpha) = \sum_{n=1}^{N} (R'_x)_n\mathbf{h}_x e_n(\lambda, \phi) + \sum_{n=1}^{N} (R'_\alpha)_n h_\alpha e_n(\lambda, \phi), \quad (2.25)$$

where $(R'_x)_n$ and $(R'_\alpha)_n$ are P-dimensional vectors that depend on time.

Using an orthogonal basis expansion $w_m(t)$, $m = 1, 2\cdots, mm$, expand $(R'_x)_n\mathbf{h}_x$ as before, in a truncated series to obtain

$$(R'_x)_n\mathbf{h}_x(t) = \sum_{m=1}^{mm} a_m^n w_m^n(t), \quad (2.26)$$

with

$$a_m^n = \langle (R'_x)_n(t)\mathbf{h}_x(t),\, w_m^n(t)\rangle. \quad (2.27)$$

Since a_m^n are functionals (i.e., scalar-valued quantities) of $\mathbf{h}_x(t)$, we may rewrite (2.27) as

$$\langle (R'_x)_n(t)\mathbf{h}_x(t),\, w_m^n(t)\rangle = \mathbf{h}_x(t),\, (R'_x)_n^{\mathrm{T}}(t)w_m^n(t). \quad (2.28)$$

Defining the $mm \times N$ adjoint system as

$$\mathbf{L}^*(\mathbf{x}^0(t),\, \alpha^0)\mathbf{q}_m^n(t) = (R'_x)_n^{\mathrm{T}}(t)w_m^n(t), \quad (2.29a)$$

$$\mathbf{q}_m^n(t_a) = 0, \quad (2.29b)$$

and proceeding along the same lines as those leading to (2.22) gives

$$\mathbf{h}_x(t),\, (R'_x)_n^{\mathrm{T}}(t)w_m^n(t) = [\mathbf{q}_m^n,\, \mathbf{Q}(\mathbf{x}^0(t),\, \alpha^0)\mathbf{h}_\alpha] + [\mathbf{h}_x(0),\, \mathbf{q}_m^n(0)]. \quad (2.30)$$

Finally, we note that the sensitivity VR for a response that is both a time- and space-dependent operator is given by the expression

$$VR(\mathbf{x}^0(t),\, \alpha^0;\, \mathbf{h}_x(t),\, \mathbf{h}_\alpha) = \sum_{n=1}^{N}\sum_{m} [\mathbf{q}_m^n(t),\, \mathbf{Q}(t)\mathbf{h}_\alpha]w_m^n(t) + [\mathbf{h}_x(0),\, \mathbf{q}_m^n(0)]w_m^n(t)e_n(\lambda,\, \phi) + [R'_\alpha\mathbf{h}_\alpha]. \quad (2.31)$$

We observe that the foregoing derivations are similar to those in section 2b except that one has to carry out a space expansion of the response function. The calculation of the sensitivity of a time- and space-dependent response in the adjoint-function sensitivity analysis method involves the integration of the adjoint model for $N \times mm$ times, where mm is the maximum truncation of the time expansion for the coefficient in (2.24), while N is the total number of the nonzero space expansion coefficients in (2.24). This could be computationally expensive. However, for a given application, even though the state variables $\mathbf{x}$ are expanded spatially in spherical harmonics $Y_n^l(\lambda, \phi)$, the indirect term $R'_x h_x$ need not be expanded in $Y_n^l(\lambda, \phi)$ if a more suitable basis is available. Furthermore, even if $R'_x h_x$ is expanded in $Y_n^l(\lambda, \phi)$, a few spectral terms may suffice. The choice of the optimum basis for expanding $R'_x h_x$ will depend on the shape of $R(\lambda, \phi, t)$. In fact, it usually turns out that expanding R or $R'_x h_x$ in the same basis functions as used to calculate the dependent variables $\mathbf{x}$ is a poor choice. The whole point in having used different inner products in Cacuci's (1981b) sensitivity theory (for the response, adjoint function, etc.) is precisely to enable us to use different basis functions when solving for the dependent variables $\mathbf{x}$, adjoint variables $\mathbf{q}$, and response sensitivities (i.e., $R'_x h_x$), respectively.

To decide what basis functions should be used for the $R'_x h_x$, we need to examine some representative plots of $R(\lambda, \phi, t)$, both as a function of t for some fixed and interesting values of λ and ϕ [see, for example, (3.3)], and as a function of λ and ϕ (separately and together) for some fixed and interesting values of t.

3. Numerical experiments

Although this work primarily focuses on the mathematical foundations of generalized sensitivity studies of certain classes of responses, in this section we offer (i) a meteorological illustration of the method (section 2b) applied to a blocking index obtained in the model previously discussed in section 2 and appendix A, and (ii) specific mathematical approximations for the study of the sensitivity of that blocking index. Again, we do

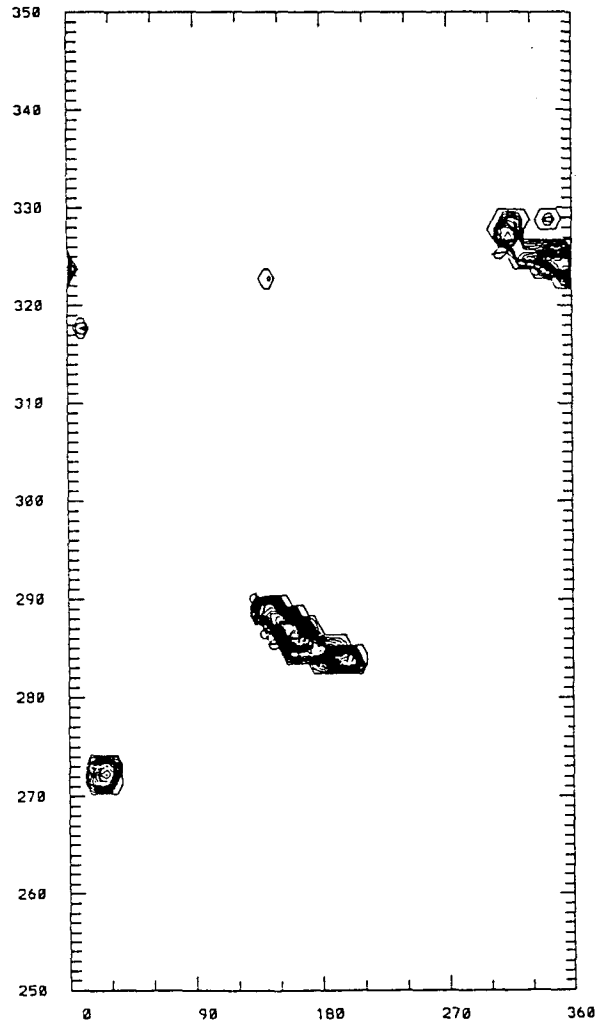

FIG. 1. Time (day 250 to 350) versus longitude (Hovmöller) plot of blocking index derived from full upper-layer streamfunction field. Contour interval is 2×10^6, with positive values suppressed.

(a)

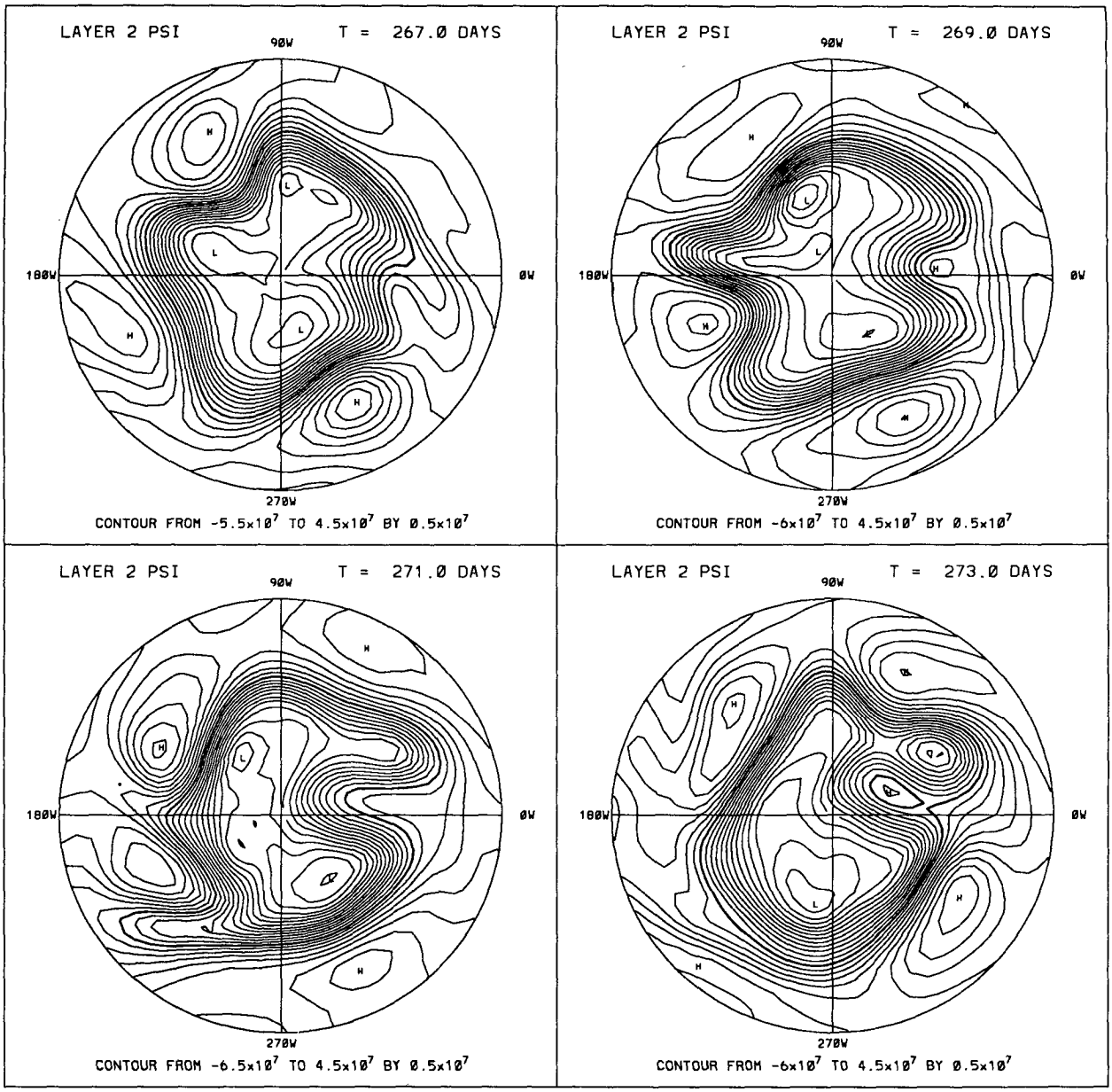

FIG. 2. The (a) upper- and (b) lower-layer full-field streamfunction distribution obtained on day 267, day 269, day 271, and day 273.

not view this section as a thorough analysis of the blocking dynamics but rather as an illustration of the previously discussed method.

The time integration of the model starts at time $t = 0$ from arbitrary initial values of the dependent variables, and the first 100 or so days are discarded. The integration proceeds until the final time t_a is reached. For the results presented in this paper, the total time span of the model is 1100 days. For the sensitivity analysis, a 32-day time window $[t_0, t_a]$ was used when a blocking event occurred.

a. Meteorological preliminaries

We define a blocking index analogous to the one introduced by Lejenäs and Okland (1983):

$$R(\lambda, t) = (\psi_{\phi_l}(\lambda, t) - \psi_{\phi_h}(\lambda, t)) \times H_-(\psi_{\phi_h}(\lambda, t) - \psi_{\phi_l}(\lambda, t)), \quad (3.1)$$

where

$$H_-(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases}, \quad (3.2)$$

(b)

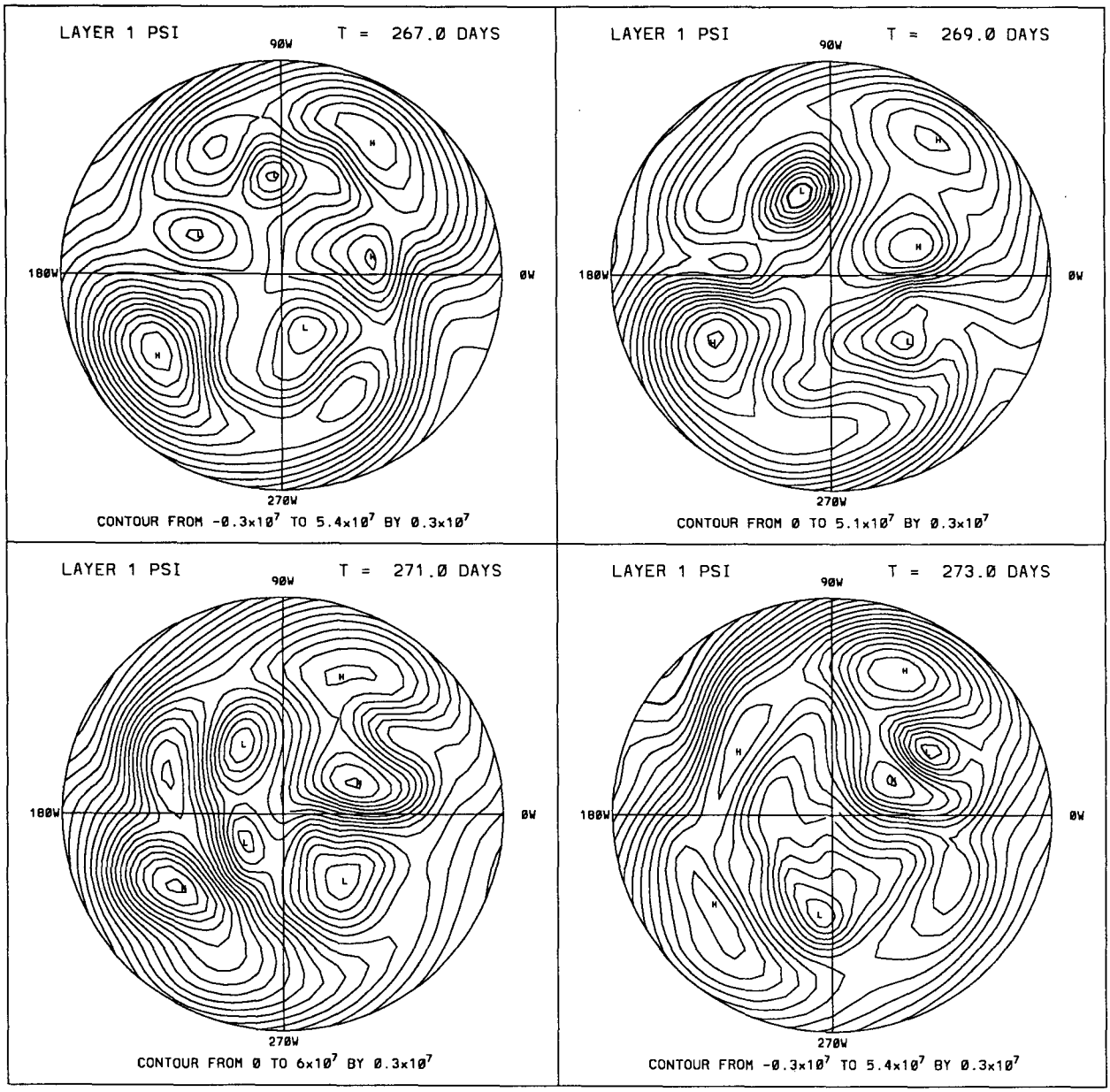

FIG. 2. (*Continued*)

is the unsymmetrical Heaviside function and $\psi_{\phi_{l,h}}$ is the upper-level streamfunction at latitudes ϕ_l, ϕ_h, which were empirically selected as the Gaussian latitudes 39° and 65°N. Figure 1 shows a Hovmöller plot of $R(t)$ in which λ is plotted horizontally and time increases upward from 250 to 350 model days. In that figure the reader should note how longitudes are measured; in particular, the zonal wavenumber 2 mountain ridges are located at $\lambda = 45°$ and 225°W. In Fig. 2, $\lambda = 0$ corresponds with the conventional positive "x axis" in the first quadrant. During the interval, day 250 to day 350, three model blocks form.

The formation of blocks is particularly difficult to predict (Tracton 1990; Tibaldi and Molteni 1990), and several theories have attempted to offer mechanisms for this process; no single theory is presently universally accepted. We do not claim to resolve the differences between blocking theories in this paper but rather present results that point at the power of the adjoint sensitivity method previously described.

From Figs. 1–3 we notice that although we make no claim that the topography chosen is in any way "earth-like," the coherent structures that develop resemble those observed in the atmosphere. We will discuss the

(a)

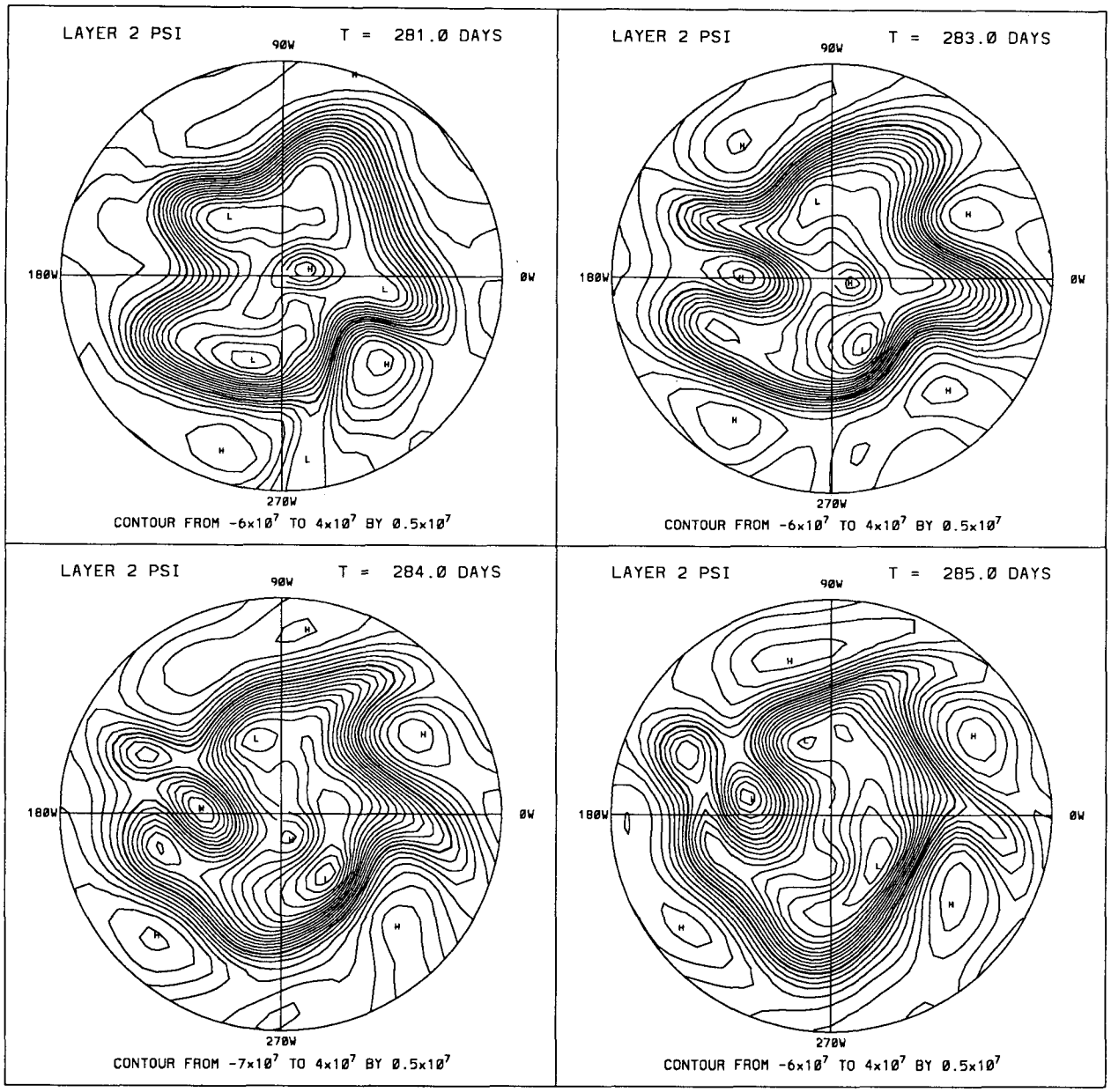

FIG. 3. The (a) upper- and (b) lower-layer full-field streamfunction distribution obtained on day 281, day 283, day 284, and day 285.

blocks that formed around day 272 (block 272) and around day 283 (block 283). Figures 2a,b and 3a,b show the streamfunction maps of synoptic situations in upper [panel (a)] and lower [panel (b)] layers just prior to the formation of these blocks. In these and similar figures the outmost latitude is 20°N, and at that latitude the flow is easterly. For block 272 (Fig. 2), in the lower layer, starting from day 269, the lee cyclone west of 0°W intensifies, while the zone of the anticyclone near the mountaintop propagates westward, resulting in a strong northward heat transport along 0°W. In the upper layer, the trough above that cyclone intensifies and the ridge narrows, extends northward while the downstream trough intensifies and extends southeastward. For block 283 (Fig. 3), for example, at day 281 the circulation over 90°–180°W (second quadrant) is zonal with no significant meanders. At day 283, a meridional elongation of the disturbance leads to the formation of the block over 150°–180°W (days 284 and 285), which will last for some 8 days. Visual inspections of the figures in the lower layer (Fig. 2b) show that the preblocking situation in-

(b)

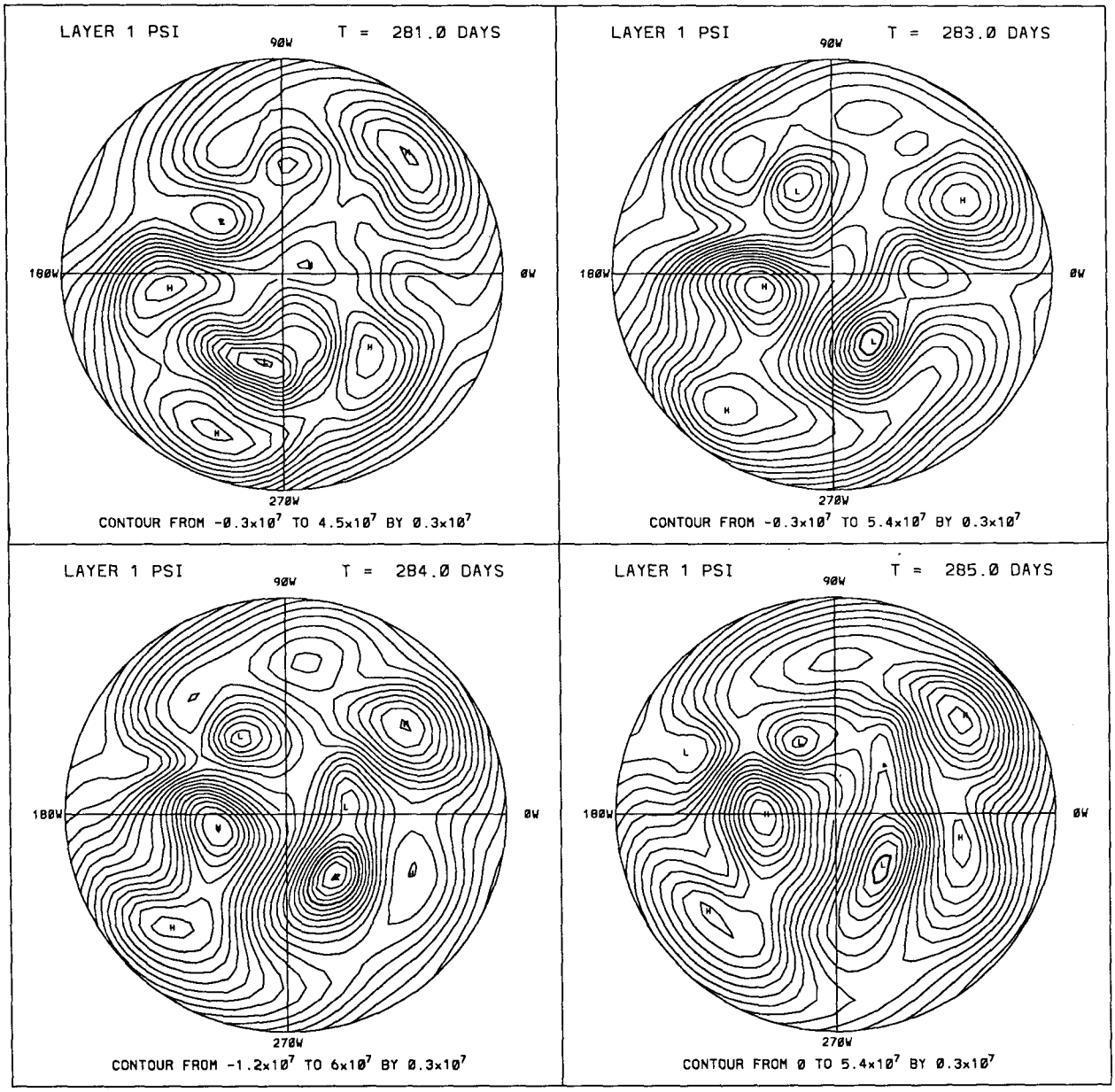

FIG. 3. (*Continued*)

volves intense thickness advection northward, this advection being due to the proximity, in the lower layer, of a high–low dipole. Thus, the initiation phase of blocking is highly baroclinic. Once formed, these blocks have a tendency to slowly retrograde and do not maintain the modon shape advocated by simple baroclinic models of blocking (Haines and Marshall 1987). Observations of streamfunction model fields in both layers indicate that a barotropic character prevails in mature and decaying blocks.

The index given by (3.1) is both a function of time and space. To simplify we select a fixed longitude λ for a given block and redefine the response as a function of time only as

$$R(t) = (\psi_{\phi_l}(t) - \psi_{\phi_h}(t))H_-(\psi_{\phi_h}(t) - \psi_{\phi_l}(t)), \quad (3.3)$$

where we select the constant longitude $\lambda_1 = 23°$W for block 272, $\lambda_1 = 150°$W for block 283, and $\lambda_1 = 360°$W for block 322. In the experiments to be described, a time window $t_0 \leq t \leq t_a$ will be selected, with t_0 being varied from a few to several days before blocking, and

t_a chosen past the duration time of blocking where $R(t_a) = 0$. For time intervals greater than the limit of validity of the tangent linear model, the use of the adjoint model can be questioned (Lacarra and Talagrand 1988; Rabier and Courtier 1992). Therefore, it is very important to select a reasonable time window for a sensitivity study of a specific problem.

As a representative measure of the blocking index we first select a time window extending from $t_0 = 267$ days to $t_a = 298$ days for block 283 and calculate the response once a day, shown as a solid line in Fig. 4. The finite Fourier sine expansion

$$R(t) = \frac{2}{t_a - t_0} \sum_{m=1}^{mm} a_m \sin \frac{\pi m(t - t_0)}{t_a - t_0} \quad (3.4)$$

is also shown in Fig. 4 as a dash-dotted line, for $mm = 8$, and as a dotted line, for $mm = 16$. The $mm = 16$ representation matches the response extremely well in the region where the computed response differs from zero, that is, between day $t_0 + 18$ and $t_0 + 22$. When the response is zero, the approximation given by (3.4) suffers from a Gibbs phenomena, and therefore, one should be careful in interpreting preblocking results that are tainted by this shortcoming. In what follows we selected the truncated basis function $w_m(t) = \sin[\pi m(t - t_0)/t_a - t_0]$ and $m = 1, \cdots, 16$. Since the response does not depend upon the model parameters, the direct effect $R'_\alpha \mathbf{h}_\alpha$ is zero and the indirect effect is

$$R'_x(x^0(t), \alpha^0)\mathbf{h}_x = (h_{\phi l} - h_{\phi h})H_-(\psi^0_{\phi h} - \psi^0_{\phi l}), \quad (3.5)$$

where we used $x\delta(x) = 0$, a property of the Dirac delta function.

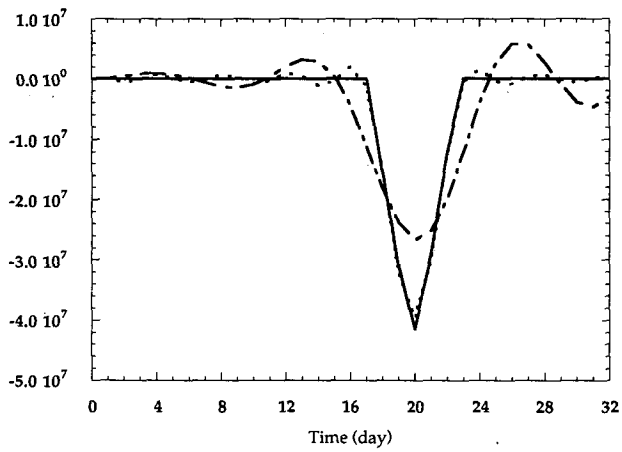

FIG. 4. Variation of the blocking index in time (solid line) and its approximation of the finite Fourier sine expansion truncated at $mm = 8$ (dash-dotted line) and $mm = 16$ (dotted line).

TABLE 1. Sensitivity of R to the increment $\epsilon\psi(t_0)$ in the field of streamfunction $\psi(t_0)$.

α	Predicted change: Adjoint method[a]	Predicted change: Direct method[b]	Actual change[c]
10^{-2}	4.683E17	4.683E17	−6.311E17
10^{-3}	4.683E16	4.683E16	4.721E16
10^{-4}	4.683E15	4.683E15	4.709E15
10^{-5}	4.683E14	4.683E14	4.686E14
10^{-6}	4.683E13	4.683E13	4.683E13

[a] Predicted change (adjoint method) = $[q(0), h_x(0)]$.
[b] Predicted change (direct method) = $[R'_x, h_x]$.
[c] Actual change = $R(x^0 + h_x, \alpha^0) - R(x^0, \alpha^0)$.

b. A simple test: The functional response

For code verification, before dealing with $R(t)$, we consider a simple functional response not related to blocking. For simplicity, we write

$$R(\mathbf{x}(t_0), \alpha) = \sum_{i,j,r} \psi^2(\lambda_i, \phi_j, k; t_r),$$

$$(k = 2, \text{ upper layer}), \quad (3.6)$$

where the summation is over longitude, latitude, and time, once a day for 32 days starting at $t_0 = 267$ days.

The sensitivity of the response to changes in model state $\mathbf{h}_x(t_0)$ and in parameters $\mathbf{h}_\alpha$, can be obtained in three different ways: 1) using the adjoint formulation

$$VR^{\text{adj}} = \sum_r q(t_r)\mathbf{Q}(\mathbf{x}^0(t_r), \alpha^0)\mathbf{h}_\alpha + \mathbf{h}_x(t_0)q(t_0), \quad (3.7)$$

2) using the direct forward integration

$$VR^{\text{dir}} = \sum_r 2\psi^0(t_r)\mathbf{h}_\psi(t_r), \quad (3.8)$$

and 3) using the total variation of the response

$$dR = R(\mathbf{x} + \mathbf{h}_x, \alpha + \mathbf{h}_\alpha) - R(\mathbf{x}, \alpha)$$

$$= \sum_r \psi^2(\mathbf{x}^0 + \mathbf{h}_x, \alpha^0 + \mathbf{h}_\alpha; t_r) - \psi^2(\mathbf{x}^0, \alpha^0; t_r). \quad (3.9)$$

We anticipate that (3.7) and (3.8) should produce, to within numerical accuracy, identical results but that (3.9) can be compared only with the previous two for sufficiently small changes; that is, in some sense $\|\mathbf{h}_x\|$ and $\|\mathbf{h}_\alpha\|$ must be small. In the foregoing expressions, q is the adjoint variable obtained by integrating the adjoint model (2.11) with forcing terms $r'_x = 2\psi^0(t_r)$

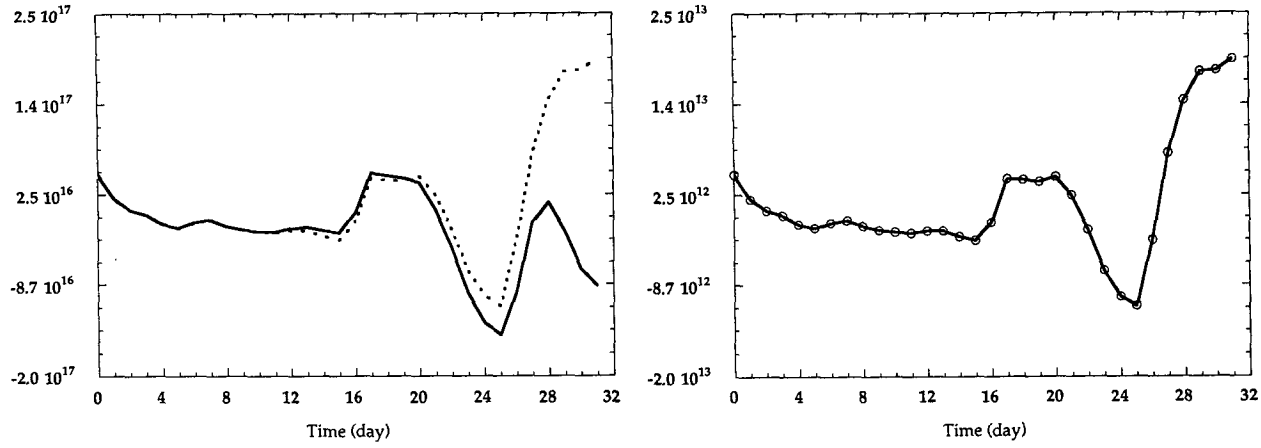

FIG. 5. Variations of $dR(t)$ (dotted line for $\epsilon = 10^{-2}$ and solid line for $\epsilon = 10^{-6}$) and $VR^{\text{dir}}(t)$ (solid line for $\epsilon = 10^{-2}$ and circles for $\epsilon = 10^{-6}$) with time in the time window of $[t_0, t_a]$ for variation of $\epsilon\psi(t_0)$ in the field of ψ at time t_0 for (a) $\epsilon = 10^{-2}$ and (b) $\epsilon = 10^{-6}$.

present at time t_r, $r = t_a, \cdots, t_0$, where $t_a = 298$ days and $t_0 = 267$ days; $\mathbf{h}_x$ is obtained by integrating the tangent linear model (2.7) directly.

Table 1 shows VR^{adj}, VR^{dir}, and dR for $\mathbf{h}_x = (\epsilon\psi^0(t_0), 0, 0)^{\mathrm{T}}$ and $\mathbf{h}_\alpha = 0$, where ϵ ranges from 10^{-2} to 10^{-6}. The results presented in the columns labeled "actual change" are the changes in the blocking index obtained by rerunning the nonlinear model after ψ is increased by $\epsilon\psi$. The columns "predicted change (direct)" are sensitivity values obtained by using VR^{dir}, while the columns "predicted change (adjoint)" display results due to VR^{adj}. The good agreement between the predicted changes obtained by the direct and adjoint methods supports the belief that the numerical method has been correctly applied to solve the adjoint system. For ϵ large ($\sim 10^{-2}$), the nonlinear dR results differ from those obtained using the linear method. Figure 5a plots, for $\epsilon = 10^{-2}$, the values of $2\psi^0(t_r)\mathbf{h}_\psi(t_r)$ and

$$\psi^2(\mathbf{x}^0 + \mathbf{h}_x, \alpha^0; t_r) - \psi^2(\mathbf{x}^0, \alpha^0; t_r). \quad (3.10)$$

The two curves are very close for the first 9 days of integration and begin diverging afterward, indicating that the cumulative effects of the nonlinear terms, for this value of ϵ, are no longer negligible. This result implies that the tangent linear model approximation is valid for 9 days of integration for this size ($\epsilon = 10^{-2}$) of perturbation. However, when $\epsilon = 10^{-6}$, the agreement between these same curves (Fig. 5b) is excellent, confirming results of Table 1 and verifying that the tangent linear model approximation holds. We feel confident that there are no gross errors in the code, and proceed to a more representative response.

c. *A time-dependent blocking index*

We now focus our attention on the index $R(t)$ as given by (3.3). We can view R as being dependent upon the vector $\mathbf{X}(t)$ in *physical* space or upon the vector of time-dependent *spectral* coefficients $\mathbf{x}$. We will consider the mathematical formulations from both points of view.

1) RESPONSE IN PHYSICAL SPACE

Expressing the response $R(t)$, (3.3), and the indirect effect $R'_X\mathbf{h}_X$, (3.5), as functions of the model variables in Gaussian grid space, we write

$$R(t) = \mathbf{a}^{\mathrm{T}} X H_-(-\mathbf{a}^{\mathrm{T}} X), \quad (3.11)$$

$$R'_X\mathbf{h}_X = \mathbf{a}^{\mathrm{T}}\mathbf{h}_X H_-(-\mathbf{a}^{\mathrm{T}} X^0), \quad (3.12)$$

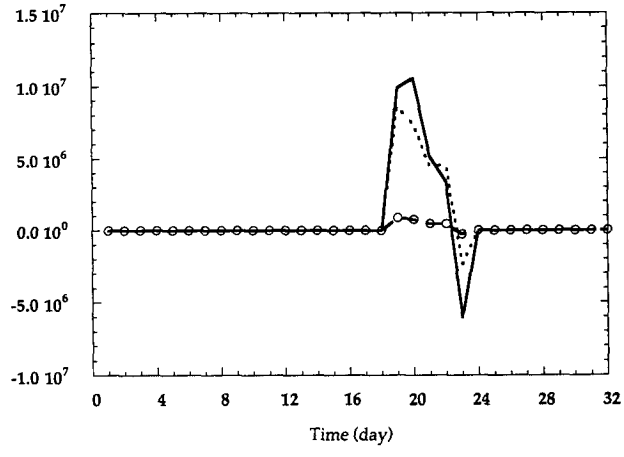

FIG. 6. The actual changes (solid line for $\epsilon = 0.01$ and dashed line for $\epsilon = 0.001$) and the predicted changes (dotted line for $\epsilon = 0.01$ and circle line for $\epsilon = 0.001$) by the direct method when ψ at time t_0 has a perturbation of $\epsilon\psi(t_0)$.

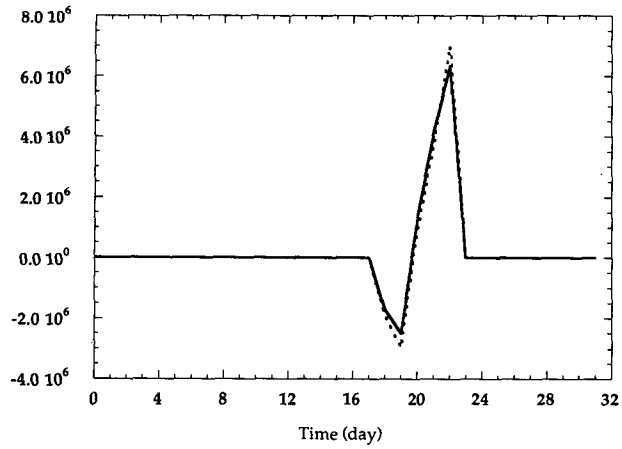

FIG. 7. The actual changes (solid line) and the predicted changes (dotted line) by the direct method for 1% changes in τ_{drag}.

where the vector $\mathbf{a}$ is defined as

$$a_l = \begin{cases} 1, & \text{if } l = (j1 - 1)I + i1; \\ -1, & \text{if } l = (j2 - 1)I + i1; \\ 0, & \text{otherwise,} \end{cases} \quad (3.13)$$

$j1, j2$, and $i1$ correspond to ϕ_l, ϕ_h, and λ_1, respectively, and I = NLONS is the total number of grid points in the longitudinal direction. It follows that the partial derivative of the response with respect to the model variables X is

$$R'_X = \left(\cdots, H_-(-\mathbf{a}^T X^0)\mathbf{a}^T \frac{\partial X}{\partial X_i}, \cdots \right). \quad (3.14)$$

Express the sensitivity of the response in terms of the basis function $\{\sin[\pi m(t - t_0)/t_a - t_0], m = 1, \cdots, mm\}$ and obtain

$$VR(X^0(t), \alpha^0, \mathbf{h}_X(t), \mathbf{h}_\alpha) = R'_X h_X = \frac{2}{t_a - t_0} \sum_{m=1}^{mm} a_m \sin \frac{\pi m(t - t_0)}{t_a - t_0}, \quad (3.15)$$

where

$$a_m = \int_{t_0}^{t_a} R'_X h_X \sin \frac{\pi m(t - t_0)}{t_a - t_0} dt \equiv \left\langle R'_X h_X, \sin \frac{\pi m(t - t_0)}{t_a - t_0} \right\rangle = \left[h_X, (R'_X)^T \sin \frac{\pi m(t - t_0)}{t_a - t_0} \right]. \quad (3.16)$$

Define the adjoint equation as

$$\mathbf{G}^*\mathbf{L}^*\mathbf{S}^* q_m(t) = (R'_X)^T \sin \frac{\pi m(t - t_0)}{t_a - t_0}, \quad (3.17)$$

where $\mathbf{G}$ is a transform from Gaussian grid space of $(\psi, D, \Delta\pi')$ to the spectral space of $(\zeta, D, \Delta\pi')$ and $\mathbf{S} = \mathbf{G}^{-1}$. Then, we can write

$$a_m = [h_X, \mathbf{G}^*\mathbf{L}^*\mathbf{S}^* q_m(t)] = [q_m(t)\mathbf{SLG}h_X(t)] + [h_X(t_0), q_m(t_0)]. \quad (3.18)$$

Using the G differential of the model equation,

$$\mathbf{SLG}h_X = \mathbf{SQ}h_\alpha, \quad (3.19)$$

a_m can be expressed as

$$a_m = [q_m, \mathbf{SQ}h_\alpha] + [h_X(t_0), q_m(t_0)]. \quad (3.20)$$

Substituting (3.20) into (3.15), we finally obtain

$$VR(X^0(t), \alpha^0, \mathbf{h}_X(t), \mathbf{h}_\alpha) = \frac{2}{t_a - t_0} \sum_{m=1}^{mm} \{[q_m, \mathbf{SQ}h_\alpha] + [h_X(t_0), q_m(t_0)]\} \sin \frac{\pi m(t - t_0)}{t_a - t_0}. \quad (3.21)$$

We run the nonlinear model once to obtain the value of response function (3.11) and the partial derivative of the response $(R'_X)^T$, (3.14). Then, we integrate the adjoint model backward $mm = 16$ times with different forcing terms $(R'_X)^T \sin[\pi m(t - t_0)/t_a - t_0]$ to obtain the values of the adjoint variables $q_m(t)$. Having obtained the values of all $q_m(t)$, $m = 1, \cdots, mm$, we use (3.21) to calculate the sensitivities of the response to model parameters and to model state at a certain time.

The inner product in (3.21) is

$$[\mathbf{g}^{(1)}(t), \mathbf{g}^{(2)}(t)] = \sum_{i=1}^{P} \langle g_i^{(1)}(t), g_i^{(2)}(t) \rangle = \sum_{i=1}^{P} \Delta t \left(\frac{1}{2} g_i^{(1)}(t_0) g_i^{(2)}(t_0) + \sum_{j=1}^{N-2} g_i^{(1)}(t_j) g_i^{(2)}(t_j) + \frac{1}{2} g_i^{(1)}(t_{N-1}) g_i^{(2)}(t_{N-1}) \right), \quad (3.22)$$

where $\mathbf{g}^{(1)}$ and $\mathbf{g}^{(2)}$ are two arbitrary vectors in physical space; an extended trapezoidal rule was used for the time integration and $t_j = t_0 + j[(t_a - t_0)/N - 1]$, $j = 0, \cdots, N - 1$, where $N - 1 = 31$ is the total number of subintervals in the integration period $[t_0, t_a]$.

The summation in the sensitivity formula (3.21)

$$VR|_{t=t_j} = \frac{2}{t_a - t_0} \sum_{m=1}^{mm} a_m \sin \frac{\pi m j}{N-1}, \quad j = 0, \cdots, N-1 \quad (3.23)$$

is obtained by a fast Fourier sine transform.

2) Response in spectral space

Expressing the response $R(t)$ and the indirect effect $R'_x \mathbf{h}_x$ as functions of the model variables in spectral space enables us to study the sensitivity of the blocking response to changes in different wavenumber amplitudes. Then (3.11) and (3.12) are written as

$$R(t) = \mathbf{a}^T \mathbf{S}\mathbf{x} H_-(-\mathbf{a}^T \mathbf{S}\mathbf{x}), \quad (3.24)$$

$$R'_x \mathbf{h}_x = \mathbf{a}^T \mathbf{S}\mathbf{h}_x H_-(-\mathbf{a}^T \mathbf{S}\mathbf{x}^0), \quad (3.25)$$

where the vector $\mathbf{a}$ is defined in (3.13). The partial derivative of the response with respect to the spectral model variables $\mathbf{x}$ is

$$R'_x = \left(\cdots, H_-(-\mathbf{a}^T \mathbf{S}\mathbf{x}^0)\mathbf{a}^T \mathbf{S} \frac{\partial \mathbf{x}}{\partial x_i}, \cdots \right). \quad (3.26)$$

Following the derivation in section 2b, the sensitivity of the response to variations $\mathbf{h}_x$ and $\mathbf{h}_\alpha$ can be obtained by integrating the adjoint model

$$\mathbf{L}^*(\mathbf{x}^0(t), \alpha^0)\mathbf{q}_m(t) = (R'_x)^T \sin \frac{\pi m(t - t_0)}{t_a - t_0} \quad (3.27a)$$

$$\mathbf{q}_m(t_a) = 0, \quad (3.27b)$$

and using the formula

$$VR(\mathbf{x}^0(t), \alpha^0; \mathbf{h}_x(t), \mathbf{h}_\alpha) = \sum_m \{[\mathbf{q}_m(t), \mathbf{Q}(\mathbf{x}^0(t), \alpha^0)\mathbf{h}_\alpha] + [\mathbf{h}_x(t_0), \mathbf{q}_m(t_0)]\} \sin \frac{\pi m(t - t_0)}{t_a - t_0}. \quad (3.28)$$

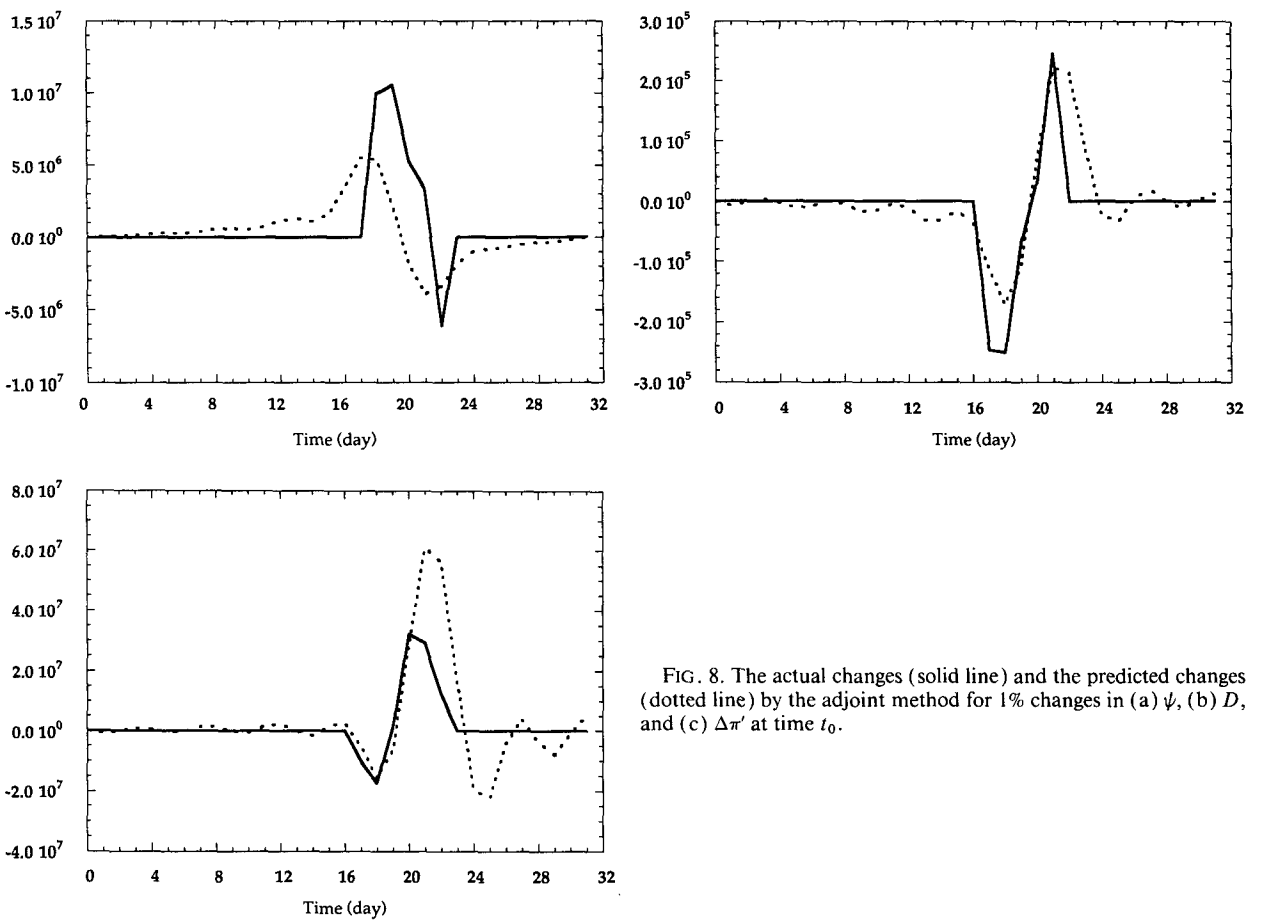

FIG. 8. The actual changes (solid line) and the predicted changes (dotted line) by the adjoint method for 1% changes in (a) ψ, (b) D, and (c) $\Delta\pi'$ at time t_0.

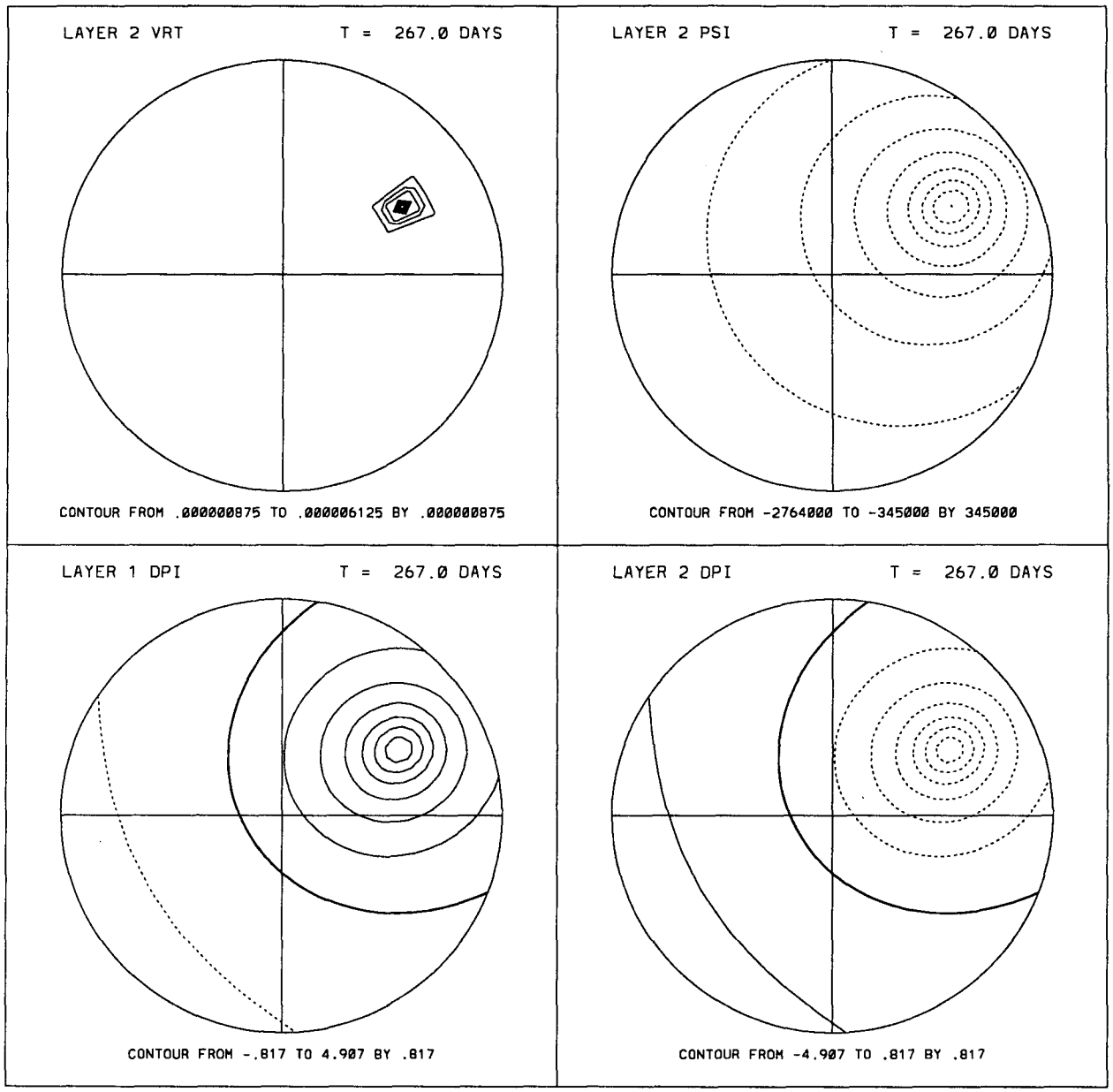

FIG. 9. Distribution of (a) an isolated vorticity source used as the perturbation for the vorticity field at time t_0, to which a sensitivity analysis is applied, (b) the corresponding streamfunction distribution, and (c) and (d) the layer thicknesses.

d. Additional verification procedures

Before carrying out the adjoint sensitivity analysis, we present several additional checks that ensure that the development of the matrices **L**, **Q**, and **L*** is free of errors. Figure 6 shows the actual change and the predicted change by the direct method when $\mathbf{h}_x$ assumes the following values:

$$\mathbf{h}_x^{\epsilon} = (\epsilon\psi^0(t_0), 0, 0)^{\mathrm{T}}, \quad \epsilon = 0.01, 0.001,$$

and $\mathbf{h}_\alpha = 0$. The predicted change (dotted line) is slightly smaller than the actual change (solid line) when $\epsilon = 0.01$, and they are nearly the same (dashed line and circles, respectively) when we decrease ϵ by another order of magnitude. This ensures that the **L** matrix is constructed correctly. We then compare the predicted change and the actual change when $\mathbf{h}_x = 0$ and $\mathbf{h}_\alpha = (0.01\tau^0_{\mathrm{drag}}, 0, 0, 0)^{\mathrm{T}}$ (Fig. 7). The predicted change matches the actual change very well. Similar results are obtained when $\mathbf{h}_\alpha = (0, 0.01\tau^0_{\mathrm{diab}}, 0, 0)^{\mathrm{T}}$, $\mathbf{h}_\alpha = (0, 0, 0.01\pi^0_\beta, 0)^{\mathrm{T}}$, or $\mathbf{h}^0_\alpha = (0, 0, 0, 0.01h_0)^{\mathrm{T}}$. These results ensure that the **Q** matrix calculation is correct. The

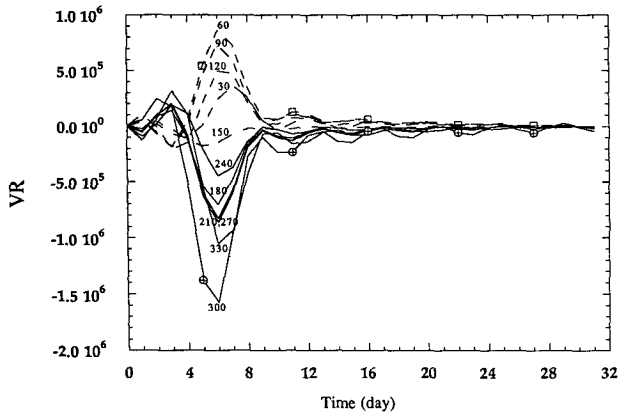

FIG. 10. Values of sensitivity of the blocking index response at longitude 22° between days 267 and 298 to perturbation in Fig. 8 placed at day 267. The numbers on the line indicate the longitudes of the vorticity perturbation center.

correctness of the adjoint operator $\mathbf{L}^*$ was verified by applying an identity check (see Navon et al. 1992).

In Fig. 8 we show the predicted changes using the adjoint method and the actual change for three values of $\mathbf{h}_x$:

$$\mathbf{h}_x^1 = (0.01\psi^0(t_0), 0, 0)^T$$
$$\mathbf{h}_x^2 = (0, 0.01D^0(t_0), 0)^T$$
$$\mathbf{h}_x^3 = (0, 0, 0.01\Delta\pi'^0(t_0))^T, \qquad (3.29)$$

and $\mathbf{h}_\alpha = 0$. We observe that the adjoint-predicted changes due to $\mathbf{h}_x^1$ and $\mathbf{h}_x^2$ are smooth and slightly smaller than the corresponding actual changes, and the adjoint-predicted change due to $\mathbf{h}_x^3$ is smooth and slightly greater that the actual change. These discrepancies are partly due to the nonlinear effect (see Fig. 6) and partly due to the nondifferentiable property of the response function.

From now on, we will use only the adjoint formulations (3.21) and (3.28) for all the numerical experiments that follow.

e. *Sensitivity analysis*

1) SENSITIVITY TO ISOLATED VORTICITY PERTURBATION IN THE UPPER LEVEL

The first sensitivity experiment deals with the longitudinal placement of an isolated vorticity perturbation at a fixed latitude, 43°N, in the upper layer (Fig. 9) at time t_0 for block 272; the thickness field is calculated by solving the nonlinear balance equation (Charney 1962) at the initial time $t_0 = 267$ or $t_0 = 315$. Figure 10 shows the various sensitivities of the blocking index when the placement longitude is varied in steps of 30° of longitude. The isolated vorticity perturbation placed upstream of the block at 300°W seems to produce the maximum change in favor of a more intense block 4 days later near 22°W. A small vorticity source placed downstream of the block at 60° of longitude weakens the block.

The streamfunction maps in both layers near day 267 are shown in Fig. 2. A naive explanation may go as follows. With the mean upper-level wind (about 30 m s^{-1}), it takes about 1 day for a quantity placed near the center of the westerly current to be advected 30° of longitude; day 267 is 3 or 4 days from block formation. Therefore, a midlatitude vortex placed 90° of longitude upstream of the block will reach longitude 22°W at day 267 + 3. Colucci (1985, 1987), while studying the role of synoptic scales on block formation, found that rapidly deepening cyclones can play a major role in block formation or demise. He found that not only is the cyclone intensity important but its phase in relation to the planetary waves is crucially important; the importance of the phase placement is responsible for enhanced potential vorticity transports associated with the synoptic-scale motion.

2) SENSITIVITY ANALYSIS TO CHANGES IN STREAMFUNCTION AT TIME t_0

We use (3.21) to perform sensitivity analysis when at time t_0 we introduce a 1% change in ψ at each Gaussian grid point, respectively. This kind of sensitivity analysis will indicate where the most sensitive regions for a certain blocking event are. Since the total number of the grid points in one layer is 96×24, more than 4000 sensitivity calculations are required. The computational advantage of using the adjoint sensitivity analysis method is thus clearly evident. Again, we require the system to be in balance, and plot spatial maps of relative sensitivity.

Figures 11a,b show, for $t_0 = 267$, upper- and lower-layer relative sensitivity of the blocking index at longitude 22°W at days $t_0 + (5, 6, 7)$ for block 272; thus, we are in preblock conditions. The results in Fig. 11 display the sensitivity of the blocking index 5, 6, and 7 days after t_0 to alterations of the structure of the streamfunctions at t_0. The dominant sensitivity occurs upstream of the block in the upper layer and slightly downstream of the block in the lower layer. If we look at Fig. 2 we see that both the upstream trough in the upper layer and the downstream cyclone in the lower layer intensifies in the preblock stage. The positions of the contours in the upper and lower layers are consistent with advection calculations using typical mean winds in each layer (about 30 m s^{-1} and 10 m s^{-1} in the upper and lower layers, respectively). The features shown have meridional scales of some 10^3 km, with longitudinal scale several times larger. It seems that the presence of jet streak in the fourth quadrant at day 267

(a)

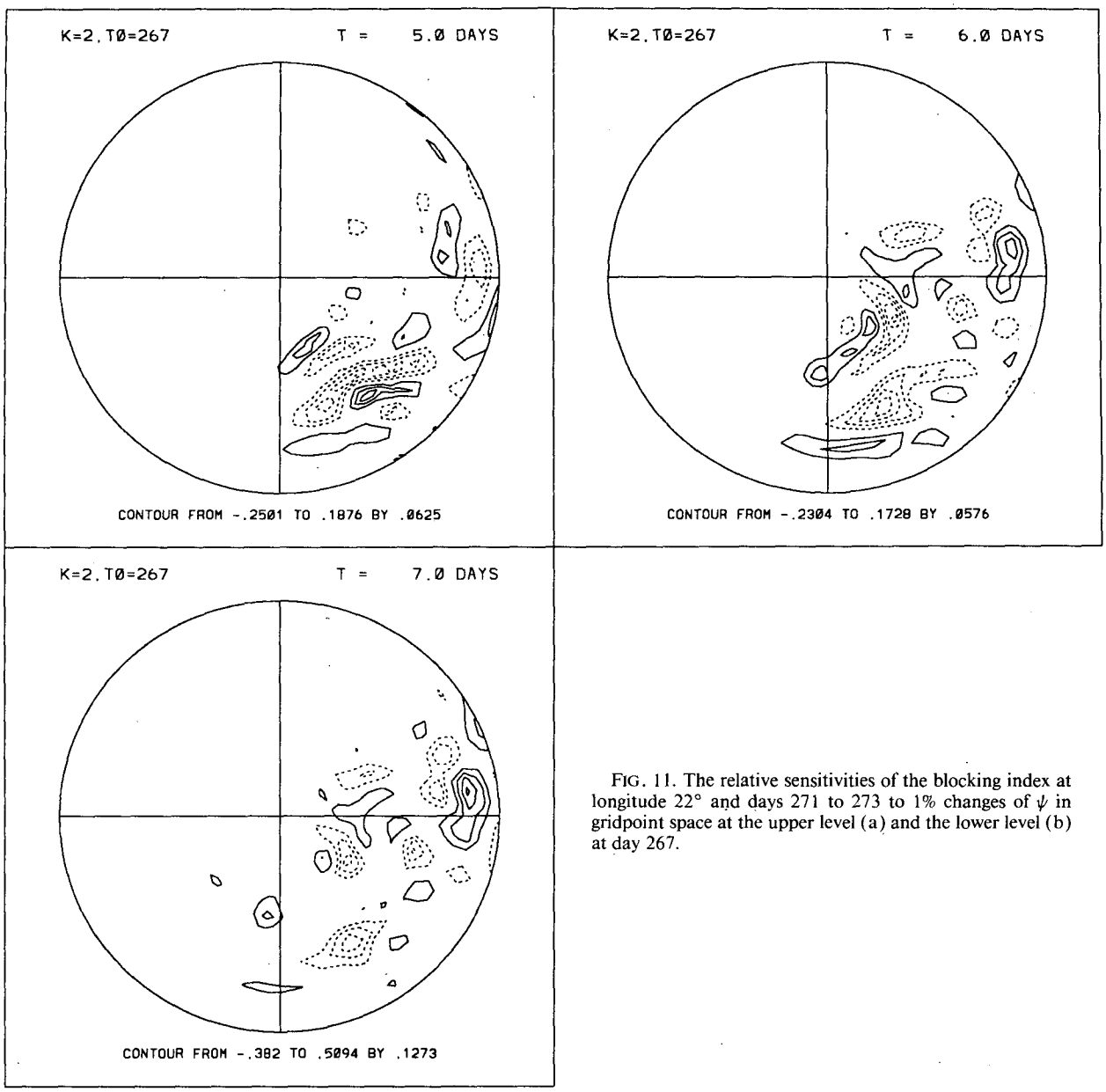

FIG. 11. The relative sensitivities of the blocking index at longitude 22° and days 271 to 273 to 1% changes of ψ in gridpoint space at the upper level (a) and the lower level (b) at day 267.

when the perturbation is introduced contributes to a large number of contours in that quadrant. Thus, block formation may be more sensitive to certain flow features.

Figures 12a,b show maps of the relative sensitivity for block 283 and t_0 = 280, that is, sensitivity of the blocking index at longitude 150°W to changes in streamfunction 6 days before block formation. There is a large number of contours, at least in the upper layer, that emanate from the tropics. In the lower layer, some of the contours lie in the wake of the block. Also, the features in Fig. 12 have larger longitudinal scale than that in Fig. 11. All this may imply that the blocks at days 272 and 283 might be caused by different mechanisms. Visual observation of the streamfunction distribution in upper and lower layers (Figs. 2 and 3) tends to support this conjecture. For block 272, the cyclogenesis upstream of the block seems to play a major role. For the block 283, the block seems to be the result of a Rossby wave breaking, since both the lee cyclone near 270° and the upstream high intensify from day 281 and move westward. The ridge in the upper layer is rather stable, and the downstream trough extends southwestward.

(b)

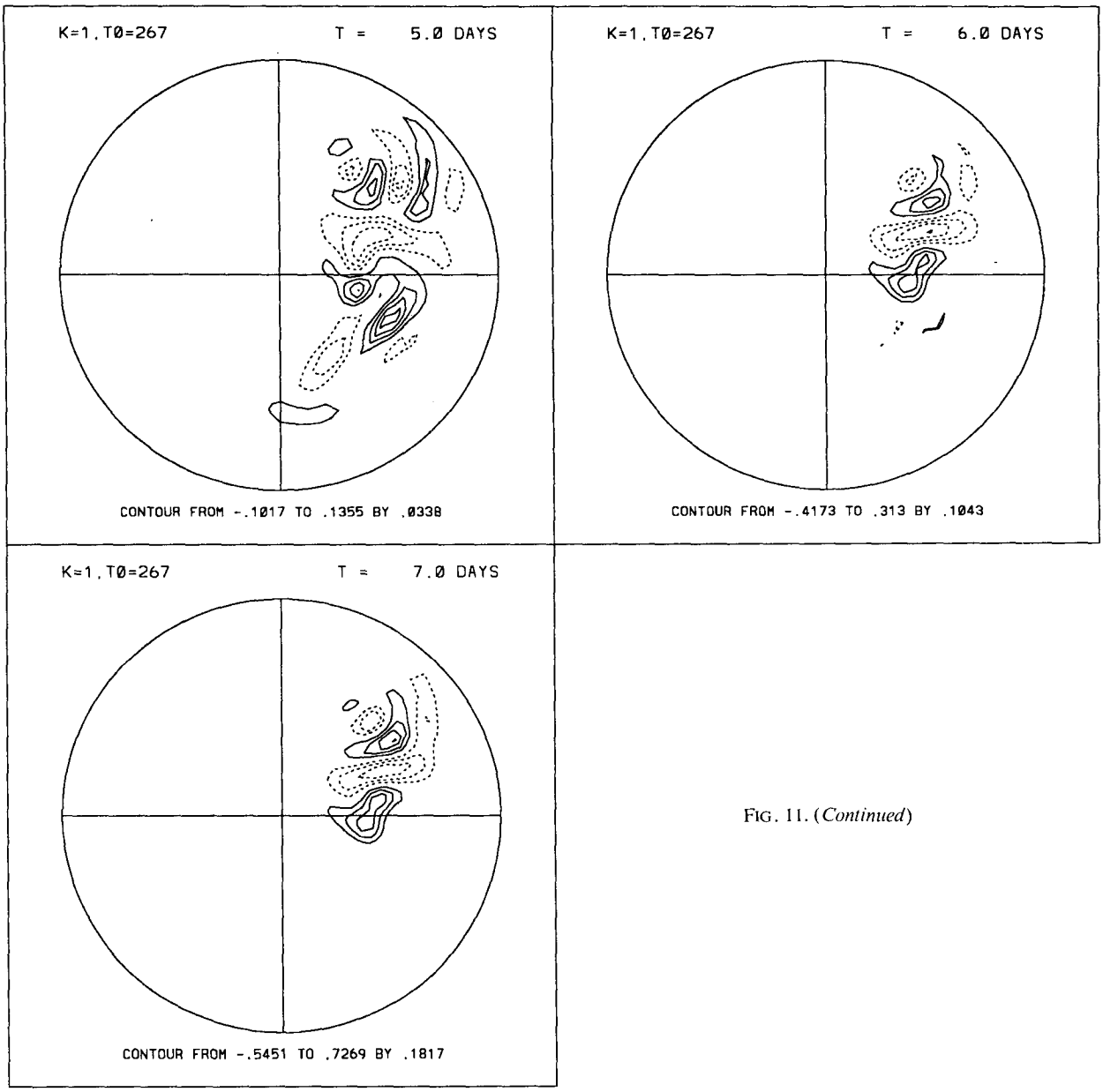

FIG. 11. (*Continued*)

3) SENSITIVITY ANALYSIS TO CHANGES IN SPECTRAL SPACE

We here address the sensitivity of the blocking index to different scale processes using (3.28) for block 272. The relative sensitivity to different wavenumbers in the upper layer is calculated and shown in Table 2.

The results of sensitivity analysis in spectral space seem to confirm, for the simple model considered, the importance of zonal flow, planetary waves, as well as synoptic-scale features: positive relative sensitivities that imply block intensification and scale $(m, n) = (2, 7; 0, 11; 6, 11)$ have a dominant contribution. The quantities (m, n) represent the zonal wavenumber and the number of nodes in the meridional direction. A Cartesian meridional wavenumber, l, could be approximately obtained by writing $l \approx (m - n)/2$. These sets of waves that contribute most to block intensification could then be characterized as being planetary $(m = 0, n = 7; l \approx 2)$, synoptic $(m = 6, n = 11; l \approx 2)$, and zonal $(m = 0, n = 11; l \approx 5\text{–}6)$. The waves producing negative values of relative sensitivity that contribute most to block demise are $(m, n) = (0, 3; 0, 7; 4, 7; 5, 6)$. Again, zonal flow seems to be very important, in agreement with the findings in the next section.

(a)

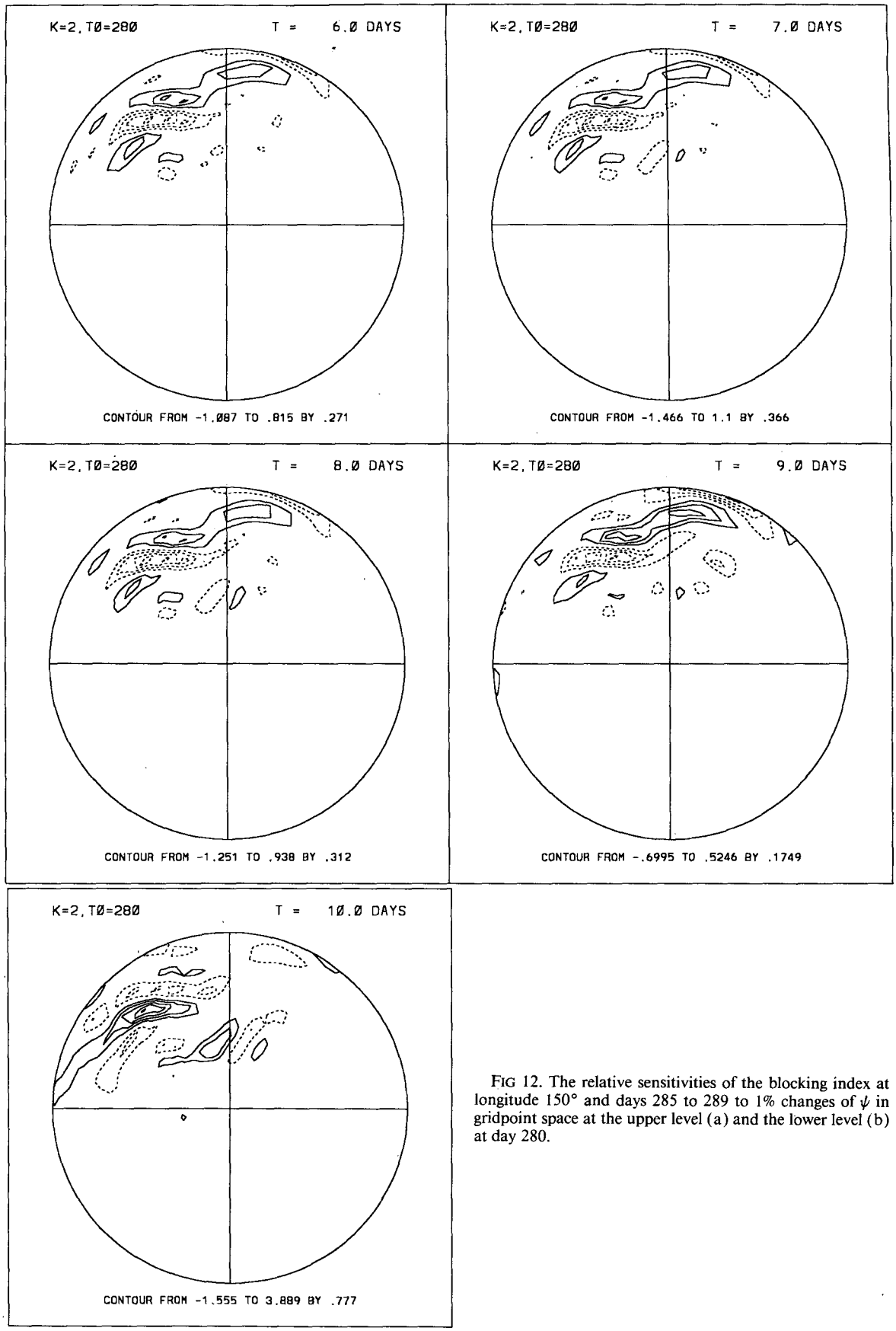

FIG 12. The relative sensitivities of the blocking index at longitude 150° and days 285 to 289 to 1% changes of ψ in gridpoint space at the upper level (a) and the lower level (b) at day 280.

(b)

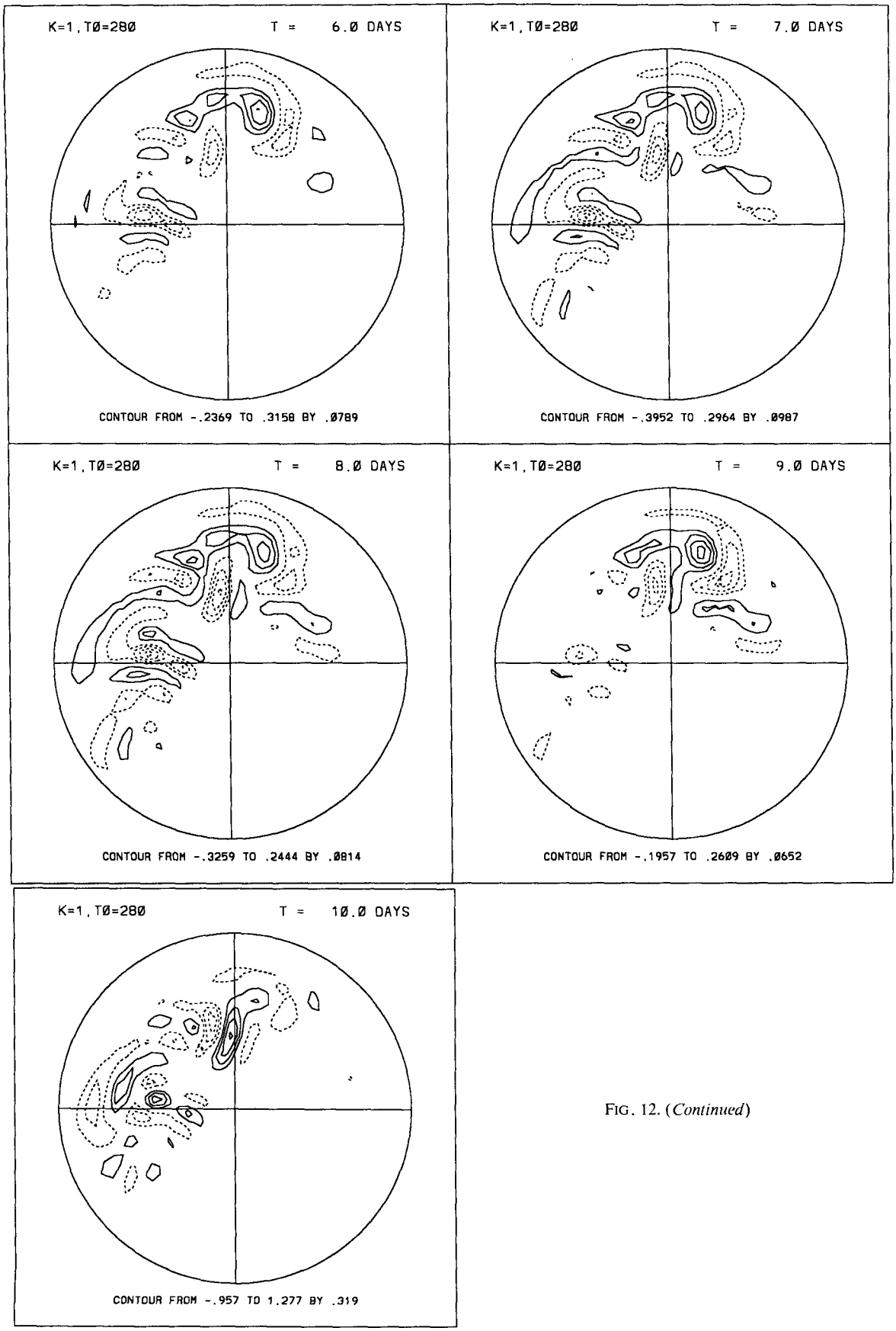

FIG. 12. (*Continued*)

TABLE 2. Values of relative sensitivities of the blocking index at longtiude 150°W and day 271 to 1% perturbations of ψ in spectral space at the upper layer at time $t_0 = 267$ (day).

	$m=0$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$
$n=0$								
$n=1$	−3.0E−2*							
$n=2$		−4.6E−4						
$n=3$	−2.7E−1		−2.3E−2					
$n=4$		−4.4E−2		2.7E−2				
$n=5$	8.6E−3		−4.9E−2		−5.0E−2			
$n=6$		−5.0E−2		−3.4E−2		−1.7E−3		
$n=7$	−1.2E−1		8.8E−2		−1.2E−1		1.3E−2	
$n=8$		−4.1E−3		6.4E−2		−8.7E−2		8.8E−3
$n=9$	−2.5E−2		3.5E−2		−5.0E−2		−2.2E−2	
$n=10$		−4.0E−3		−1.3E−2		−4.8E−2		−1.2E−2
$n=11$	9.5E−2		−7.7E−3		−2.0E−2		8.4E−2	
$n=12$		3.1E−2		−9.8E−2		−1.0E−2		−8.4E−3
$n=13$	−4.6E−4		−8.1E−4		−6.4E−3		−1.9E−2	
$n=14$		2.0E−3		−7.8E−3		2.9E−3		−4.2E−4

* −3.0E−2 means -3.0×10^{-2}.

The main sensitivities in the wave band $m = 0$–6 occur mainly in meridional wavenumber $n = 0$–9. The sensitivities in wavenumbers $m = 7$–12 are one order of magnitude smaller than those in wavenumbers $m = 0$–6. The major sensitivities in this wave band ($m = 7$–12) have smaller meridional scale ($9 \leqslant n \leqslant 12$).

4) SENSITIVITY ANALYSIS TO MODEL PARAMETERS

If a variation occurs solely in the nth parameter, we denote by $\mathbf{h}_\alpha^n$ the corresponding vector of parameter variation

$$\mathbf{h}_\alpha^n = (0, \cdots, h_\alpha^n, \cdots, 0)^{\mathrm{T}}, \qquad (3.30)$$

and denote the corresponding sensitivity by VR^n. The relative sensitivity s_n is defined as the dimensionless quantity

$$s_n = \frac{VR^n}{R}\left(\frac{h_\alpha^n}{\alpha_n^0}\right)^{-1}. \qquad (3.31)$$

Note that because $R \leqslant 0$ and $h_\alpha^n/\alpha_n^0 > 0$, s_n and VR^n have opposite signs; $s_n > 0$ implies block intensification. The magnitudes of the relative sensitivities serve as a guide to ranking the importance of model parameters.

Figure 13 shows the relative sensitivity of the blocking index at longitude $\lambda_1 = 23°$W in the time window of [267, 298]. Only the relative sensitivities for the 5 days when the blocking index is not equal to zero are plotted. We observe an antisymmetric distribution of the time evolution of the relative sensitivity, which is produced by the symmetric-like distribution of the response function (see Fig. 4). Of the four parameters in the model, the largest sensitivity corresponds to the mountain height h_0, as can be seen from Fig. 13. The second most important parameter affecting the blocking index is π_β, which is a measure of the slope of the layer interface and therefore a measure of the upper-level zonal jet in the radiative drive. Next comes the surface drag that controls the baroclinic life cycle via the barotropic governor mechanism of James and Gray (1986) and James (1987). Mechanical drag is an order of magnitude more important than the Newtonian damping, τ_{diab}. The relative importance of these parameters could be justified on physical grounds. We suspect that the blocks are a result of the interaction between synoptic-scale baroclinic waves and planetary-scale zones of diffluence associated with forced mountain waves. An increase in the mountain height produces a larger-amplitude stationary wave and stronger regions of diffluence (Shutts 1983). A decrease in the

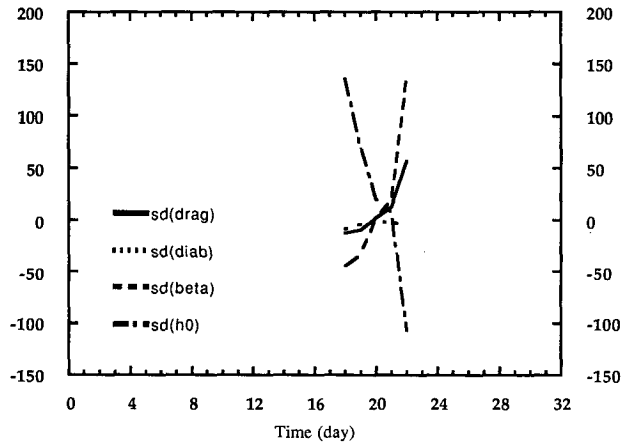

FIG. 13. The relative sensitivities of the blocking index to model parameters τ_{drag} (solid line), τ_{diab} (dotted line), π_β (dashed line), and h_0 (dash-dotted line).

surface drag suggests a stronger forced response, stronger diffluence, and more blocks. From the work of James and Gray (1986) and James (1987) we know that it does not necessarily lead to stronger baroclinic waves, and the maintenance mechanism of blocks may be made less effective.

f. Computational cost

Table 3 presents the computational cost of the generalized adjoint sensitivity analysis. For one time integration, the CPU time spent in the adjoint model (3.17) integration is double the time required for the nonlinear model integration (2.11). The high cost of integrating the G-differential equation (3.19) originates from the left-hand-side term $\mathbf{Qh}_\alpha$, which is calculated at every time step. The CPU times for calculating the predicted change by the adjoint method and the direct method are 1455 (with 16 times adjoint integrations) and 1119 s, respectively. The CPU time for calculating the actual change is 92 s. It seems that the adjoint sensitivity analysis requires the highest CPU time. However, once the values of the adjoint variables q_m, $m = 1, \cdots, 16$ are obtained, the subsequent sensitivity analysis is computationally very cheap. The sensitivity calculation of the response to any changes in the model variables takes only 1 s, and the sensitivity calculation of the response to any changes in the model parameters requires only 33 s, which is the time required to compute the quadrature in (3.21). Therefore, the adjoint sensitivity analysis is most economic for sensitivity analysis when the dimension of the model-state variables or the number of the model parameters is large.

The computational cost is rather small to compute sensitivity of the blocking index to vorticity sources in different locations. Once the values of the adjoint variables q_m, $m = 1, \cdots, 16$ have been obtained for this blocking index, the adjoint sensitivity calculation to any perturbations of the model dynamic fields takes only few seconds of CPU time. The CPU time required to calculate sensitivity to perturbation at different grid points of the model variables—that is, $2 \times \text{NLONS} \times \text{NLATSH} = 4608$ time calculations of VR in (3.21) or $2 \times \text{NMDIM} = 1024$ times calculations of (3.28)—requires only 10–40 s, which would be otherwise impossible using the forward sensitivity formulation.

TABLE 3. Computational cost of the adjoint sensitivity analysis.

Models	32-day integration	Sensitivity calculation
Adjoint model	86 s	1455 s
Adjoint model		variable (1 s) parameter (33 s)
G-differential model	1074 s	1119 s
Nonlinear model	46 s	92 s

4. Discussion and conclusions

This work has demonstrated the application of the sensitivity analysis method developed by Cacuci (1981a,b) for nonlinear systems and operator-type responses to a two-layer isentropic spectral model. The results for this model have shown the method to be both efficient and accurate. A few adjoint calculations, each involving a similar amount of computation as would be required for solving once the tangent linear of the original model, sufficed to obtain sensitivities of a blocking index to all the relevant model parameters and model states (the total dimension is 3076 in spectral space and 13 828 in grid space). The use of sensitivities of a time-dependent response has been discussed and illustrated in detail. In this context, the adjoint sensitivities accurately predicted the effect of small variations in the model state and parameters. Relative sensitivities were used to rank the importance of all the parameters. The sensitivity of the blocking index to mountain height was found to be the largest. The sensitivity analysis in grid space and spectral space has shown that the significant sensitivities for blocking occur in some preferred regions and spectral wave bands.

A marked difference between the generalized adjoint sensitivity analysis presented in this paper and the adjoint sensitivity used until now in meteorology consists in the following two issues.

(a) The response function is not a scalar value but an operator depending on time and/or space. Due to this basic difference, more comprehensive expressions for the sensitivity were obtained.

(b) In our case, the model parameters are taken as the control variables in addition to the model state variables, whereas in the case presented by Errico and Vukicevic (1992), for instance, the model-state variables are the only control variables. This results in a more complete tangent linear model with a right-hand side consisting of the first derivative of the nonlinear model with respect to the model parameters.

The two-layer isentropic spectral model includes many aspects found in numerical weather prediction models. This work has shown that the efficient application of the adjoint model to sensitivity analysis is feasible. The response can be a function of both space and time. When derived via the adjoint sensitivity formalism, the exact expression of the sensitivity of an operator-type response contains as many adjoint functions as there are nonzero terms in the expansion in time and space of the indirect term [see (2.17)] and the response [see (2.24)]. While adjoints of NWP models are increasingly applied in variational data assimilation and parameter estimation, the present application of the adjoint method to sensitivity analysis shows that large-scale exhaustive sensitivity analyses of any 3D NWP models are feasible without any con-

ceptual difficulties. It is a useful diagnostic tool and helps explore the physical details of sensitivity. The generalized adjoint sensitivity analysis is also a useful tool for synoptic examination and data impact studies, since it quantifies the antecedent conditions that most affect a specified forecast aspect. It also helps assess the reliability of forecasts. In particular, this approach can help assess reliability of forecasts in data-void versus data-rich regions, especially if the sensitivity with respect to a response function is found to be larger over a data-rich region. This will enhance our confidence in the forecast. All these prove that the generalized adjoint sensitivity analysis constitutes a useful and powerful tool for numerical weather prediction.

Acknowledgments. The first and third authors were funded by NSF Grant ATM-9102851, while the second author was funded by NSF Grant ATM-9112564. We would like to thank Dr. Pamela Stephens of NSF for continued support. The second author also acknowledges support of AFOSR Grants 49620-92-J-0498 and NAG8-760. Additional support was provided by the Supercomputer Computations Research Institute at The Florida State University, which is partially funded by the Department of Energy through Contract DE-FC0583ER250000. The insightful contributions of two anonymous reviewers that helped clarify the presentation of this paper is acknowledged.

APPENDIX A

Description of the Isentropic Two-Layer Model

The isentropic primitive equations can be written as

$$\frac{\partial(\zeta+f)}{\partial t}+\nabla\cdot\mathbf{J}=0 \quad \text{(vorticity equation)}, \tag{A.1}$$

$$\frac{\partial D}{\partial t}-\mathbf{k}\cdot\nabla\times\mathbf{J}+\nabla^2\left(M+\frac{\mathbf{V}\cdot\mathbf{V}}{2}\right)=0$$

$$\text{(divergence equation)}, \tag{A.2}$$

$$\frac{\partial\sigma}{\partial t}+\nabla\cdot(\mathbf{V}\sigma)+\frac{\partial(\dot\theta\sigma)}{\partial\theta}=0 \quad \text{(continuity equation)}, \tag{A.3}$$

$$\frac{\partial M}{\partial\theta}=\pi \quad \text{(hydrostatic equation)}, \tag{A.4}$$

where $\zeta+f$—absolute vorticity, or potential vorticity (PV) per unit volume; $\mathbf{J}=[u(\zeta+f)+\dot\theta\partial v/\partial\theta-F_j]\mathbf{i}+[v(\zeta+f)-\dot\theta\partial u/\partial\theta+F_i]\mathbf{j}$—the flux of PV per unit volume due to advective, diabatic and frictional effects; $\pi=C_p(p/p_0)^\kappa$—the Exner function; $M=\pi\theta+gz$—the Montgomery potential; $\sigma=-g^{-1}\partial p/\partial\theta$—the isentropic mass density; and $\mathbf{F}=F_i\mathbf{i}+F_j\mathbf{j}$—the local friction force per unit mass.

As a direct consequence of (A.1), "PV substance", that is, the amount of PV per unit volume, can neither be created nor destroyed in an isentropic layer except where that layer intersects a boundary. The PV substance is simply concentrated within that layer by advective, diabatic, and friction forces (Haynes and McIntyre 1987).

For simplicity, the nonlinearity in the relationship between σ and M is eliminated by making the Boussinesq approximation. This is done by noting that

$$\sigma=\rho_a\frac{\partial z_a}{\partial\theta}, \tag{A.5}$$

where

$$\rho_a=\frac{p_0}{R\theta_0}\left(\frac{\pi}{C_p}\right)^{C_v/R}, \quad z_a=\frac{\theta_0}{g}(C_p-\pi) \tag{A.6}$$

are the density and height in an isentropic atmosphere with potential temperature θ_0. Assuming $\rho_a=\rho_0=p_0/R\theta_0$ gives

$$\sigma=-\frac{p_0}{R\theta_0}\frac{\partial\pi}{\partial\theta}. \tag{A.7}$$

After the Boussinesq approximation, (A.3) becomes

$$\frac{\partial}{\partial t}\left(\frac{\partial\pi}{\partial\theta}\right)+\nabla\cdot\left(\mathbf{V}\frac{\partial\pi}{\partial\theta}\right)+\frac{\partial}{\partial\theta}\left(\dot\theta\frac{\partial\pi}{\partial\theta}\right)=0. \tag{A.8}$$

The fluid system to be studied consists of two layers of constant potential temperature on a rotating sphere. The subscripts 1 and 2 are used to denote the lower and upper layers, respectively, and 1/2, 3/2, and 5/2 the surface, layer interface, and upper boundary, respectively. The boundary conditions are taken as

$$\dot\theta_{5/2}=\dot\theta_{1/2}=0. \tag{A.9}$$

The condition (A.9) means that there is no mass exchange through the upper and lower boundaries. In addition, the upper boundary is assumed to be a free surface so that $\pi_{5/2}$ = constant. Therefore, the continuity equation (A.8) in the upper layer becomes an equation for the layer interface pressure $\pi_{3/2}$.

Using the hydrostatic equation, $\partial\pi/\partial z=-g/\theta$, it can be shown that $\partial M/\partial z=\pi\partial\theta/\partial z$. Therefore, M can change only in the vertical when θ changes, and so is independent of height within each layer. The horizontal momentum equations then require that u and v (hence ζ and D) must remain independent of height within each layer if they are initially so. The dynamics of each layer is then similar to that of a shallow-water model.

By the definition of M at the surface, we then have

$$M_1=\pi_{1/2}\theta_1+gz_s, \tag{A.10}$$

where z_s is the topographic height. Integrating the hydrostatic equation across the layer interface results in

$$M_2=M_1+\pi_{3/2}\Delta\theta=\theta_1(\Delta\pi_1+\Delta\pi_2+\pi_{5/2})+gz_s+\Delta\theta(\Delta\pi_2+\pi_{5/2}). \tag{A.11}$$

Using (A.10) and (A.11) and following the discretization scheme of Hsu and Arakawa (1990) for the vertical advection terms appearing in **J**, the discretized equations (A.1) and (A.2), and (A.8) can be written as

$$\frac{\partial \zeta_k}{\partial t} + \nabla \cdot [(A_k, B_k)] = 0, \tag{A.12}$$

$$\frac{\partial D_k}{\partial t} + \nabla \cdot [(-B_k, A_k)] + \nabla^2 \left[M_k + \frac{1}{2}(u_k^2 + v_k^2) \right] = 0, \tag{A.13}$$

$$\frac{\partial \Delta\pi_k}{\partial t} + \nabla \cdot [(u_k \Delta\pi_k, v_k \Delta\pi_k)] = \frac{(-1)^k \dot{\theta}_{3/2}}{\Delta\theta}(\Delta\pi_2 + \Delta\pi_1), \quad k = 1, 2, \tag{A.14}$$

where

$$A_k = u_k(\zeta_k + f) + \frac{\dot{\theta}_{3/2}(\Delta\pi_1 + \Delta\pi_2)}{2\Delta\theta\Delta\pi_k} \times (v_2 - v_1) - \left(\alpha\nabla^v - \frac{\delta_{k1}}{\tau_{\text{drag}}} \right) v_k, \tag{A.15a}$$

$$B_k = v_k(\zeta_k + f) - \frac{\dot{\theta}_{3/2}(\Delta\pi_1 + \Delta\pi_2)}{2\Delta\theta\Delta\pi_k} \times (u_2 - u_1) + \left(\alpha\nabla^v - \frac{\delta_{k1}}{\tau_{\text{drag}}} \right) u_k, \tag{A.15b}$$

$$\Delta\pi_k = \pi_{k-1/2} - \pi_{k+1/2}, \tag{A.15c}$$

$$\Delta\theta = \theta_2 - \theta_1. \tag{A.15d}$$

Here, we have parameterized the frictional force **F** using a linear mechanical damping applied in the lower layer only (to simulate surface drag), and a hyperdiffusion in both layers (to control the enstrophy cascade).

The diabatic heating $\dot{\theta}_{3/2}$ is parameterized as a Newtonian relaxation toward an equilibrium interface Exner function π_e. Therefore, for $k = 2$, (A.14) becomes

$$\frac{\partial \pi_{3/2}}{\partial t} + \nabla \cdot [(u_2 \Delta\pi_2, v_2 \Delta\pi_2)] = \frac{\dot{\theta}_{3/2}}{\Delta\theta}(\Delta\pi_2 + \Delta\pi_1) \equiv \frac{\pi_e - \pi_{3/2}}{\tau_{\text{diab}}}. \tag{A.16}$$

Thus, the diabatic heating at the interface is

$$\dot{\theta}_{3/2} = \frac{(\pi_e - \pi_{3/2})\Delta\theta}{\tau_{\text{diab}}(\Delta\pi_2 + \Delta\pi_1)}, \tag{A.17}$$

implying that mass can be transferred between the layers.

The prognostic equations (A.12)–(A.14), together with the diagnostic relations (A.10), (A.11), (A.15), and (A.17), are integrated numerically in spherical geometry using a spectral transform technique similar to that described by Browning et al. (1989). Spherical harmonic basis functions with triangular truncation, truncated at wavenumber 31, and a semi-implicit time integration scheme with a time step of 45 min are used. The semi-implicit scheme requires that the layer thickness $\Delta\pi$ be split into a horizontally uniform reference state $\Delta\bar{\pi}$ and a horizontally varying part $\Delta\pi'$:

$$\Delta\pi = \Delta\bar{\pi} + \Delta\pi', \tag{A.18}$$

where

$$\bar{\Delta}\pi_1 = C_p - \pi_{\text{mid}},$$

$$\bar{\Delta}\pi_2 = \pi_{\text{mid}} - \pi_{5/2},$$

$$\pi_{\text{mid}} = \frac{1}{2}(C_p + \pi_{\text{top}}).$$

The divergence equation (A.13) and the continuity equation (A.14) are then written as

$$\frac{\partial D_k}{\partial t} + \nabla \cdot [(-B_k, A_k)] + \nabla^2 \left[gz_s + \frac{1}{2}(u_k^2 + v_k^2) \right] + \nabla^2 M'_k = 0, \tag{A.19}$$

$$\frac{\partial \Delta\pi'_k}{\partial t} + \nabla \cdot [(u_k \Delta\pi'_k, v_k \Delta\pi'_k)] + \Delta\bar{\pi}_k D_k = (-1)^k \frac{\pi_e - \pi_{\text{mid}} - \Delta\pi'_2}{\pi_{\text{diab}}}. \tag{A.20}$$

$$k = 1, 2$$

All dependent variables are expressed in terms of a spherical harmonic representation of the form

$$F_k(\phi, \lambda, t) = \sum_{l=-J}^{J} \sum_{n=|l|}^{J} F_{l,n,k}(t) Y^l_n, \tag{A.21}$$

where $J = 31$ and the spherical harmonics Y^l_n are given by

$$Y^l_n(\phi, \lambda) = P^l_n(\sin\phi) e^{il\lambda}, \tag{A.22}$$

where P^l_n are the associated Legendre polynomials.

The spectral form of the model equations (A.12), (A.19), and (A.20) is

$$\frac{d\zeta_{l,n,k}}{dt} = A_{l,n,k} + \delta_1(n)\zeta_{l,n,k}, \tag{A.23}$$

$$\frac{dD_{l,n,k}}{dt} = B_{l,n,k} + \delta_1(n)D_{l,n,k} + \delta_2(n)M'_{l,n,k}, \tag{A.24}$$

$$\frac{d\Delta\pi'_{l,n,k}}{dt} = C_{l,n,k} - D_{l,n,k}\Delta\bar{\pi}_k, \tag{A.25}$$

where $A_{l,n,k}$ represents the spectral coefficient of all terms except the diffusion term in (A.12), $B_{l,n,k}$ the

spectral coefficient of all terms except the diffusion and $\nabla^2 M'$ terms in (A.19), $C_{l,n,k}$ the spectral coefficient of all terms except the $\Delta\bar{\pi}_k D_{l,n,k}$ term in (A.20), and

$$\delta_1(n) = -\alpha\left[\frac{n(n+1)}{a^2}\right]^{\nu/2}, \qquad \text{(A.26a)}$$

$$\delta_2 = -\frac{n(n+1)}{a^2}. \qquad \text{(A.26b)}$$

The last terms on the left-hand side of (A.24) and (A.25) are responsible for gravity-wave propagation and are treated implicitly using the Crank–Nicholson time-differencing scheme, while the other terms are treated explicitly using the leapfrog scheme. The vorticity equation (A.23) is integrated explicitly using the leapfrog scheme. A weak time filter is applied to all three prognostic equations to damp the computational mode.

The parameters settings used here are

$$\theta_1 = 280 \text{ K},$$
$$\theta_2 = 320 \text{ K},$$
$$p_0 = 1000 \text{ mb},$$
$$z_{\text{top}} = 10^4 \text{ m},$$
$$\pi_{5/2} = C_p - \frac{gz_{\text{top}}}{\theta_1},$$
$$\nu = 12,$$
$$\tau_{\text{drag}} = 5 \text{ days},$$
$$\tau_{\text{diab}} = 15 \text{ days},$$

where α is chosen so that the smallest resolvable scale is damped within an e-folding time scale of 3 h. The topographic height and equilibrium interface Exner function are given by

$$z_s = 4h_0(\mu^2 - \mu^4)\sin 2\lambda,$$

$$\pi_e = \pi_{\text{mid}} - \frac{1}{2}\pi_\beta(C_p - \pi_{\text{mid}})\cos 2\phi(\sin^2 2\phi + 2),$$

where $\mu = \sin\phi$, $h_0 = 2000$ m, $\pi_{\text{mid}} = \frac{1}{2}(C_p + \pi_{5/2})$, and $\pi_\beta = 0.75$.

APPENDIX B

Numerical Aspects of the Sensitivity Analysis: Mathematical Expression for the Components of the Matrix Q

From (A.23)–(A.25) we can derive an expression for the matrix **Q** as follows:

$$QZ_{\tau_{\text{drag}}} = \left(\cdots, \frac{\partial A_{l,n,k}}{\partial \tau_{\text{drag}}}, \cdots\right)^{\mathrm{T}}$$

$$\frac{\partial A_{l,n,k}}{\partial \tau_{\text{drag}}} = -\{\nabla\cdot[(A_k^{\text{drag}}, B_k^{\text{drag}})]\}_{l,n,k} = \frac{\delta_{k1}}{\tau_{\text{drag}}^2}\zeta$$

$$A_1^{\text{drag}} = -\frac{1}{\tau_{\text{drag}}^2}v_1, \quad A_2^{\text{drag}} = 0$$

$$B_1^{\text{drag}} = \frac{1}{\tau_{\text{drag}}^2}u_1, \quad B_2^{\text{drag}} = 0 \qquad \text{(B.1)}$$

$$QD_{\tau_{\text{drag}}} = \left(\cdots, \frac{\partial B_{l,n,k}}{\partial \tau_{\text{drag}}}, \cdots\right)^{\mathrm{T}}$$

$$\frac{\partial B_{l,n,k}}{\partial \tau_{\text{drag}}} = -\{\nabla\cdot[(-B_k^{\text{drag}}, A_k^{\text{drag}})]\}_{l,n,k} = \frac{\delta_{k1}}{\tau_{\text{drag}}^2}D \qquad \text{(B.2)}$$

$$QP_{\tau_{\text{drag}}} = 0 \qquad \text{(B.3)}$$

$$QZ_{\tau_{\text{diab}}} = \left(\cdots, \frac{\partial A_{l,n,k}}{\partial \tau_{\text{diab}}}, \cdots\right)^{\mathrm{T}}$$

$$\frac{\partial A_{l,n,k}}{\partial \tau_{\text{diab}}} = -\{\nabla\cdot[(A_k^{\text{diab}}, B_k^{\text{diab}})]\}_{l,n,k}$$

$$A_k^{\text{diab}} = -\frac{\dot{\theta}_{3/2}(\Delta\pi_1 + \Delta\pi_2)}{2\tau_{\text{diab}}\Delta\theta\Delta\pi_k}(v_2 - v_1)$$

$$B_k^{\text{diab}} = \frac{\dot{\theta}_{3/2}(\Delta\pi_1 + \Delta\pi_2)}{2\tau_{\text{diab}}\Delta\theta\Delta\pi_k}(u_2 - u_1) \qquad \text{(B.4)}$$

$$QD_{\tau_{\text{diab}}} = \left(\cdots, \frac{\partial B_{l,n,k}}{\partial \tau_{\text{diab}}}, \cdots\right)^{\mathrm{T}}$$

$$\frac{\partial B_{l,n,k}}{\partial \tau_{\text{diab}}} = -\{\nabla\cdot[(-B_k^{\text{diab}}, A_k^{\text{diab}})]\}_{l,n,k} \qquad \text{(B.5)}$$

$$QP_{\tau_{\text{diab}}} = \left(\cdots, \frac{\partial C_{l,n,k}}{\partial \tau_{\text{diab}}}, \cdots\right)^{\mathrm{T}}$$

$$\frac{\partial C_{l,n,k}}{\partial \tau_{\text{diab}}} = -\left\{(-1)^k\frac{\pi_e - \pi_{\text{mid}} - \Delta\pi'_2}{\tau_{\text{diab}}^2}\right\}_{l,n,k} \qquad \text{(B.6)}$$

$$QZ_{\pi_\beta} = \left(\cdots, \frac{\partial A_{l,n,k}}{\partial \pi_\beta}, \cdots\right)^{\mathrm{T}}$$

$$\frac{\partial A_{l,n,k}}{\partial \pi_\beta} = -\{\nabla\cdot[(A_k^{\pi_\beta}, B_k^{\pi_\beta})]\}_{l,n,k}$$

$$A_k^{\pi_\beta} = \frac{\partial\dot{\theta}_{3/2}}{\partial\pi_\beta}\frac{\Delta\pi_1 + \Delta\pi_2}{2\Delta\theta\Delta\pi_k}(v_2 - v_1)$$

$$B_k^{\pi_\beta} = -\frac{\partial\dot{\theta}_{3/2}}{\partial\pi_\beta}\frac{\Delta\pi_1 + \Delta\pi_2}{2\Delta\theta\Delta\pi_k}(u_2 - u_1)$$

$$\frac{\partial\dot{\theta}_{3/2}}{\partial\pi_\beta} = \frac{(\partial\pi_e/\partial\pi_\beta)\Delta\theta}{\tau_{\text{diab}}(\Delta\pi_1 + \Delta\pi_2)} \qquad \text{(B.7)}$$

$$QD_{\pi_\beta} = \left(\cdots, \frac{\partial B_{l,n,k}}{\partial \pi_\beta}, \cdots\right)^{\mathrm{T}}$$

$$\frac{\partial B_{l,n,k}}{\partial \pi_\beta} = -\{\nabla\cdot[(-B_k^{\pi_\beta}, A_k^{\pi_\beta})]\}_{l,n,k} \qquad \text{(B.8)}$$

$$QP_{\pi_\beta} = \left(\cdots, \frac{\partial C_{l,n,k}}{\partial \pi_\beta}, \cdots \right)^T$$

$$\frac{\partial C_{l,n,k}}{\partial \pi_\beta} = \left\{ (-1)^k \frac{\partial \pi_e / \partial \pi_\beta)}{\tau_{\text{diab}}} \right\}_{l,n,k} \tag{B.9}$$

$$QZ_{h_0} = 0 \tag{B.10}$$

$$QD_{h_0} = \left(\cdots, \frac{\partial C_{l,n,k}}{\partial h_0}, \cdots \right)^T$$

$$\frac{\partial C_{l,n,k}}{\partial h_0} = -\left\{ \nabla^2 \left(g \frac{\partial z_s}{\partial h_0} \right) \right\}_{l,n,k}$$

$$\frac{\partial z_s}{\partial h_0} = 4(\mu^2 - \mu^4) \sin 2\lambda \tag{B.11}$$

$$QP_{h_0} = 0. \tag{B.12}$$

REFERENCES

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

Browning, G. L., J. J. Hack, and P. N. Swarztrauber, 1989: A comparison of 3 numerical methods for solving differential equations on the sphere. *Mon. Wea. Rev.,* **117,** 1058–1075.

Cacuci, D. G., 1981a: Sensitivity theory for nonlinear systems. I: Nonlinear functional analysis approach. *J. Math. Phys.,* **22,** 2794–2802.

——, 1981b: Sensitivity theory for nonlinear systems. II: Extensions to additional classes of responses. *J. Math. Phys.,* **22,** 2803–2812.

——, 1988: The forward and adjoint methods of sensivity analysis. *Uncertainty Analysis,* Yigal Ronen, Ed., CRC Press, Inc., 71–144.

——, E. Greenspan, J. H. Marable, and M. L. Williams, 1980: Developments in sensitivity theory. *Proc. Conf. of 1980 Advances in Reactor Physics and Shielding,* Sun Valley, ID, American Nuclear Society, 692–704.

Charney, J. G., 1962: Integration of the primitive and balance equations. *Proc. Int. Symp. Numerical Weather Prediction,* Tokyo, Meteor. Soc. Japan, 131–152.

Colucci, S. J., 1985: Explosive cyclogenesis and large-scale circulation changes: Implications for atmospheric blocking. *J. Atmos. Sci.,* **42,** 2701–2717.

——, 1987: Comparative diagnosis of blocking versus nonblocking planetary-scale circulation changes during synoptical-scale cyclogenesis. *J. Atmos. Sci.,* **44,** 124–139.

Errico, R. M., and T. Vukicevic, 1992: A sensitivity analysis using an adjoint of the PSU–NCAR Mesoscale Model. *Mon. Wea. Rev.,* **120,** 1644–1660.

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

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

Hall, M. C. G., and D. G. Cacuci, 1983: Physical interpretation of the adjoint functions for sensitivity analysis of atmospheric models. *J. Atmos. Sci.,* **40,** 2537–2546.

——, D. G. Cacuci, and M. E. Schlesinger, 1982: Sensitivity analysis of a radiative–convective model by the adjoint method. *J. Atmos. Sci.,* **39,** 2038–2050.

Haynes, P. H., and M. E. McIntyre, 1987: On the evolution of vorticity and potential vorticity in the presence of diabatic heating and frictional or other forces. *J. Atmos. Sci.,* **44,** 828–841.

Hsu, Y. J., and A. Arakawa, 1990: Numerical modeling of the atmosphere with an isentropic vertical coordinate. *Mon. Wea. Rev.,* **118,** 1933–1959.

James, I. N., 1987: Suppression of baroclinic instability in horizontally sheared flows. *J. Atmos. Sci.,* **44,** 3710–3720.

——, and L. J. Gray, 1986: Concerning the effect of surface drag on the circulation of a baroclinic planetary atmosphere. *Quart. J. Roy. Meteor. Soc.,* **112,** 1231–1250.

Kimoto, M., H. Mukougawa, and S. Yoden, 1992: Medium-range forecast skill variation and blocking transition: A case study. *Mon. Wea. Rev.,* **120,** 1616–1627.

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

LeDimet, F. X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus,* **38A,** 91–110.

Lejenäs, H., and H. Okland, 1983: Characteristics of Northern Hemisphere blocking as determined from a long time series of observational data. *Tellus,* **35A,** 350–362.

Levine, H., and J. Schwinger, 1949: On the theory of diffraction by an aperture in an infinite plane screen 2. *Phys. Rev.,* **75,** 1423–1432.

Marchuk, G. I., 1974: Osnovnye i sopryazhennye uravneniya dinamiki atmosfery i okeana. *Meteor. Gidrol.,* **2,** 9–37.

Navon, I. M., X. Zou, J. Derber, and J. Sela, 1992: Variational data assimilation with an adiabatic version of the NMC spectral model. *Mon. Wea. Rev.,* **120,** 1433–1446.

Rabier, F., and P. Courtier, 1992: Four-dimensional assimilation in the presence of baroclinic instability. *Quart. J. Roy. Meteor. Soc.,* **118,** 649–672.

Roussopolos, R., 1953: Methodes variationnelles en theories des collisions. *C. R. Acad. Sci.,* **236,** 1858–1860.

Shutts, G., 1983: The propagation of diffluent jet streams: Eddy vorticity forcing of "blocking" flow fields. *Quart. J. Roy. Meteor. Soc.,* **109,** 737–761.

Smedstad, O. M., and O'Brien, J. J., 1991: Variational data assimilation and parameter estimation in an equatorial Pacific Ocean model. *Prog. Oceanogr.,* **26,** 179–241.

Thépaut, J. N., and P. Courtier, 1992: Four-dimensional variational data assimilation using the adjoint of a multilevel primitive equation model. *Quart. J. Roy. Meteor. Soc.,* **117,** 1225–1254.

Tibaldi, S., and F. Molteni, 1990: On the operational predictability of blocking. *Tellus,* **42A,** 343–364.

Tracton, M. S., 1990: Predictability and its relationship to scale interaction processes in blocking. *Mon. Wea. Rev.,* **118,** 1666–1695.

——, K. Mo, W. Chen, E. Kalnay, R. Kistler, and G. White, 1989: Dynamical extended range forecasting (DERF) at the National Meteorological Center. *Mon. Wea. Rev.,* **117,** 1606–1637.

Tribbia, J., 1991: On the problem of prediction beyond the deterministic range. *Proc. NATO Workshop on Predictions of Interannual Climate Variations,* Trieste, Italy, NATO/ASI (in press).

Wigner, E. P., 1945: Effects of small perturbations on pile period. Chicago Report CP-G-3048.

Zou, X., I. M. Navon, and F. X. LeDimet, 1992: An optimal nudging data assimilation scheme using parameter estimation. *Quart. J. Roy. Meteor. Soc.,* **118,** 1163–1186.