An Adjoint Sensitivity Method for the Adaptive Location of the Observations in Air Quality Modeling

DACIAN N. DAESCU

Institute for Mathematics and Its Applications, University of Minnesota, Minneapolis, Minnesota

GREGORY R. CARMICHAEL

Department of Chemical and Biochemical Engineering, and The Center for Global and Regional Environmental Research, The University of Iowa, Iowa City, Iowa

(Manuscript received 6 December 2001, in final form 31 May 2002)

ABSTRACT

The spatiotemporal distribution of observations plays an essential role in the data assimilation process. An adjoint sensitivity method is applied to the problem of adaptive location of the observational system for a nonlinear transport-chemistry model in the context of 4D variational data assimilation. The method is presented in a general framework and it is shown that in addition to the initial state of the model, sensitivity with respect to emission and deposition rates and certain types of boundary values may be obtained at a minimal additional cost. The adjoint modeling is used to evaluate the influence function and to identify the domain of influence associated with the observations. These essential tools are further used to develop a novel algorithm for targeting observations that takes into account the interaction among observations at different instants in time and spatial locations. Issues related to the case of multiple observations are addressed and it is shown that by using the adjoint modeling an efficient implementation may be achieved. Computational and practical aspects are discussed and this analysis indicates that it is feasible to implement the proposed method for comprehensive air quality models. Numerical experiments performed with a two-dimensional test model show promising results.

1. Introduction

As our understanding of the complex processes in the atmosphere has evolved, comprehensive atmospheric chemistry models have been developed. At the same time, our ability to measure the concentration of various chemical species in the atmosphere has significantly improved over the past decades. A forecast model and observational data are combined in the data assimilation process in order to provide an optimal estimate of the true atmospheric state. Since available observations are usually distributed unevenly and sparsely, the spatiotemporal distribution of the observations plays an essential role in the data assimilation process. While the location of many observations is a priori fixed (such as ground stations and satellite observations), it is often possible to include in an analysis a few additional observations whose locations may be selected in a flexible manner. For example, if additional observations are to be taken using dropsondes from an aircraft mission, a strategy to select the times and locations of these observations must be specified. Adaptive methods search for spatiotemporal locations of the observations that minimize the error in the analysis and subsequent forecasts. In practice the design of adaptive strategies is constrained by the limited number of resources available and physical considerations (e.g., only a limited area may be covered in a given interval of time).

Significant research has been dedicated to the problem of adaptive location of the observations in numerical weather prediction. The experiments performed in the data assimilation context show that the success of the adaptive strategies relies on the identification of the areas where the errors in the initial conditions are large and/or are rapidly growing (usually referred to as target areas). Berliner et al. (1999) provide a rigorous statistical formulation and mathematical framework for the adaptive design problem. The Fronts and Atlantic Storm Tracks Experiment (FASTEX; Joly et al. 1997) and the North Pacific Experiment (NORPEX-1998) provided real-life applications where several methods for targeting observations were tested. A targeting technique based on adjoint sensitivity was formulated by Langland and Rohaly (1996). Palmer et al. (1998) and Buizza and Montani (1999) describe adaptive techniques using the dominant singular vectors of the integral tangent propagator associated with a nonlinear dynamical system. Implementation of the gradient and singular vector methods relies on adjoint modeling. Bishop and Toth (1999) use an ensemble transform technique based on

Corresponding author address: Dr. Dacian N. Daescu, Institute for Mathematics and Its Applications, University of Minnesota, 207 Church Street SE, 400 Lind Hall, Minneapolis, MN 55455.
E-mail: daescu@ima.umn.edu

nonlinear ensemble forecasts to approximate the prediction error covariance matrices. Recently, an ensemble transform Kalman filter method was proposed by Bishop et al. (2001) for the adaptive observations problem. Bishop (2000, manuscript submitted to *Quart. J. Roy. Meteor. Soc.*, hereafter BIS) introduced an extended observation gradient targeting technique and provided a comparative analysis of several adaptive observing methods in the context of estimation theory. Baker and Daley (2000) consider an adjoint-based technique to determine the sensitivity of the forecast to the observations and the background field and apply this method to the adaptive targeting problem.

In the past few years variational methods have been successfully used in data assimilation for comprehensive 3D atmospheric chemistry models (Elbern and Schmidt 1999; Errera and Fonteyn 2001). Satellite observations of the chemical species are beginning to provide global datasets that facilitate an improved understanding of the natural and human influence on the changes in the atmosphere. In addition, valuable measurements are performed during extensive field experiments using aircraft missions, ships, and balloons. The Aerosol Characterization Experiment (ACE-Asia, spring 2001) and National Aeronautics and Space Administration (NASA) Transport and Chemical Evolution over the Pacific Experiment (TRACE-P, spring 2001) have recently produced datasets of the trace gases and aerosol distribution over East Asia and the western Pacific. At the same time, these experiments revealed the necessity of developing computationally feasible methods for targeting observations in air quality modeling. The adjoint sensitivity technique we consider in this paper in the context of the 4D variational data assimilation was inspired by the work of Fisher and Lary (1995) for atmospheric chemistry models. Associated with a single observation, they introduced an ''influence function'' and used it to provide information about the sensitivity of the cost functional with respect to intermediate states of the model. Their study of a stratospheric photochemical box model with trajectories showed that observations of some chemical species may provide useful information about unobserved species. Similar experiments were performed later by Elbern et al. (1997) using a tropospheric chemistry box model. We extend this technique by taking into account the spatial dimension and a general set of parameters, and we use the influence functions to develop an algorithm for the adaptive location of the observational system. An adjoint-based gradient method for a transport-chemistry model is implemented and it is shown that the set parameters can be naturally extended to include emissions and certain types of boundary values with minimal additional cost.

The paper is organized as follows. In section 2 we briefly review the four-dimensional variational data assimilation (4DVAR) for transport-chemistry models. The adjoint sensitivity method is presented in section 3. The problem of the adaptive location of the observations is formulated in section 4 in the 4DVAR context. In section 5 we describe an adjoint sensitivity approach and propose an algorithm for targeting observations. Computational and implementation issues are addressed. Preliminary numerical experiments performed with an Eulerian two-dimensional atmospheric chemistry model are presented in section 6. An outline of the results and further research directions are given in section 7.

2. 4DVAR data assimilation for air quality models

The time evolution of the chemical composition of the atmosphere is determined by various processes such as transport, diffusion, chemical transformations, emissions, and depositions. Using the mass balance equations, the dynamical model is expressed as the coupled system of nonlinear partial differential equations

$$\frac{\partial}{\partial t}c_i = -\nabla \cdot (\mathbf{u}c_i) + \nabla \cdot \left[\rho \mathbf{K} \cdot \nabla\left(\frac{c_i}{\rho}\right)\right] + f_i(\mathbf{c}) + E_i - D_i, \qquad i = \overline{1, s} \tag{1}$$

We consider the spatial domain $\{\mathbf{x} = (x, y, z)\} = \Omega \subset \mathcal{R}^3$ and the analysis time interval $[t_0, T]$. The solution $\mathbf{c}(t, \mathbf{x}) \in \mathcal{R}^s$ of problem (1) represents the concentration vector of the chemical species in the model, $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ is the gradient operator, $\mathbf{u}$ is the wind field, $\mathbf{K}$ is a second-order, diagonal, eddy diffusivity tensor, and the air density is denoted by ρ. We will use the notation $\mathbf{c}_k$ to specify $\mathbf{c}(t_k, \mathbf{x})$ and $c_i(t_k, \mathbf{x})$ for the component of the vector $\mathbf{c}_k$ corresponding to species i in the chemical model. The chemical reactions are modeled by the nonlinear terms $f_i(\mathbf{c}) = P_i(\mathbf{c}) - L_i(\mathbf{c})c_i$ of polynomial form, with $P_i(\mathbf{c})$, $L_i(\mathbf{c})$ the chemical production and loss terms; E_i and D_i represent source and deposition processes, respectively. Since chemical reactions have characteristic timescales that differ by orders of magnitude, the chemistry component introduces stiffness in the system (1). This is an additional difficulty that arises during the numerical integration of (1) as explained by McRae et al. (1982). Space and time dependence is assumed for all terms, but for simplicity the explicit notation is omitted. The initial condition associated with (1) is

$$\mathbf{c}(t_0, \mathbf{x}) = \mathbf{c}_0(\mathbf{x}) \tag{2}$$

Let $\delta\Omega_l$ be the lateral boundary of Ω, $\delta\Omega_0$ the bottom boundary, $\delta\Omega_h$ the top boundary, and $\partial\Omega = \delta\Omega_l \cup \delta\Omega_0 \cup \delta\Omega_h$. One possible specification of the boundary values is

$$\mathbf{c} = \mathbf{c}^{\leftarrow} \quad \forall(t, \mathbf{x}) \in [t_0, T] \times \delta\Omega_l^{\leftarrow} \tag{3}$$

$$\frac{\partial \mathbf{c}}{\partial \mathbf{n}} = 0 \quad \forall(t, \mathbf{x}) \in [t_0, T] \times \delta\Omega_l^{\rightarrow} \tag{4}$$

$$-\mathbf{n}_0 \cdot (\mathbf{K} \cdot \nabla c_i) = Q_i - \nu_i c_i \quad \forall(t, \mathbf{x}) \in [t_0, T] \times \delta\Omega_0, \quad i = \overline{1, s} \tag{5}$$

$$\frac{\partial \mathbf{c}}{\partial z} = 0 \quad \forall(t, \mathbf{x}) \in [t_0, T] \times \delta\Omega_h, \tag{6}$$

where $\delta\Omega_l^{\rightarrow}$ and $\delta\Omega_l^{\leftarrow}$ are the outflow and the inflow regions, respectively, of the lateral boundary,

$$\delta\Omega_l = \delta\Omega_l^{\rightarrow} \cup \delta\Omega_l^{\leftarrow},$$
$$\delta\Omega_l^{\rightarrow} = \{\mathbf{x} \in \delta\Omega_l, \mathbf{u} \cdot \mathbf{n} \geq 0\},$$
$$\delta\Omega_l^{\leftarrow} = \{\mathbf{x} \in \delta\Omega_l, \mathbf{u} \cdot \mathbf{n} < 0\}, \tag{7}$$

where $\mathbf{n}$ denotes the outward unit vector normal to the lateral boundary surface, $\mathbf{n}_0$ is the inward vector normal to the earth's surface, Q_i and ν_i are the surface emission rate and deposition velocity of species i, respectively.

The numerical solution of the problem (1)–(6) is a reliable representation of the evolution of the true atmospheric state provided that accurate values for the various model input parameters are specified. In air quality modeling, uncertainties in the initial state ($\mathbf{c}_0$), emission (E_i), and deposition (D_i) rates, and boundary values ($\mathbf{c}^{\leftarrow}$, Q_i, ν_i), to name only a few, must be considered. In the variational data assimilation, information provided by the observations is used to find an optimal set of model parameters through a minimization process. For a complete description of the various assumptions used by the data assimilation techniques, including the continuum formulation and a probabilistic interpretation, we will refer to Jazwinski (1970), Tarantola (1987), Daley (1991), Cohn (1997), and to the recent work of Wang et al. (2001) for applications of 4D variational data assimilation to atmospheric chemistry. A rigorous mathematical framework of the adjoint parameter estimation, identifiability issues and regularization techniques are presented by Navon (1998). Since for most practical purposes a discrete model must be considered, we formulate a discrete 4DVAR data assimilation problem that takes into account a general set of model parameters. After semidiscretization on a spatial grid $n_x \times n_y \times n_z$, the problem (1)–(6) is written

$$\frac{d\mathbf{c}}{dt} = F(\mathbf{c}, \mathbf{v}) \tag{8}$$

$$\mathbf{c}(t_0, \mathbf{x}) = \mathbf{c}_0(\mathbf{x}), \tag{9}$$

where $\mathbf{v}$ is a time-dependent vector of discrete model parameters uncorrelated with the initial state $\mathbf{c}_0$. The dimension of the discrete state vector $\mathbf{c}(t, \mathbf{x})$ is $N = s \times n_x \times n_y \times n_z$. Using interpolation techniques, the parameters $\mathbf{v}$ are determined by the values at discrete nodes in the time–space domain. We denote by $\mathbf{p} = (\mathbf{c}_0^{\mathrm{T}}, \mathbf{v}^{\mathrm{T}})^{\mathrm{T}}$ the complete set of model parameters and assume that the solution $\mathbf{c}(t, \mathbf{x})$ is uniquely determined once the parameter vector $\mathbf{p}$ is specified. The time integration of the problem (8)–(9) provides the evolution of the state vector

$$\mathbf{c}_{k+1} = \overline{F}_k(\mathbf{c}_k, \mathbf{v}), \qquad k = 0, 1, \ldots, \tag{10}$$

where $\overline{F}_k$ is determined by F and the numerical integration method.

Consider the set of observations

$$O = \{\mathbf{c}_k^o \in R^{n_k}, k = \overline{0, m}\}, \tag{11}$$

which are taken at discrete moments in time t_k, $k = \overline{0, m}$ over the analysis interval and assume that observations are linearly related with the state

$$\mathbf{c}_k^o = \mathbf{H}_k \mathbf{c}_k + \boldsymbol{\epsilon}_k, \tag{12}$$

where the observational operator $\mathbf{H}_k$ is assumed to be state independent and the total observation error $\boldsymbol{\epsilon}_k$ is determined by the measurement error and the error of representativeness (Cohn 1997; Lorenc 1986). The covariance matrix of the total observation error $\mathbf{R}_k = \langle \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^{\mathrm{T}} \rangle$ is assumed to be known. Additional information on the model parameters may be taken into account as a ''background'' estimate $\mathbf{p}^b$ of the true parameter values. In practice the covariance matrix $\mathbf{B} = \langle \boldsymbol{\epsilon}^b \boldsymbol{\epsilon}^{b\mathrm{T}} \rangle$ of the errors in the background estimate is not known and suboptimal approximations are used to provide it. The 4D variational data assimilation seeks to minimize the discrepancy between the model forecast and observations expressed by the cost function

$$\mathcal{J} = \mathcal{J}^b + \mathcal{J}^o = \frac{1}{2}(\mathbf{p} - \mathbf{p}^b)^{\mathrm{T}} \mathbf{B}^{-1} (\mathbf{p} - \mathbf{p}^b) + \frac{1}{2} \sum_{k=0}^{m} (\mathbf{H}_k \mathbf{c}_k - \mathbf{c}_k^o)^{\mathrm{T}} \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{c}_k - \mathbf{c}_k^o), \tag{13}$$

If the solution of the problem (8)–(9) is expressed as a function of the parameters $\mathbf{c}_k = \mathbf{c}(t_k, \mathbf{x}, \mathbf{p})$ the 4DVAR data assimilation problem is formulated

$$\min_{\mathbf{p}} \mathcal{J}(\mathbf{p}). \tag{14}$$

3. Adjoint sensitivity analysis

Mathematical foundations of the adjoint sensitivity for nonlinear dynamical systems and various classes of response functionals are presented by Cacuci (1981a,b). Marchuk (1995) and Marchuk et al. (1996) describe in detail the adjoint modeling and the construction of the adjoint operators for linear and nonlinear atmospheric dynamics. Pudykiewicz (1998) shows the derivation of the continuous adjoint operator for a tracer transport model and an application to source parameters estimation. Further applications of the adjoint sensitivity

analysis to variational data assimilation are presented by Le Dimet et al. (1997).

Associated with the dynamical model (8)–(9) we consider a response functional of the form

$$\mathcal{F} = \langle \mathbf{w}, \mathbf{c}(t_n) \rangle, \tag{15}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathcal{R}^N$, $t_0 \leq t_n \leq T$, and $\mathbf{w}$ is a specified state independent vector of weights. Then

$$\nabla_{\mathbf{p}} \mathcal{F} = \nabla_{\mathbf{p}} \langle \mathbf{w}, \mathbf{c}(t_n) \rangle \tag{16}$$

is the sensitivity of the scalar expression $\langle \mathbf{w}, \mathbf{c} \rangle$at a given moment in time t_n with respect to the parameters $\mathbf{p}$. To a perturbation in the input parameters $\mathbf{p}' = (\mathbf{c}'^{\mathrm{T}}_0, \mathbf{v}'^{\mathrm{T}})^{\mathrm{T}}$ corresponds a perturbation in the response functional

$$\mathcal{F}' = \langle \mathbf{w}, \mathbf{c}'(t_n) \rangle \tag{17}$$

and to first-order approximation the time evolution of the perturbation $\mathbf{c}'$ is obtained by solving the tangent linear model problem

$$\frac{d\mathbf{c}'}{dt} = \mathbf{F}'_{\mathbf{c}}(\mathbf{c}, \mathbf{v})\mathbf{c}' + \mathbf{F}'_{\mathbf{v}}(\mathbf{c}, \mathbf{v})\mathbf{v}' \tag{18}$$

$$\mathbf{c}'(t_0) = \mathbf{c}'_0, \tag{19}$$

where $\mathbf{F}'_{\mathbf{c}}(\mathbf{c}, \mathbf{v})$ and $\mathbf{F}'_{\mathbf{v}}(\mathbf{c}, \mathbf{v})$ are the partial derivatives (Jacobian matrices) of F with respect to the model state and the model parameters $\mathbf{v}$ at time t, respectively. We introduce an adjoint variable $\lambda(t) \in \mathcal{R}^N$, to be defined later for convenience, and multiply (18) in $(\mathcal{R}^N, \langle \cdot, \cdot \rangle)$ by λ, then integrate on $[t_0, t_n]$ to obtain

$$\int_{t_0}^{t_n} \left\langle \lambda, \frac{d\mathbf{c}'}{dt} \right\rangle dt = \int_{t_0}^{t_n} \langle \lambda, \mathbf{F}'_{\mathbf{c}}(\mathbf{c}, \mathbf{v})\mathbf{c}' + \mathbf{F}'_{\mathbf{v}}(\mathbf{c}, \mathbf{v})\mathbf{v}' \rangle \, dt, \tag{20}$$

which may be written using matrix transposition

$$\int_{t_0}^{t_n} \left\langle \lambda, \frac{d\mathbf{c}'}{dt} \right\rangle dt = \int_{t_0}^{t_n} \langle \mathbf{F}'^{\mathrm{T}}_{\mathbf{c}}(\mathbf{c}, \mathbf{v})\lambda, \mathbf{c}' \rangle + \langle \mathbf{F}'^{\mathrm{T}}_{\mathbf{v}}(\mathbf{c}, \mathbf{v})\lambda, \mathbf{v}' \rangle \, dt. \tag{21}$$

After integrating by parts in the left side of (21) and arranging the terms

$$\langle \lambda, \mathbf{c}' \rangle|_{t_0}^{t_n} = \int_{t_0}^{t_n} \left\langle \frac{d\lambda}{dt} + \mathbf{F}'^{\mathrm{T}}_{\mathbf{c}}(\mathbf{c}, \mathbf{v})\lambda, \mathbf{c}' \right\rangle + \langle \mathbf{F}'^{\mathrm{T}}_{\mathbf{v}}(\mathbf{c}, \mathbf{v})\lambda, \mathbf{v}' \rangle \, dt. \tag{22}$$

Therefore, if λ is defined as the solution of the adjoint problem

$$\frac{d\lambda}{dt} = -\mathbf{F}'^{\mathrm{T}}_{\mathbf{c}}(\mathbf{c}, \mathbf{v})\lambda \tag{23}$$

$$\lambda(t_n) = \mathbf{w} \tag{24}$$

we obtain from (17), (19), and (22)

$$\mathcal{F}' = \langle \lambda(t_0), \mathbf{c}'_0 \rangle + \int_{t_0}^{t_n} \langle \mathbf{F}'^{\mathrm{T}}_{\mathbf{v}}(\mathbf{c}, \mathbf{v})\lambda, \mathbf{v}' \rangle \, dt \tag{25}$$

such that the sensitivities of the response functional are

$$\nabla_{\mathbf{c}_0} \mathcal{F} = \lambda(t_0) \tag{26}$$

$$\nabla_{\mathbf{v}} \mathcal{F} = \mathbf{F}'^{\mathrm{T}}_{\mathbf{v}}(\mathbf{c}, \mathbf{v})\lambda. \tag{27}$$

The adjoint problem (23)–(24) must be integrated backward in time to obtain the sensitivity with respect to the initial state as $\lambda(t_0)$. It is essential to notice that during this process all the values $\lambda(t)$, $t_0 \leq t \leq t_n$ are computed, and using (27) the sensitivities with respect to the time-dependent parameters $\mathbf{v}$ are obtained.

For the purpose of this paper, we will assume that the model (8) takes the particular form

$$\frac{d\mathbf{c}}{dt} = F(\mathbf{c}) + \mathbf{G}\mathbf{v}, \tag{28}$$

where $\mathbf{G} = \mathbf{G}(t)$ is a state independent matrix operator. In this formulation, emission and deposition rates and certain boundary values, such as $\mathbf{c}^{\leftarrow}$ and Q, may be considered in the vector of parameters $\mathbf{v}$. For example, when emission rates are considered, $\mathbf{G}$ is a diagonal matrix with nonzero entries corresponding to the index of the emitted chemical species and their spatial locations. Corresponding to (28), Eq. (27) is written

$$\nabla_{\mathbf{v}} \mathcal{F} = \mathbf{G}^{\mathrm{T}}\lambda \tag{29}$$

such that the sensitivities with respect to the time-dependent parameters $\mathbf{v}$ are obtained at a minimal additional cost.

4. The adaptive observations problem

The adaptive observation problem is presented next in the 4DVAR context. For the remaining part of this paper, unless otherwise specified, by ''location'' of an observation we will understand the 4D coordinate $(t, \mathbf{x})$ of the observation. An observation will be considered as ''located'' if both the time and the spatial coordinate of this observation are determined. To formulate the adaptive location of the observations problem, let

$$O^f = \cup_{k=1}^{m} O^f_k \tag{30}$$

represent the set of observations whose location over the analysis interval $[t_0, T]$ is fixed and a priori known at the initial moment t_0. We consider a ''verification'' domain $\mathcal{D}_v \subseteq \Omega$ and the verification time $t_v > T$. We assume that at discrete instants in time $t_0 \leq t_i \leq T$, $i = 1, I$ it is possible to take n_i additional observations, which must be selected from the set of all possible locations where additional observational resources may be deployed at moment t_i. If the set of all feasible spatial locations is $\mathcal{L}_i$, then a subset of n_i locations $\mathcal{L}^o_i \subset \mathcal{L}_i$ must be selected. An adaptive observational strategy searches for a selection of an observational path $O_p = \{\mathcal{L}^o_1, \mathcal{L}^o_2, \ldots, \mathcal{L}^o_I\}$ such that the solution of the corre-

sponding 4DVAR will minimize the error of some aspect of the forecast at the verification time t_v over the verification domain $\mathcal{D}_v$. The problem can be generalized to fully take into account the time coordinate by allowing the time instant t_i, $i = 1, I$ to be selected from a feasible time set $\mathcal{T}$.

Several practical aspects must be taken into account while designing the adaptive strategy and we outline a few of them. Since the actual values of the observations that will be taken at fixed and adaptive locations are not known at the moment when the decision must be made they cannot be included in the adaptive strategy. The adaptive strategy must take into account the influence of the observations that will be taken from the fixed locations. The relationship between the selection of the locations at different moments in time must be considered. In particular, regarded independently, $\mathcal{L}_i^o$ and $\mathcal{L}_j^o$, $i \neq j$ may be feasible selections, but $\{\mathcal{L}_i^o, \mathcal{L}_j^o\}$ is not. The notion of a feasible set of locations must then take into account the temporal interdependence, and the order in which the adaptive observations are selected becomes important. An attempt to globally search for an optimal solution from the set of all feasible paths may easily lead to a problem that is computationally impractical. BIS discusses computational and practical aspects related to several adaptive strategies and shows with a simple example that for moderately complex practical applications a serial observation processing must be considered. In the adaptive strategy method and the algorithm we describe in the next section, information provided by the observations at fixed locations is globally taken into account, while the adaptive observations are selected sequentially. If the adaptive selected observational path is denoted by $\mathcal{O}^a$, the forecast at the verification time t_v is obtained by integrating (8)–(9) on $[t_0, t_v]$ with the parameter values $\mathbf{p}^*$ the solution of the problem

$$\min_{\mathbf{p}} (\mathcal{J}^b + \mathcal{J}^{of} + \mathcal{J}^{oa})(\mathbf{p}). \tag{31}$$

5. An adjoint method for adaptive observations

In the adjoint sensitivity approach a cost functional is defined as the measure of the forecast error at the verification time t_v over the verification domain $\mathcal{D}_v$

$$\mathcal{J}^v = \frac{1}{2}\langle \mathbf{P}(\mathbf{c}_v - \mathbf{c}_v^{\text{ref}}), \mathbf{P}(\mathbf{c}_v - \mathbf{c}_v^{\text{ref}})\rangle, \tag{32}$$

where $\mathbf{c}_v$ and $\mathbf{c}_v^{\text{ref}}$ represent, respectively, the model forecast and the verifying analysis at t_v, and $\mathbf{P}$ is a self-adjoint projection operator on the verification domain. In practice the inner product $\langle \cdot, \cdot \rangle$ defines an appropriate energy norm (Rabier et al. 1996; Palmer et al. 1998). For simplicity, we will consider $\mathbf{P}$ to be a diagonal matrix with entries corresponding to the grid points inside the verification domain being equal to one while the other entries are set to zero.

Sensitivity fields

$$\nabla_{\mathbf{c}_l}\mathcal{J}^v = \left(\frac{\partial \mathbf{c}_v}{\partial \mathbf{c}_l}\right)^{\mathrm{T}} \mathbf{P}(\mathbf{c}_v - \mathbf{c}_v^{\text{ref}}) \tag{33}$$

provided by the adjoint model may be used to identify the areas where errors in the model state at t_l have the most significant impact on the forecast error at t_v over the verification domain. By providing additional observational data in the areas where the sensitivity field has a large magnitude is expected to obtain maximum benefit in reducing the forecast error over $\mathcal{D}_v$. However, evaluation of the gradient (33) requires explicit knowledge of the verifying analysis $\mathbf{c}_v^{\text{ref}}$, which is not known at the initial time t_0. For practical purposes, the sensitivity field used to select targeted observations must be based on the forecast alone such that information provided by the gradient fields (33) may be used only for a posteriori analysis and adaptive observations design (Rabier et al. 1996).

Baker and Daley (2000) noticed that traditional adaptive strategies are based on the a priori evaluation of the sensitivity of some aspect of the forecast at the verification time to intermediate states and are completely ignorant of any existing observations. Since 4DVAR data assimilation takes into consideration *all* observations available in the assimilation window, targeted observations strategies must account for uncertainty magnitude, uncertainty growth, *and* the details of the data assimilation scheme (Bergot 2001).

The adjoint targeting strategy we propose is based on the evaluation of two sensitivity fields: the first sensitivity field, $\mathbf{\Gamma}_v$, is associated to the verification cost functional $\mathcal{J}^v$ and information provided by $\mathbf{\Gamma}_v$ does not account for any existing observations; the second sensitivity field, $\mathbf{\Gamma}_o$, is associated with the cost functional used in the data assimilation process. Here $\mathbf{\Gamma}_o$ is dynamically updated and takes into consideration information from all routine (fixed location) and adaptive observations whose locations are already determined. Next, we present a rigorous definition and the interaction mechanism between these two sensitivity fields that are used to develop a new adaptive observations algorithm.

a. The influence function at the verification time

The sensitivity field we associate to the verification cost functional (32) is intimately related to the evaluation of the sensitivity of forecast errors to initial conditions as described by Rabier et al. (1996). To eliminate the explicit dependence on $\mathbf{c}_v^{\text{ref}}$, we define an ''influence function'' as a measure of the sensitivity of forecast errors to relative changes in the model state at intermediate instants in time.

We consider the model forecast $\mathbf{c}_v$ as a function of the model state $\mathbf{c}_l$ at $t_l < t_v$ and $\mathcal{J}^v = \mathcal{J}^v(\mathbf{c}_l)$. If we consider an infinitesimal variation $\delta c_i(t_l, \mathbf{x})$ in species

i at instant t_l and location $\mathbf{x}$, then the induced variation in the cost functional (32) may be expressed

$$\delta \mathcal{J}^v = (\nabla_{\mathbf{c}_l} \mathcal{J}^v)_{(i,\mathbf{x})} \delta c_i(t_l, \mathbf{x}) = \{(\nabla_{\mathbf{c}_l} \mathcal{J}^v)_{(i,\mathbf{x})} c_i(t_l, \mathbf{x})\} \frac{\delta c_i(t_l, \mathbf{x})}{c_i(t_l, \mathbf{x})} \quad (34)$$

such that

$$(\nabla_{\mathbf{c}_l} \mathcal{J}^v)_{(i,\mathbf{x})} c_i(t_l, \mathbf{x}) = \sum_{j=1}^{s} \sum_{\mathbf{x}^* \in \mathcal{D}_v} \frac{\partial c_j(t_v, \mathbf{x}^*)}{\partial c_i(t_l, \mathbf{x})} \times [c_j(t_v, \mathbf{x}^*) - c_j^{\text{ref}}(t_v, \mathbf{x}^*)] c_i(t_l, \mathbf{x}) \quad (35)$$

represents the sensitivity of the cost functional to relative changes in the parameter $c_i(t_l, \mathbf{x})$. For each $\mathbf{x}^* \in \mathcal{D}_v$, the sensitivity of the cost functional with respect to relative changes in the forecast component $c_j(t_v, \mathbf{x}^*)$ is

$$(\nabla_{\mathbf{c}_v} \mathcal{J}^v)_{(j,\mathbf{x}^*)} c_j(t_v, \mathbf{x}^*) = [c_j(t_v, \mathbf{x}^*) - c_j^{\text{ref}}(t_v, \mathbf{x}^*)] c_j(t_v, \mathbf{x}^*). \quad (36)$$

Assuming that the verification domain is reduced to one grid point, $\mathcal{D}_v = \{\mathbf{x}^*\}$, and the verifying analysis has only one component $c_j^{\text{ref}}(t_v, \mathbf{x}^*)$, we define the influence function as the normalized quantity

$$\Gamma_{j,t_v,\mathbf{x}^*}(i, t_l, \mathbf{x}) = \frac{(\nabla_{\mathbf{c}_l} \mathcal{J}^v)_{(i,\mathbf{x})} c_i(t_l, \mathbf{x})}{(\nabla_{\mathbf{c}_v} \mathcal{J}^v)_{(j,\mathbf{x}^*)} c_j(t_v, \mathbf{x}^*)} = \frac{\partial c_j(t_v, \mathbf{x}^*)}{\partial c_i(t_l, \mathbf{x})} \frac{c_i(t_l, \mathbf{x})}{c_j(t_v, \mathbf{x}^*)}, \quad (37)$$

which is independent of the verifying analysis $c_i^{\text{ref}}(t_v, \mathbf{x}^*)$. We will interpret the value of the influence function as a measure of the sensitivity of the forecast error of $c_j(t_v, \mathbf{x}^*)$ to relative changes in the parameter $c_i(t_l, \mathbf{x})$. From this definition it follows that

$$\Gamma_{j,t_v,\mathbf{x}^*}(i, t_v, \mathbf{x}) = \delta_{ij} \delta_{\mathbf{x}\mathbf{x}^*}, \quad (38)$$

where δ represents the Kronecker delta function $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. A straightforward extension to the case when the verification domain includes multiple grid points and the verifying analysis has various components is presented in section 5d.

b. The influence function of a single observation

The second sensitivity field we define is associated with the cost functional (13) to be minimized in the data assimilation process.

If we consider the truncated cost function (13) at moment t_l as a function of $\mathbf{c}_l$ only

$$\mathcal{J}_l(\mathbf{c}_l) = \frac{1}{2} \sum_{k=l}^{m} (\mathbf{H}_k \mathbf{c}_k - \mathbf{c}_k^o)^{\mathrm{T}} \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{c}_k - \mathbf{c}_k^o) \quad (39)$$

then

$$(\nabla_{\mathbf{c}_l} \mathcal{J}_l)_{(i,\mathbf{x})} c_i(t_l, \mathbf{x}) = \sum_{k=l}^{m} \left\langle \mathbf{H}_k^{\mathrm{T}} \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{c}_k - c_k^o), \left(\frac{\partial \mathbf{c}_k}{\partial c_i(t_l, \mathbf{x})} \right) \right\rangle c_i(t_l, \mathbf{x}) \quad (40)$$

represents the sensitivity of the cost functional to relative changes in the parameter $c_i(t_l, \mathbf{x})$. Assuming that there is only a single observation

$$c_j^0(t_n, \mathbf{x}^*) = \langle \mathbf{H}_n^{\mathrm{T}}(j, \mathbf{x}^*), \mathbf{c}_n \rangle + \epsilon_n(j, \mathbf{x}^*) \quad (41)$$

of species j at moment $t_n \geq t_l$ and point $\mathbf{x}^*$, the influence function is defined, by analogy with (37), as the normalized quantity

$$\Gamma_{j,t_n,\mathbf{x}^*}(i, t_l, \mathbf{x}) = \frac{(\nabla_{\mathbf{c}_l} \mathcal{J}_l)_{(i,\mathbf{x})} c_i(t_l, \mathbf{x})}{(\nabla_{\mathbf{c}_n} \mathcal{J}_l)_{(j,\mathbf{x}^*)} c_j(t_n, \mathbf{x}^*)}, \quad (42)$$

which can be written explicitly

$$\Gamma_{j,t_n,\mathbf{x}^*}(i, t_l, \mathbf{x}) = \frac{\left\langle \mathbf{H}_n^{\mathrm{T}}(j, \mathbf{x}^*), \frac{\partial \mathbf{c}_n}{\partial c_i(t_l, \mathbf{x})} \right\rangle c_i(t_l, \mathbf{x})}{\left\langle \mathbf{H}_n^{\mathrm{T}}(j, x^*), \frac{\partial c_n}{\partial c_j(t_n, \mathbf{x}^*)} \right\rangle c_j(t_n, \mathbf{x}^*)}. \quad (43)$$

From this definition it follows that the influence function is independent of the observation value and the observation error, and is determined only by the forecast state and the observational operator $\mathbf{H}_n$ which is known. The relation (38) holds also for $\Gamma_{j,t_n,\mathbf{x}^*}(i, t_n, \mathbf{x})$. We will interpret the value of the influence function (43) as the sensitivity of the model fit to the observation of species c_j at $(t_n, \mathbf{x}^*)$ with respect to relative changes in the species c_i at $(t_l, \mathbf{x})$. A large absolute value of the influence function for species i due to an observation of species j indicates that observations of species j play an important part in determining the analyzed values for species i (Fisher and Lary 1995).

To simplify the presentation, we will assume for the remaining part of this paper that the concentrations of the chemical species are directly observed at the model grid points such that the observational operator $\mathbf{H}_n(j, \mathbf{x}^*)$ has only a nonzero entry on the $(j, \mathbf{x}^*)$ coordinate

$$\mathbf{H}_n(j, \mathbf{x}^*) = (0, \ldots, 0, 1, 0, \ldots, 0). \quad (44)$$

The explicit expression (43) of the influence function is then written

$$\Gamma_{j,t_n,\mathbf{x}^*}(i, t_l, \mathbf{x}) = \frac{\partial c_j(t_n, \mathbf{x}^*)}{\partial c_i(t_l, \mathbf{x})} \frac{c_i(t_l, \mathbf{x})}{c_j(t_n, \mathbf{x}^*)}. \quad (45)$$

The definition of the influence function is naturally extended with respect to the model parameters $\mathbf{v}_l$. Corresponding to the $(i, \mathbf{x})$ component of $\mathbf{v}_l(t_k)$, $t_l \leq t_k \leq t_n$, we define

$$\Gamma_{j,t_n,\mathbf{x}^*}(i, t_k, \mathbf{x}) = \frac{\partial c_j(t_n, \mathbf{x}^*)}{\partial v_i(t_k, \mathbf{x})} \frac{v_i(t_k, \mathbf{x})}{c_j(t_n, \mathbf{x}^*)}. \quad (46)$$

The influence function with respect to $\mathbf{v}_l$ has then a time-

distributed value (46) for $t_l \le t_k \le t_n$ and the analog of the relation (38) is in this case

$$\Gamma_{j,t_n,\mathbf{x}^*}(i,\, t_n,\, \mathbf{x}) = 0. \tag{47}$$

c. Adjoint computation of the influence function

The complexity of evaluating the influence function is dominated by the computation of the sensitivity values $\nabla_{\mathbf{c}_l} c_j(t_n,\, \mathbf{x}^*)$ and $\nabla_{\mathbf{v}_l} c_j(t_n,\, \mathbf{x}^*)$. The adjoint method provides an efficient way to compute *at once* the influence function $\Gamma_{j,t_n,\mathbf{x}^*}(i,\, t_l,\, \mathbf{x})$ with respect to *all* chemical species i in the model and *all* spatial points $\mathbf{x}$. With a single backward integration of the adjoint model (23) the vector value $\mathbf{\Gamma}_{j,t_n,\mathbf{x}^*}(t_l)$ (still refered to as influence function) may be computed.

Using relation (29), while computing the influence function with respect to the model state, we obtain the influence function $\mathbf{\Gamma}_{j,t_n,\mathbf{x}^*}(t_k)$, $t_l \le t_k \le t_n$ of the time-dependent model parameters with a minimal additional cost. Therefore, it is sufficient to focus our analysis on the influence function with respect to the model state.

In practice, a discrete adjoint model is often implemented directly from the numerical method used in the integration of the forward model. In this approach reverse automatic differentiation tools, for example, Tangent Linear and Adjoint Model Compiler (TAMC) (Giering 1997), may be used to facilitate the adjoint code generation. If $\mathbf{x}^*$ has the grid coordinates $\mathbf{x}^*(ix,\, iy,\, iz)$, let $\mathbf{e}$ be the $(j,\, ix,\, iy,\, iz)$ vector of the canonical base of $R^s \times R^{n_x} \times R^{n_y} \times R^{n_z}$: $\mathbf{e}(j,\, ix,\, iy,\, iz) = 1$ and all other components of $\mathbf{e}$ are zero. Assume that the forecast state at t_n is obtained from the state at t_l through a sequence of q intermediate time steps $t_l = t_l^0 < t_l^1 < \cdots t_l^q = t_n$. The adjoint method to evaluate the gradient $\nabla_{\mathbf{c}_l} c_j(t_n,\, \mathbf{x}^*)$ is implemented as the backward loop:

Initialize $\nabla_{\mathbf{c}_l} c_j(t_n,\, \mathbf{x}^*) = \mathbf{e}$; for $k = q,\, 1,\, -1$

$$\nabla_{\mathbf{c}_l} c_j(t_n,\, \mathbf{x}^*) = \left(\frac{\partial \mathbf{c}_{l^k}}{\partial \mathbf{c}_{l^{k-1}}}\right)^{\mathrm{T}} \nabla_{\mathbf{c}_l} c_j(t_n,\, \mathbf{x}^*). \tag{48}$$

Once the gradient $\nabla_{\mathbf{c}_l} c_j(t_n,\, \mathbf{x}^*)$ is computed, the value of the influence function is easily evaluated using (45). Observe that the computation of $\nabla_{\mathbf{c}_l} c_j(t_n,\, \mathbf{x}^*)$ provides also the values of the intermediate gradients with respect to $\mathbf{c}_{l^k}$, $k = 1,\, q$. In particular, while computing the value of the influence function with respect to the initial state, $\mathbf{\Gamma}_{j,t_n,\mathbf{x}^*}(t_0)$, spatial and temporal sensitivity information is provided with respect to all chemical species in the model.

d. Domain of influence and multiple observations

The influence function was defined in the case when a single observation was taken into account and its computation required only one integration of the adjoint model. However, in practice it is often the case that multiple observations are available at a moment t_n. In this section we extend the definition of the influence function (42) to include a set of observations. By a simple analogy, the extension applies also to the definition (37).

Consider a set O_n of observations at moment t_n, and assume that O_n has at least two elements. The previous definition (42) of the influence function can not be applied in this case since it will involve the observation values. We define the influence function associated with the set O_n as

$$\Gamma_{O_n}(i,\, t_l,\, \mathbf{x}) = \sum_{(j,t_n,\mathbf{x}^*)\in O_n} \omega_{j,t_n,\mathbf{x}^*} \Gamma_{j,t_n,\mathbf{x}^*}(i,\, t_l,\, \mathbf{x}) \tag{49}$$

with the positive weights $\omega_{j,t_n,\mathbf{x}^*}$ to be specified as convenient. From the computational point of view, the evaluation of $\mathbf{\Gamma}_{O_n}$ has the same complexity as the evaluation of $\mathbf{\Gamma}$ and its computation requires only one backward integration [same trajectory as in (48)]

$$\begin{aligned}
\text{Initialize} \qquad \Gamma_{O_n} &= \sum_{(j,t_n,\mathbf{x}^*)\in O_n} \frac{\omega_{j,t_n,\mathbf{x}^*}}{c_j(t_n,\, \mathbf{x}^*)} \mathbf{e}(j,\, \mathbf{x}^*) \\
&\qquad \text{for } k = q,\, 1,\, -1 \\
\Gamma_{O_n} &= \left(\frac{\partial \mathbf{c}_{l^k}}{\partial \mathbf{c}_{l^{k-1}}}\right)^{\mathrm{T}} \Gamma_{O_n} \\
\Gamma_{O_n}(i,\, t_l,\, \mathbf{x}) &= \Gamma_{O_n}(i,\, t_l,\, \mathbf{x}) c_i(t_l,\, \mathbf{x}).
\end{aligned} \tag{50}$$

The extension to the case when multiple observations are considered at different moments in time is straightforward and may be obtained by periodically adding an initialization term during the backward loop. Therefore, if $t_l \le t_m \le t_n$, evaluating $\mathbf{\Gamma}_{O_n \cup O_m}(t_l)$ has roughly the same complexity as evaluating $\mathbf{\Gamma}_{O_n}(t_l)$.

We define the ''domain of influence'' associated with Γ, and respectively $\mathbf{\Gamma}$, as

$$\mathcal{D}_{j,t_n,\mathbf{x}^*}(i,\, t_l) = \{\mathbf{x} \in \Omega \,|\, \Gamma_{j,t_n,\mathbf{x}^*}(i,\, t_l,\, \mathbf{x}) \neq 0\} \tag{51}$$

$$\tilde{\mathcal{D}}_{j,t_n,\mathbf{x}^*}(t_l) = \{\mathbf{x} \in \Omega \,|\, \mathbf{\Gamma}_{j,t_n,\mathbf{x}^*}(t_l,\, \mathbf{x}) \neq 0\}. \tag{52}$$

From this definition it follows that

$$\tilde{\mathcal{D}}_{j,t_n,\mathbf{x}^*}(t_l) = \cup_i \mathcal{D}_{j,t_n\mathbf{x}^*}(i,\, t_l) \tag{53}$$

and we can rewrite (49) as

$$\mathbf{\Gamma}_{O_n}(t_l) = \sum_{(j,t_n,\mathbf{x}^*)\in O_n} \chi_{\tilde{\mathcal{D}}_{j,t_n,\mathbf{x}^*}(t_l)} \omega_{j,t_n,\mathbf{x}^*} \mathbf{\Gamma}_{j,t_n,\mathbf{x}^*}(t_l), \tag{54}$$

where $\chi_{\tilde{\mathcal{D}}}$ is the characteristic function of the set $\tilde{\mathcal{D}}$. We will consider that two observations $(j_1,\, t_n,\, \mathbf{x}_1^*)$ and, respectively, $(j_2,\, t_n,\, \mathbf{x}_2^*)$ have an independent influence at moment t_l if

$$\tilde{\mathcal{D}}_{j_1,t_n,\mathbf{x}_1^*}(t_l) \cap \tilde{\mathcal{D}}_{j_2,t_n,\mathbf{x}_2^*}(t_l) = \varnothing. \tag{55}$$

A useful relationship between $\mathbf{\Gamma}_{O_n}$ and $\mathbf{\Gamma}$ can be shown in the case of sparse observations. Assume that any two observations in the set O_n have an independent influence. It follows then that the sum in the right-hand side of (54) has at most one nonzero term and, therefore, all information provided by any of $\mathbf{\Gamma}$ can be obtained from

$\mathbf{\Gamma}_{O_n}$. This may be particularly useful for the adaptive observations problem since in general we are interested in areas where only sparse observations are available.

It must be noted that in the case when multiple observations have a common domain of influence it is possible that when additional observations are considered the magnitude of $\mathbf{\Gamma}_{O_n}$ will decrease. A possible solution is to replace the definition (49) by

$$\Gamma_{O_n}(i, t_l, \mathbf{x}) = \sum_{(j,t_n,\mathbf{x}^*)\in O_n} \omega_{j,t_n,\mathbf{x}^*}|\Gamma_{j,t_n,\mathbf{x}^*}(i, t_l, \mathbf{x})|. \quad (56)$$

However, in this case, in order to compute Γ_{O_n} each function $\Gamma_{j,t_n,\mathbf{x}^*}$ must be evaluated individually, and this results in significant computational expense. The matrix approach to the adjoint sensitivity analysis as described by Ustinov (2001) offers a promising technique to efficiently evaluate the influence function (56).

e. An algorithm for adaptive observations

The 4D variational data assimilation takes into account all observations available in the analysis time interval. Therefore, as pointed out in section 4, the interaction between observations must be considered in the observations selection algorithm. We assume that the data assimilation analysis interval is $[t_0, T]$ and at instant $t_i \in [t_0, T]$, $i = 1, I$ a set O^i of n_i observations must be selected from the set of all possible observations at time t_i. The target area $\mathcal{A}_i = \mathcal{A}|_{t=t_i}$ is located inside the domain of influence $\tilde{\mathcal{D}}_i = \tilde{\mathcal{D}}_{t_v}(t_i)$ of the influence function $\mathbf{\Gamma}_v(t_i)$ associated with the verification cost functional $\mathcal{J}^v$. The algorithm for adaptive location of the observations that we propose in this section searches for locations where the magnitude of the influence function $\mathbf{\Gamma}_v$ is maximal, conditioned by the information accumulated from all observations whose locations were already determined. The selection of the observations is sequential in time and proceeds backward during the adjoint integration. If O^f denotes the set of observations at fixed locations, O^a denotes the adaptive observational path to be selected, and $O = O^f \cup O^a$, then the algorithm may be outlined as follows:

ALGORITHM FOR ADAPTIVE SELECTION OF THE OBSERVATIONS

$O^a = \varnothing$

$O = O^f$

Evaluate $\mathbf{\Gamma}_v(t_0)$! provides also $\tilde{\mathcal{D}}_i$ and $\mathbf{\Gamma}_v(t_i)$, $i = I, 1, -1$

Evaluate $\mathbf{\Gamma}_O(t_0)$! provides also $\mathbf{\Gamma}_O(t_i)$, $i = I, 1, -1$

FOR $i = I, 1, -1$

$\mathbf{\Gamma}_v = \mathbf{\Gamma}_v \dfrac{|\mathbf{\Gamma}_v|}{|\mathbf{\Gamma}_v| + |\mathbf{\Gamma}_O|}$! pointwise in $\tilde{\mathcal{D}}_i$ at t_i

CALL *findmax*$(n_i, \mathbf{\Gamma}_v(t_i), O^i)$

$O^a = O^a \cup O^i$

Evaluate $\mathbf{\Gamma}_{O^a}(t_{i-1})$! further explained below

$\mathbf{\Gamma}_O(t_{i-1}) = \mathbf{\Gamma}_O(t_{i-1}) + \mathbf{\Gamma}_{O^a}(t_{i-1})$

$O = O^f \cup O^a$

The subroutine *findmax*( ) defines the adaptive observations set O^i as the locations corresponding to the first n_i maximal absolute values of the updated influence function $\mathbf{\Gamma}_v(t_i)$. The evaluation of $\mathbf{\Gamma}_{O^a}(t_{i-1})$ requires an update of the adjoint variables at t_i to include O^i followed by a backward integration from t_i to t_{i-1}. By periodically updating the influence functions $\mathbf{\Gamma}_v$ and $\mathbf{\Gamma}_O$, the algorithm takes into account the cumulative influence of all observations that are already located. The updated influence function $\mathbf{\Gamma}_v$ is inversely proportional to the value of $\mathbf{\Gamma}_O$ such that the new observations are located in regions where the sensitivity of the forecast to parameters is large *and* little additional information may be obtained from previously located observations. The

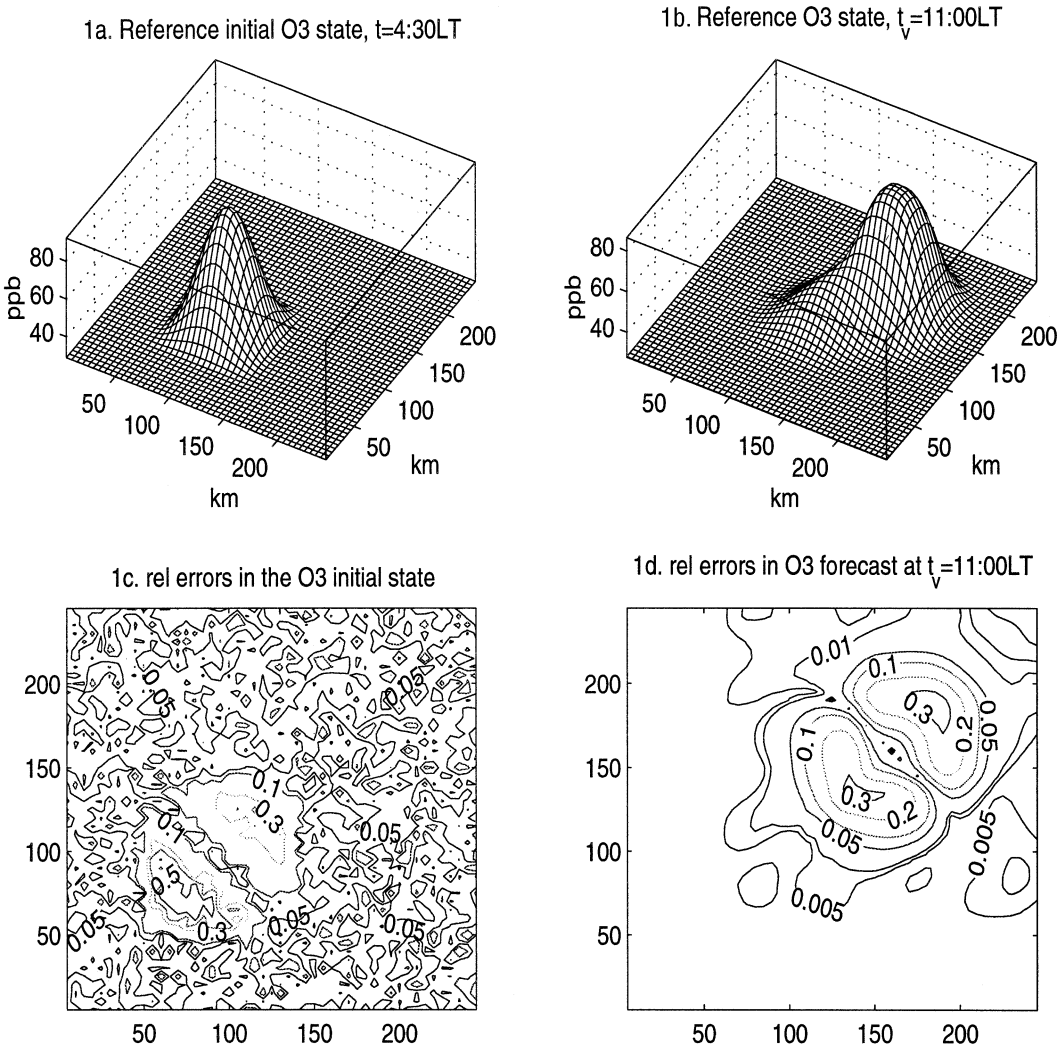

FIG. 1. (a) The reference initial state of ozone, t = 0430 LST; (b) the reference state of ozone at t_v = 1100 LST; (c),(d) isopleths of the relative errors (absolute values) in the initial guess ozone state and in the corresponding forecast at t_v.

computational cost is dominated by the evaluation of the influence functions such that the CPU time required for implementation is roughly

$$\text{CPU} \sim \text{CPU}([t_v, t_0]) + 2 \times \text{CPU}([T, t_0]), \quad (57)$$

where CPU($[t, t_0]$) is the CPU time of the adjoint model integration from t to t_0. In practice CPU ($[t, t_0]$) is a small factor (2–5) of the CPU time of the forward integration from t_0 to t (Griewank 2000) such that for most applications (57) is an acceptable complexity. Moreover, parallel processing may be used to reduce the computational cost; for example, the initial evaluation of $\mathbf{\Gamma}_v(t_0)$ and $\mathbf{\Gamma}_o(t_0)$ may be done in parallel, sharing the same forward trajectory storage. Additional memory resources must be allocated during the evaluation of the functions $\mathbf{\Gamma}_v(t_0)$ and $\mathbf{\Gamma}_o(t_0)$ to store all the intermediate values $\mathbf{\Gamma}_v(t_i)$, $\mathbf{\Gamma}_o(t_i)$, $i = I, 1, -1$. No claim is made here that the selected path is optimal among all the possible paths. However, we provided in an efficient way a good candidate. In the next section we implement this algorithm and present numerical experiments for a two-dimensional transport-chemistry model.

6. Numerical experiments

a. The test model

The numerical experiments were performed with a two-dimensional test model based on the Carbon Bond Mechanism IV (Gery et al. 1989) with 32 variable chemical species involved in 70 thermal and 11 photolytic reactions. The spatial domain is [0 250] km × [0 250] km and a uniform grid $\Delta x = \Delta y = 5$ km is considered,

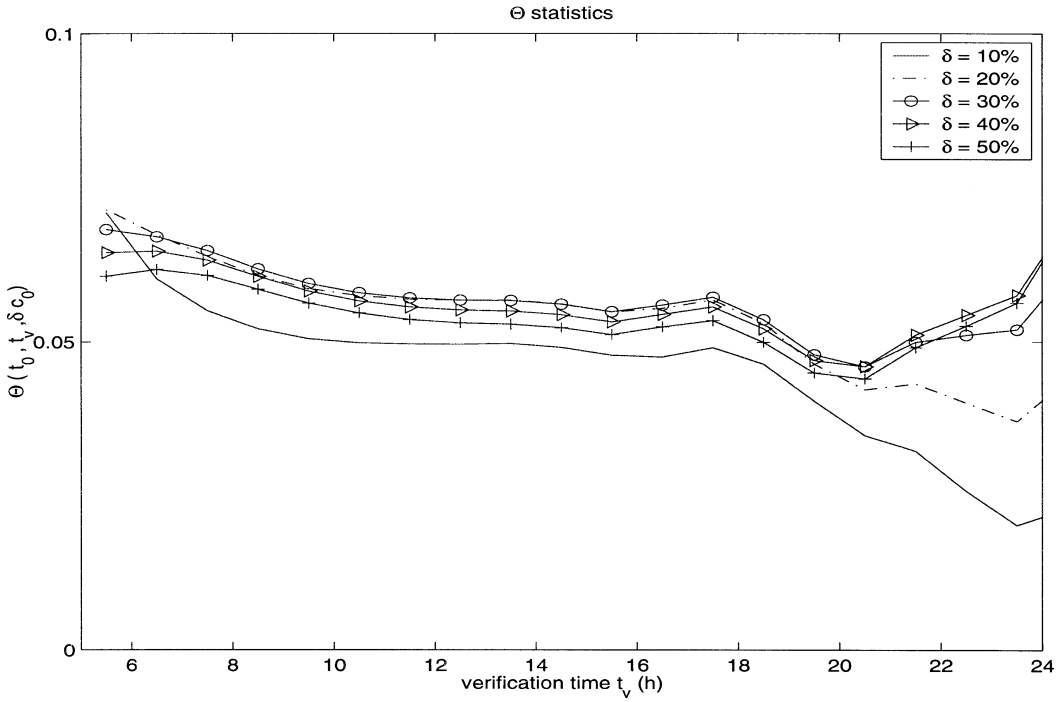

FIG. 2. The Θ statistics test for increasing longer forecast time. A random field is used to generate the initial perturbation and examples are shown for a magnitude of the intial perturbation $\delta\mathbf{c}_0$ up to 50% of the control initial state.

such that there are 49 × 49 interior grid points and the dimension of the discrete state vector is $N = 76\,832$. The wind field $\mathbf{u}(u_x, u_y)$ and the diffusion coefficient $\mathbf{K}(K_{xx}, K_{yy})$ are taken constant $u_x = u_y = 10$ km h^{-1}, $K_{xx} = K_{yy} = 10^{-3}$ km^2 s^{-1}. The initial state distribution and emissions values are obtained using the box model urban and rural scenarios described by Sandu et al. (1997). An urban region is considered in the domain [50 150] km × [50 150] km. At the center of the urban area, (100, 100) km, we consider the initial state and emissions as in the urban scenario. Outside the urban area rural initial conditions and emissions are specified. Interpolation is done between the center of the urban area and the urban boundaries to obtain the initial state and emissions inside the urban region. Emission rates are constant in time and no deposition terms are considered. Boundary conditions are prescribed according to (3) and (4) with the inflow boundary values $\mathbf{c}^{\leftarrow}$ obtained from a box model integration with rural initial conditions. The advection operator is discretized using a limited $k = 1/3$ upwind flux interpolation as presented by Koren (1993) and the diffusion operator using central differences formula. The time integration of the semidiscrete model is performed using dimensional Strang operator splitting (Strang 1968) with a splitting time step $2\Delta t = 30$ min. The advection–diffusion terms are integrated using a second-order explicit Runge–Kutta method and the chemistry-source terms using a variable step size second-order L-stable Rosenbrock method ROS-2 (Verwer et al. 1999). The reference state of ozone at the initial time $t_0 = 0430$ local standard time (LST) and at the verification time $t_v = 1100$ LST are shown in Figs. 1a and 1b, respectively. An initial guess state for the model was obtained by shifting SW two grid points from the initial reference state. After shifting, random errors up to 20% are introduced in the ozone state. Isopleths of the relative errors (absolute values) in the initial guess ozone state and the corresponding forecast state at t_v are shown in Figs. 1c and 1d, respectively. We notice that large errors in the forecast state at t_v are located around the area [100 200] km × [100 200] km with maximal errors at the locations $\mathbf{x}_1^* = (135, 135)$ km and $\mathbf{x}_2^* = (190,185)$ km.

A discrete adjoint model was generated using the adjoint model compiler TAMC (Giering 1997) for the advection–diffusion numerical integration and symbolic processing for the chemistry integration as described by Daescu et al. (2000). In the numerical experiments we present, the restriction of the verification time to $t_v = 1100$ LST was determined by the high storage requirements of the adjoint model implementation and by the limited computational resources available. For the two-dimensional example we consider, the discrete model state is a three-dimensional vector, $\mathbf{c}_k(s, n_x, n_y)$, and the adjoint code requires manipulation of four-dimensional vectors as the complete forward trajectory is required and the sensitivity fields are time dependent. To overcome this difficulty, we used a two-level checkpointing scheme (Daescu et al. 2000) to store the forward trajectory such that there are two forward integrations per

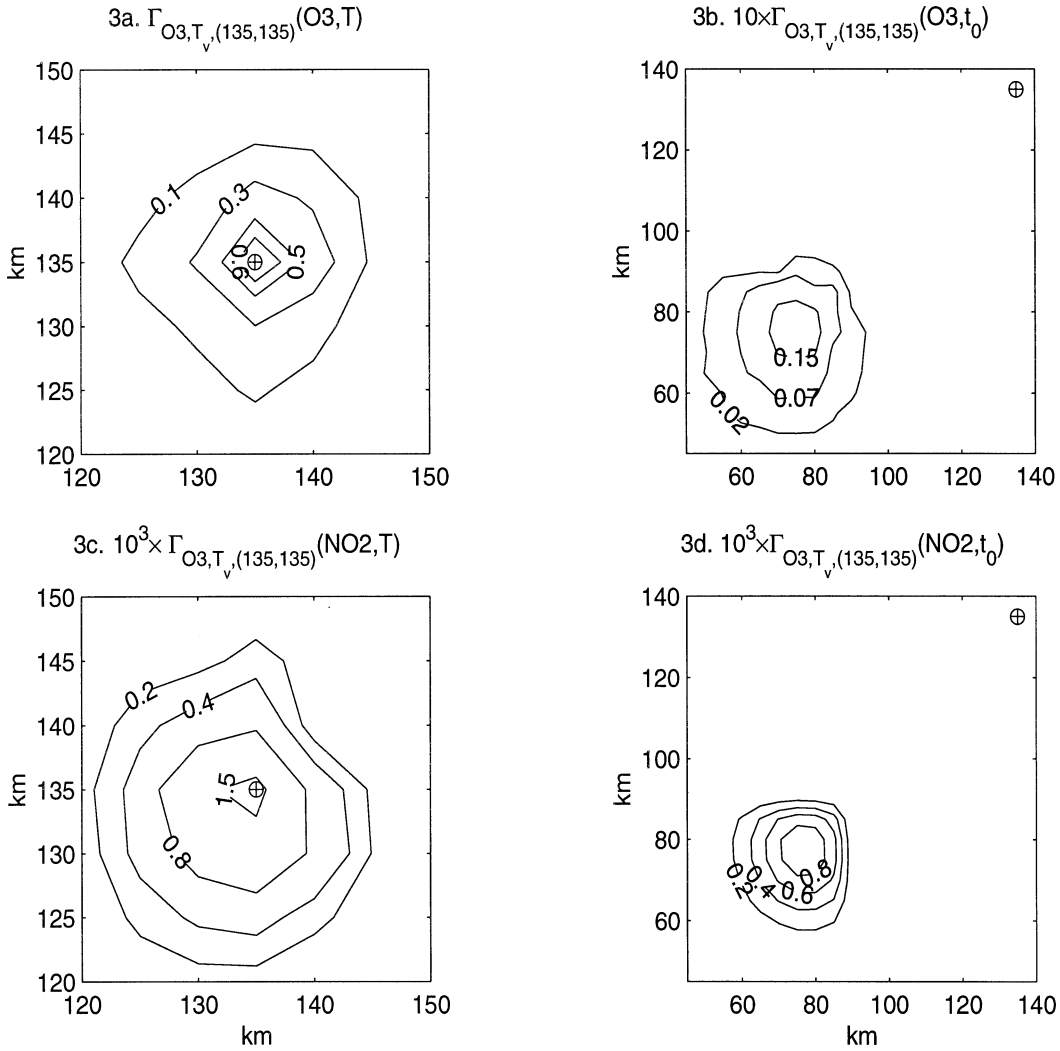

FIG. 3. Examples of influence functions for an O_3 observation located at (t_v, $\oplus$): (a),(c) with respect to O_3 and NO_2 state at T, respectively; (b),(d) with respect to O_3 and NO_2 state at t_0, respectively. Isopleths of the magnitude are shown and notice that a different scaling factor is used for each plot.

backward integration. During the first forward run we store the trajectory after each operator splitting step (30 min). During the second forward run, we store the trajectory inside each operator splitting step (the chemistry integration takes on average 8–10 steps for a 30-min interval). In double precision, our computational resources (HP-UX A 9000/778) allow only manipulations of vectors with dimension up to $\sim 10^6$ and only 14 states may be stored in fast memory. With a 30-min operator splitting step, this limits our analysis interval to 6 h, 30 min, from 0430 to 1100 LST.

b. Validity of the model linearization

The variational data assimilation and the adjoint sensitivity analysis rely on the linearization of the forward model. Therefore, it is essential to establish the validity of the linear approximation before these methods can be used. Hansen and Smith (2000) show that the validity of the linear approximation is crucial for adaptive strategies based on the linear propagator to be productive. In this section we follow the approach of Hansen and Smith (2000) and consider a Θ statistic test that is defined by examining the evolution of twin perturbations about a control trajectory. Using the notation (10), if $\mathbf{c}(t) = \overline{F}_t(\mathbf{c}_0)$, $t_0 < t$ represents the control trajectory initiated from $\mathbf{c}_0 = \mathbf{c}(t_0)$, we consider two additional trajectories: $\mathbf{c}^+(t) = \overline{F}_t(\mathbf{c}_0^+)$ initiated from $\mathbf{c}_0^+ = \mathbf{c}_0 + \delta\mathbf{c}_0$ and $\mathbf{c}^- t = \overline{F}_t(\mathbf{c}_0^-)$ initiated from $\mathbf{c}_0^- = \mathbf{c}_0 - \delta\mathbf{c}_0$. The degree to which the linear approximation of $\overline{F}$ holds at time t can be quantified as

$$\Theta(t_0, \delta\mathbf{c}_0, t) = \frac{\|\delta\mathbf{c}^+(t) + \delta\mathbf{c}^-(t)\|}{\frac{1}{2}(\|\delta\mathbf{c}^+(t)\| + \|\delta\mathbf{c}^-(t)\|)}, \quad (58)$$

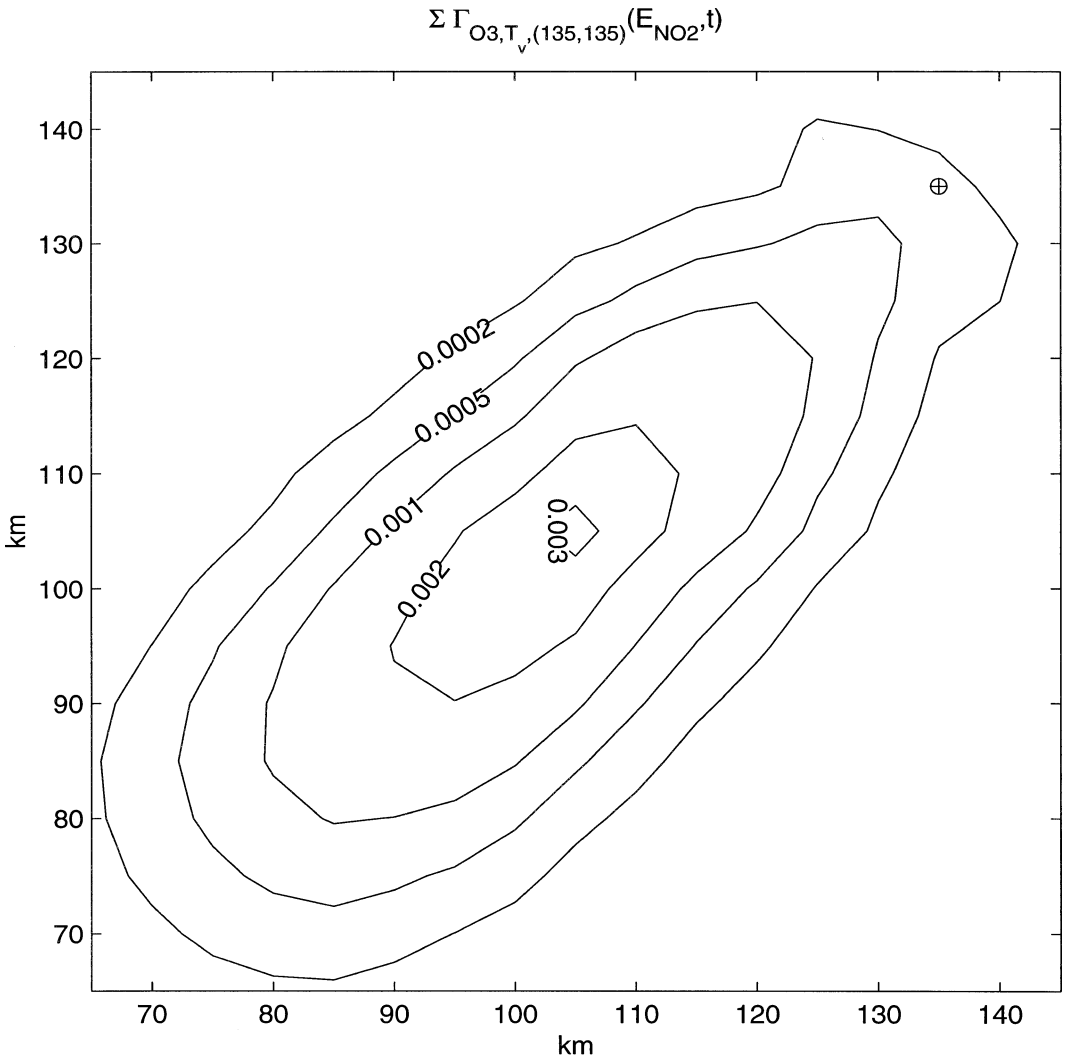

FIG. 4. Cumulative value of the influence function for an O_3 observation located at (t_v, $\oplus$) with respect to NO_2 emissions in the time interval $[t_0, t_v]$. Isopleths of the magnitude are shown.

where $\delta\mathbf{c}^+(t) = \mathbf{c}^+(t) - \mathbf{c}(t)$ and $\delta\mathbf{c}^-(t) = \mathbf{c}^-(t) - \mathbf{c}(t)$. When the linear approximation is exact ($\overline{F}$ is linear), $\Theta = 0$, $\forall t$, $\delta\mathbf{c}_0$, while when $\Theta = 1$ the errors associated with the linear approximation are equal in magnitude to the evolved perturbations themselves [see Hansen and Smith (2000) for details]. The evolution of Θ as a function of the verification time and the magnitude of the initial uncertainty $\delta\mathbf{c}_0$ is shown in Fig. 2 and indicates that the linear approximation is valid out to $t_v = 24$ h. For a stratospheric photochemical box model, Khattatov et al. (1999) performed a rigorous analysis of the validity of the model linearization using the tangent linear propagator and showed that the linear approximation remains valid for several days.

c. Examples of influence functions for ozone

To illustrate the influence functions and their domain of influence, first we consider a fixed location in the spatiotemporal domain at $t_v = 1100$ LST and $\mathbf{x}_1^* = (135, 135)$ km. The sensitivities of ozone (O_3) at (t_v, $\mathbf{x}_1^*$) with respect to relative changes in the model state and NO_2 emissions in the time interval $[t_0, T]$, $T = 1030$ LST are obtained by evaluating the influence function $\boldsymbol{\Gamma}_{O_3,t_v,\mathbf{x}_1^*}(t_l)$, $t_0 \le t_l \le T$. As shown in section 5, these values are obtained with a single backward integration of the adjoint model. Isopleths of the magnitude of the influence function $\boldsymbol{\Gamma}_{O_3,t_v,\mathbf{x}_1^*}(T)$ for O_3 and NO_2 are displayed in Figs. 3a and 3c, respectively, and it can be seen that the domain of influence is located near the observational point $\mathbf{x}_1^*$. On the other hand, as shown in Figs. 3b and 3d, the influence function with respect to the initial state $\boldsymbol{\Gamma}_{O_3,t_v,\mathbf{x}_1^*}(t_0)$ has the domain of influence located in a region far away from the observational point. An analysis of the influence function values shows that the ozone forecast state is highly sensitive to changes in the ozone state at intermediate instants in time, whereas the sensitivity with respect to NO_2 has a

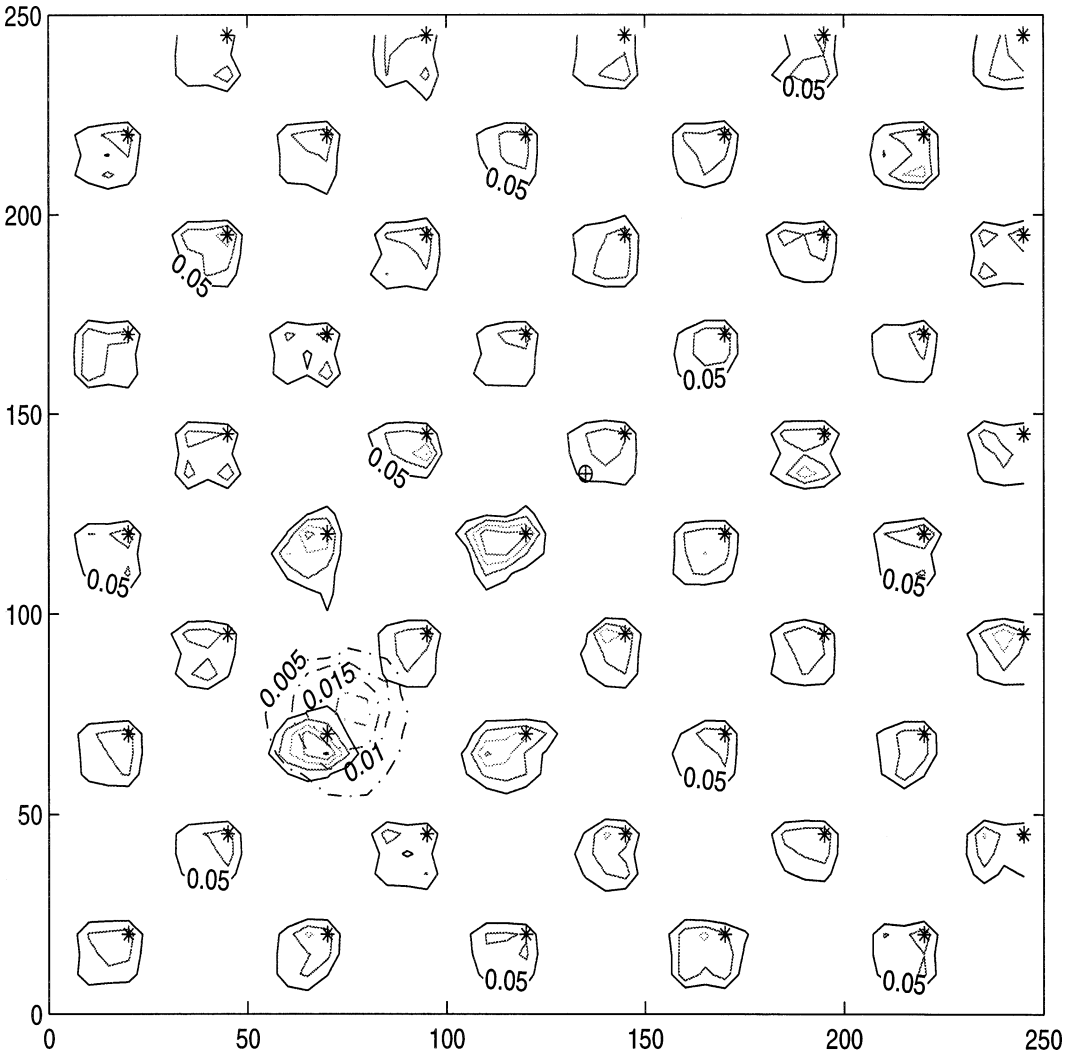

FIG. 5. The location and the influence function $\mathbf{\Gamma}_{O^f,t_1}(O_3, t_0)$ of the set of fixed observations $\mathcal{O}^f$ (marked *) at t_1 = 0500 LST with respect to the initial O_3 state. The verifying analysis at t_v = 1100 LST is located at (135,135) km (marked $\oplus$). Dashed lines show the influence function $\mathbf{\Gamma}_v(O_3, t_0)$. Isopleths of the magnitude are displayed.

much smaller magnitude. The information provided by the adjoint integration allows a similar analysis with respect to all chemical species in the model. The sensitivity with respect to NO_2 emissions is given by the time evolution of the influence function $\mathbf{\Gamma}_{O_3,t_v,\mathbf{x}_1^*}(E_{NO_2}, t_l)$, $t_0 \le t_l \le T$. If we are interested in identifying the regions where changes in NO_2 emissions over the time interval $[t_0, T]$ will influence the ozone forecast state, we may consider the cumulative value

$$\mathbf{\Gamma}_{O_3,t_v,\mathbf{x}_1^*}(E_{NO_2}, [t_0, T]) = \sum_{t=t_0}^{T} |\mathbf{\Gamma}_{O_3,t_v,\mathbf{x}_1^*}(E_{NO_2}, t)|. \quad (59)$$

The isopleths of the expression (59) are plotted in Fig. 4 and show a sensitive region located inside the square [90 110] km × [90 110] km.

d. Adaptive observations and 4DVar

The data assimilation experiment is set using model-generated data (twin experiments) over a 6-h interval $[t_0, T]$ = [0430 1030] LST with ''observations'' provided for ozone only. The error covariance matrices $\mathbf{R}_k$ are taken to be the identity matrix. The set of control parameters is considered to be the initial state of the model, $\mathbf{p} = \mathbf{c}_0$, and no background term is included in the cost functional. The initial guess state is obtained as explained in section 6a and the reference run represents the ''true'' atmospheric state $\mathbf{c}^t$. Fixed observations $\mathcal{O}^f$ for ozone are considered at t_1 = 0500 LST only at locations marked by ''*'' in Fig. 5. Isopleths of the magnitude of the influence function $\mathbf{\Gamma}_{O^f}(O_3, t_0)$ associated with the fixed observations are also displayed in Fig. 5. We assume that from t_0 to T every half-hour

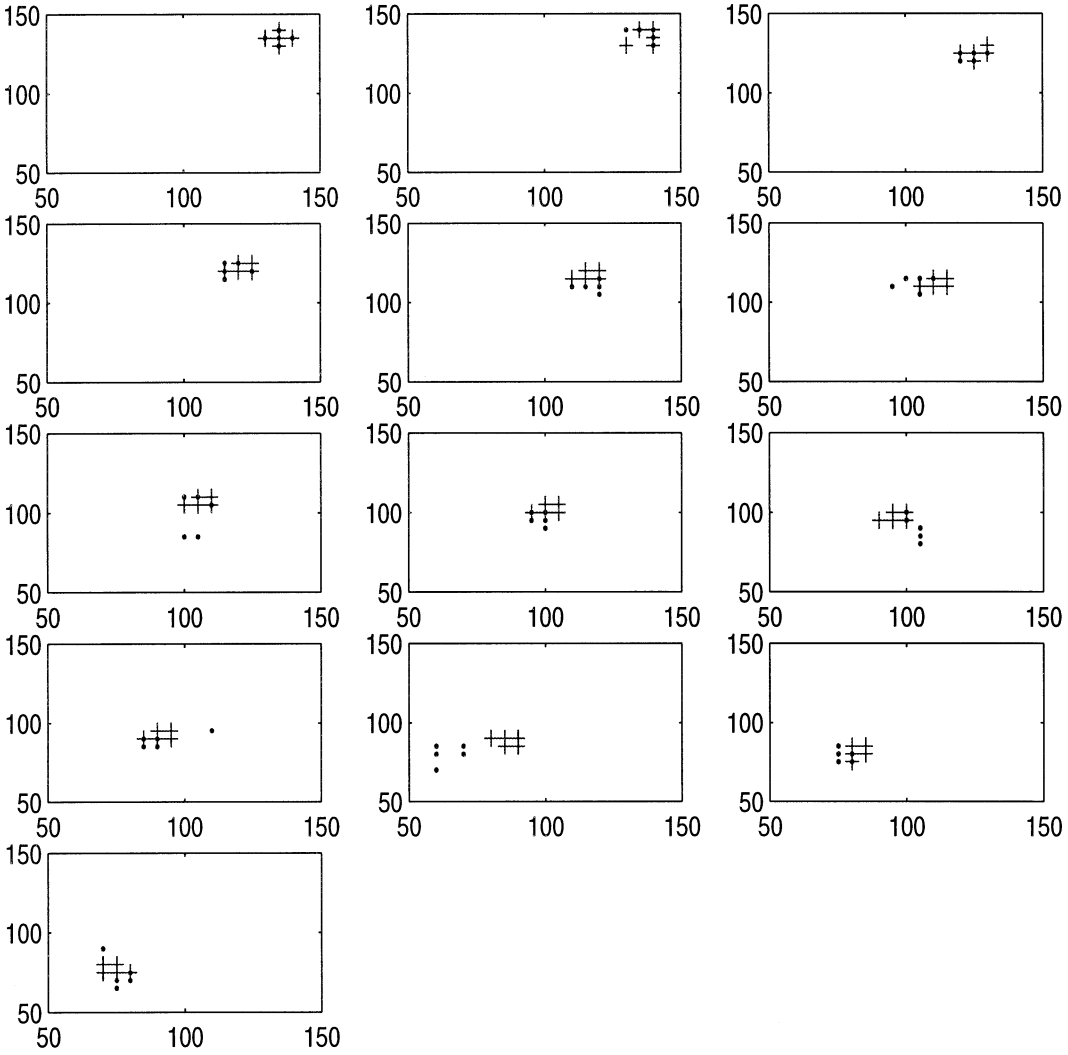

FIG. 6. The adaptive location of the observational path. Observations selected by method M1 (marked +) and the observations selected by the adaptive method M2 (marked •). The time is updated each half-hour and moves backward on the rows from the upper-left corner (t = 1030 LST) to the lower-left corner (t = 0430 LST).

five additional observations may be provided and their optimal location must be determined. Since twin experiments are performed, the evolution of the true state ($\mathbf{c}^t$) is known. The goal of the experiment is to select an observational path such that the solution provided by the 4DVAR data assimilation will minimize the forecast error for ozone at t_v = 1100 LST defined by the functional

$$\mathcal{J}^v = [\mathrm{O}_3^f(t_v, \mathbf{x}_1^*) - \mathrm{O}_3^t(t_v, \mathbf{x}_1^*)]^2, \qquad (60)$$

where $\mathbf{x}_1^*$ = (135, 135) km is the location where the forecast error at t_v was determined to be maximal. In Fig. 5 this location is marked with "⊕" and the isopleths of the influence function $\mathbf{\Gamma}_v(\mathrm{O}_3, t_0)$ are also shown with dotted line.

Two methods are tested for the adaptive observations: the first method (M1) is based only on the a priori evaluation of the sensitivity field $\mathbf{\Gamma}_v$ and fits in the traditional adjoint sensitivity framework. The influence function $\mathbf{\Gamma}_v$ is evaluated once, with no update, and the observations at t_i are always located at the points where the magnitude of $\mathbf{\Gamma}_v(t_i)$ is maximal. These

TABLE 1. The CPU time (s) of the forward and adjoint integration and the CPU time (s) required to implement the methods M1 and M2. An additional forward integration time is included in the CPU time of the backward integration.

CPU ($[t_0, t_v]$)	CPU ($[t_v, t_0]$)	CPU ($\mathbf{\Gamma}_{O^f}$)	CPU ($\mathbf{\Gamma}_{O^a}$)	CPU (M1)	CPU (M2)
10.1	30.6	2.6	30.7	40.8	74.3

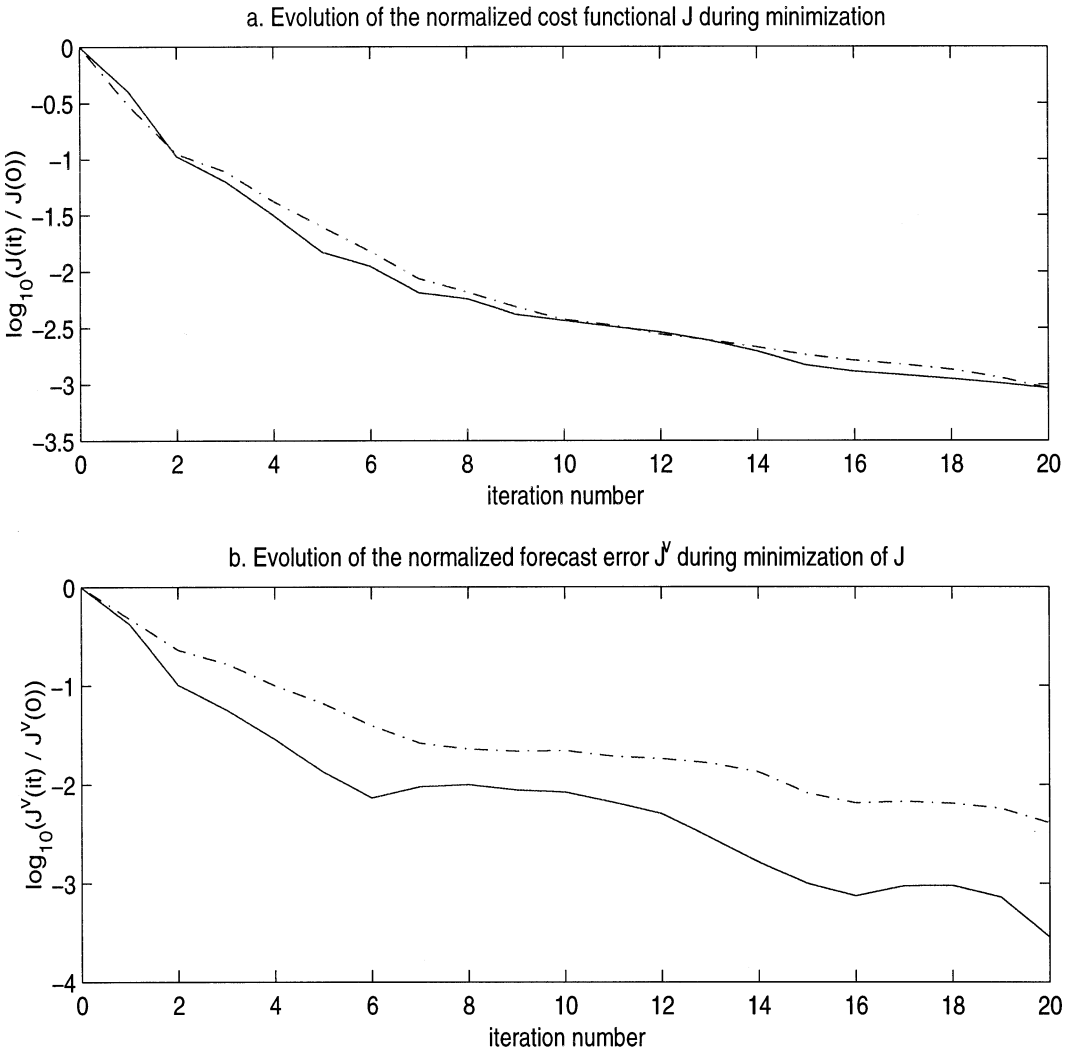

FIG. 7. The optimization process. (a) Evolution of the cost functional. (b) Evolution of the ozone forecast error at t_v = 1100 LST. Normalized values are shown on a $\log_{10}$ scale with dashed line for method M1 and with solid line for the proposed adaptive method M2. While both methods provide the same relative reduction in $\mathcal{J}$, the adaptive method M2 provides a much smaller forecast error at t_v.

locations are marked in Fig. 6 with "+" as time goes backward moving on the rows from the upper left corner ($t = T$) to the lower left corner ($t = t_0$). The second method (M2) implements the algorithm presented in section 5d with a periodic update of the values of $\mathbf{\Gamma}_v$. The selected locations are marked in Fig. 6 with a solid dot (locations marked in Fig. 6 by "+" and "·" were selected by both M1 and M2 methods). The computational cost (as CPU time) of the forward and adjoint model integration and to implement each of the methods M1 and M2 is presented in Table 1. The CPU time required to implement method M1 is dominated by the expense of evaluating $\mathbf{\Gamma}_v$, which is roughly given by the cost of a forward–backward integration. The CPU time to implement method M2 requires in addition the evaluation of the influence function of the fixed observations $\mathbf{\Gamma}_{Of}$, CPU $\sim$ CPU($[t_1, t_0]$), and a succesive evaluation/update of the influence function of the adaptive observations $\mathbf{\Gamma}_{Of}$. For each selected path the limited memory L-BFGS method (Liu and Nocedal 1989) is used to minimize the corresponding functional (31) until a reduction $\mathcal{J}_{\text{opt}}/\mathcal{J}_{\text{init}} = 10^{-3}$ is achieved. The evolution of the cost functional (31) during the minimization process in shown in Fig. 7a with dashed line for method M1 and with solid line for method M2. At the same time we monitor the forecast error at t_v by evaluating the functional (60) at each iteration. The evolution of the forecast error is shown in Fig. 7b with dashed line for method M1 and with solid line for method M2. It can be seen that the adaptive strategy method M2 consistently provides a much more accurate forecast at t_v than the method M1.

7. Conclusions and further research

Strategies for targeting observations have been considered mostly in numerical weather prediction and applications to atmospheric chemistry are at a very incipient stage. The problem of the adaptive location of the observations in atmospheric chemistry research becomes increasingly important as transport-chemistry models begin to be used in forecast mode to enhance flight planning during large-scale field experiments. Expensive field-deployed resources can be utilized more effectively and the science success can be maximized by selecting an optimal observational path.

Strategies for targeting observations must take into account the properties of the data assimilation algorithm. With the current computing resources, variational methods based on adjoint modeling may be used to perform data assimilation for comprehensive atmospheric chemistry models. We described an adjoint sensitivity method and applied it to the problem of adaptive selection of the observations for a transport-chemistry model. Our results show that using the adjoint approach, sensivities with respect to various model parameters such as emission and deposition rates or boundary values may be obtained at a reduced computational cost. The influence functions associated with the observations and their domain of influence were shown to be essential tools in developing a strategy for adaptive observations in the 4D variational data assimilation context. At the same time, our results indicate that by using a periodical update of the sensitivity values to include the influence from all previously located observations, an observational path with significant benefits for the model forecast may be determined. The novel algorithm for adaptive observations we presented may be efficiently implemented at a computational cost equivalent with the cost of a few forward model integrations and our preliminary numerical experiments show promising results. Further research is needed to implement this algorithm for a comprehensive 3D Sulfate Transport Eulerian Model (STEM) model (Carmichael et al. 1986); test the algorithm performance on real observational datasets; and apply the new adaptive technique to future field experiments. Future work will also include a comparative study with targeting methods using the dominant singular vectors and an analysis of the interaction between the information provided by the ''background'' parameter estimation and adaptive observations.

Acknowledgments. This work was supported in part by funds from the National Science Foundation, under the Information Technology Research program.

REFERENCES

Baker, N. L., and R. Daley, 2000: Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Quart. J. Roy. Meteor. Soc.*, **126,** 1431–1454.

Bergot, T., 2001: Influence of the assimilation scheme on the efficiency of adaptive observations. *Quart. J. Roy. Meteor. Soc.*, **127,** 635–660.

Berliner, L. M., Z.-Q. Lu, and C. Snyder, 1999: Statistical design for adaptive weather observations. *J. Atmos. Sci.*, **56,** 2536–2552.

Bishop, C. H., and Z. Toth, 1999: Ensemble transformation and adaptive observations. *J. Atmos. Sci.*, **56,** 1748–1765.

——, B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129,** 420–436.

Buizza, R., and A. Montani, 1999: Targeted observations using singular vectors. *J. Atmos. Sci.*, **56,** 2965–2985.

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.

Carmichael, G. R., L. K. Peters, and T. Kitada, 1986: A second generation model for regional-scale transport/chemistry/deposition. *Atmos. Environ.*, **20,** 173–188.

Cohn, S. E., 1997: An introduction to estimation theory. *J. Meteor. Soc. Japan*, **75B,** 257–288.

Daescu, D. N., G. R. Carmichael, and A. Sandu, 2000: Adjoint implementation of Rosenbrock methods applied to variational data assimilation problems. *J. Comput. Phys.*, **165,** 496–510.

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press, 457 pp.

Elbern, H., and H. Schmidt, 1999: A four-dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling. *J. Geophys. Res.*, **104** (D15), 18 583–18 598.

——, ——, and A. Ebel, 1997: Variational data assimilation for tropospheric chemistry modeling. *J. Geophys. Res.*, **102** (D13), 15 967–15 985.

Errera, Q., and D. Fonteyn, 2001: Four-dimensional variational chemical assimilation of CRISTA stratospheric measurements. *J. Geophys. Res.*, **106** (D11), 12 253–12 265.

Fisher, M., and D. J. Lary, 1995: Lagrangian four-dimensional variational data assimilation of chemical species. *Quart. J. Roy. Meteor. Soc.*, **121,** 1681–1704.

Gery, M. W., G. Z. Whitten, J. P. Killus, and M. C. Dodge, 1989: A photochemical kinetics mechanism for urban and regional scale computer modeling. *J. Geophys. Res.*, **94,** 12 925–12 956.

Giering, R., cited 1997: Tangent linear and adjoint model compiler, users manual 1.2. [Available online at http://puddle.mit.edu/~ralf/tamc.]

Griewank, A., 2000: *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation.* Frontiers in Applied Mathematics, Vol. 19, SIAM, 369 pp.

Hansen, J. A., and L. A. Smith, 2000: The role of operational constraints in selecting supplementary observations. *J. Atmos. Sci.*, **57,** 2859–2871.

Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory.* Academic Press, 376 pp.

Joly, A., and Coauthors, 1997: The Fronts and Atlantic Storm-Track Experiment (FASTEX): Scientific objectives and experimental design. *Bull. Amer. Meteor. Soc.*, **78,** 1917–1940.

Khattatov, B. V., J. C. Gille, L. V. Lyjak, G. P. Brasseur, V. L. Dvortsov, A. E. Roche, and J. Waters, 1999: Assimilation of photochemically active species and a case analysis of UARS data. *J. Geophys. Res.*, **104,** 18 715–18 737.

Koren, B., 1993: A robust upwind discretization method for advection, diffusion and source terms. *Numerical Methods for Advection–Diffusion Problems*, C. B. Vreugdenhil and B. Koren, Eds., Notes on Numerical Fluid Mechanics, Vol. 45, Vieweg, 117–137.

Langland, R., and G. Rohaly, 1996: Adjoint-based targeting of observations for FASTEX cyclones. Preprints, *Seventh Conf. on Mesoscale Process*, Reading, United Kingdom, Amer. Meteor. Soc., 369–371.

Le Dimet, F.-X., H. E. Ngodock, B. Luong, and J. Verron, 1997: Sensitivity analysis in variational data assimilation. *J. Meteor. Soc. Japan*, **75B,** 245–255.

Liu, D. C., and J. Nocedal, 1989: On the limited memory BFGS method for large scale minimization. *Math. Prog.*, **45,** 503–528.

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

Marchuk, G. I., 1995: *Adjoint Equations and Analysis of Complex Systems.* Kluwer Academic, 466 pp.

——, I. V. Agoshkov, and P. V. Shutyaev, 1996: *Adjoint Equations and Perturbation Algorithms in Nonlinear Problems.* CRC Press, 288 pp.

McRae, G. J., W. R. Goodin, and J. H. Seinfeld, 1982: Numerical solution of the atmospheric diffusion equation for chemically reactive flows. *J. Comput. Phys.,* **45,** 1–42.

Navon, I. M., 1998: Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dyn. Atmos. Oceans,* **27,** 55–79.

Palmer, T. N., R. Gelaro, J. Barkmeijer, and R. Buizza, 1998: Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.,* **55,** 633–653.

Pudykiewicz, J. A., 1998: Application of the adjoint tracer transport equations for evaluating source parameters. *Atmos. Environ.,* **32,** 3039–3050.

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.

Sandu, A., J. G. Verwer, M. Loon, G. R. Carmichael, A. F. Potra, D. Dabdub, and J. H. Seinfeld, 1997: Benchmarking stiff ODE solvers for atmospheric chemistry problems. I: Implicit versus explicit. *Atmos. Environ.,* **31,** 3151–3166.

Strang, G., 1968: On the construction and comparison of difference schemes. *SIAM J. Numer. Anal.,* **5,** 506–517.

Tarantola, A., 1987: *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation.* Elsevier Science, 613 pp.

Ustinov, E. A., 2001: Adjoint sensitivity analysis of atmospheric dynamics: Application to the case of multiple observables. *J. Atmos. Sci.,* **58,** 3340–3348.

Verwer, J. G., E. J. Spee, J. G. Blom, and W. H. Hundsdorfer, 1999: A second-order Rosenbrock method applied to photochemical dispersion problems. *SIAM J. Sci. Comput.,* **20,** 1456–1480.

Wang, K. Y., D. J. Lary, D. E. Shallcross, S. M. Hall, and J. A. Pyle, 2001: A review on the use of the adjoint method in four-dimensional atmospheric-chemistry data assimilation. *Quart. J. Roy. Meteor. Soc.,* **127B,** 2181–2204.