

Research Article

Forward Sensitivity Approach to Dynamic Data Assimilation

S. Lakshmivarahan¹ and J. M. Lewis^{2,3}

¹ School of Computer Sciences, University of Oklahoma, Norman, OK 73072, USA

² Forecast R&D, National Severe Storms Laboratory, Norman, OK 73072, USA

³ Division of Atmospheric Sciences, Desert Research Institute, Reno, NV 89512, USA

Correspondence should be addressed to S. Lakshmivarahan, varahan@ou.edu

Received 3 December 2009; Accepted 15 February 2010

Academic Editor: Zhaoxia Pu

Copyright © 2010 S. Lakshmivarahan and J. M. Lewis. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The least squares fit of observations with known error covariance to a strong-constraint dynamical model has been developed through use of the time evolution of sensitivity functions—the derivatives of model output with respect to the elements of control (initial conditions, boundary conditions, and physical/empirical parameters). Model error is assumed to stem from incorrect specification of the control elements. The optimal corrections to control are found through solution to an inverse problem. Duality between this method and the standard 4D-Var assimilation using adjoint equations has been proved. The paper ends with an illustrative example based on a simplified version of turbulent heat transfer at the sea/air interface.

1. Introduction

Sensitivity function analysis has proved valuable as a mean to both build models and to interpret their output in chemical kinetics (Rabitz et al. [1], Seefeld and Stockwell [2]) and air quality modeling (Russell et al. [3]). Yet, the ubiquitous systematic errors that haunt dynamical prediction cannot be fully understood with sensitivity functions alone. We now include an optimization component that leads to an improved fit of model to observations. The methodology is termed forward sensitivity method (FSM)—a method based on least squares fit of model to data, but where algorithmic structure and correction procedure are linked to the sensitivity functions. In essence, corrections to control (the initial conditions, the boundary conditions, and the physical and empirical parameters) are found through solution to an inverse problem.

In this paper we derive the governing equations for corrections to control and show their equivalence to equations governing the so-called 4D-Var assimilation method (four-dimensional variational method)—least squares fit of model to observations under constraint (LeDimet and Talagrand [4]). Beyond this equivalence, we demonstrate the value of the FSM as a diagnostic tool that can be used to understand the relationship between sensitivity and correction to control.

We begin our investigation by laying down the dynamical framework for the FSM: general form of the governing dynamical model, the type and representation of model error that can be identified through the FSM, and the evolution of the sensitivity functions that are central to execution of the FSM. The dual relationship between 4D-Var/adjoint equations is proved. The step-by-step process of assimilating data by FSM is outlined, and we demonstrate its usefulness by application to a simplified air-sea interaction model.

2. Foundation Dynamics for the FSM

We have included a list of mathematical symbols used in this paper. These symbols and associated nomenclature are found in Table 1.

2.1. Prediction Equations. Let $x(t) \in R^n$ denote the state and let $\alpha \in R^p$ denote the parameters of a deterministic dynamical system, where $x(t) = (x_1(t), x_2(t), \dots, x_n(t))^T$ and $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_p)^T$ are column vectors, n and p are positive integers, $t \geq 0$ denotes the time, and superscript T denotes the transpose of the vector or matrix. Let $f : R^n \times R^p \times R \rightarrow R^n$ be a mapping, where $f(x, \alpha, t) = (f_1, f_2, \dots, f_n)^T$ with $f_i = f_i(x, \alpha, t)$ for $1 \leq i \leq n$. The vector spaces R^n and R^p are called the model space and parameter space, respectively.

TABLE 1: Symbolism and nomenclature.

Symbol	Nomenclature
n	Dimension of state vector
m	Dimension of observation vector
p	Dimension of parameter vector
$q = n + p$	Dimension of control vector
N	Number of observation vectors
t	Time
$\hat{x}(t) \in R^n$	True model state vector
$x(t) = (x_1, x_2, \dots, x_n)^T \in R^n$	Model state vector
$f : R^n \times R^p$	Vector field of the model
$x(0) \in R^n$	Initial condition for model state vector
$\alpha \in R^p$	Parameter vector
$c = (x(0), \alpha) \in R^n \times R^p$	Control vector
$z(t) \in R^m$	Observation vector
$h(x(t)) : R^n \rightarrow R^m$	Forward operator relating model state and observations
$h(x(t)) \in R^m$	Model counterpart of the observation
$v(t) \in R^m$	Observation error vector
$\mathbf{R}(t) \in R^{m \times n}$	Covariance of observation error vector $v(t)$
e_F	Forecast error
$b(t)$	Systematic component of forecast error
$D_x(f) = [\partial f_i / \partial x_j] \in R^{n \times n}$	Model Jacobian w. R. T. x
$D_\alpha(f) = [\partial f_i / \partial \alpha_j] \in R^{n \times p}$	Model Jacobian w. R. T. A
$D_\alpha(x) = [\partial x_i / \partial \alpha_j] \in R^{n \times p}$	Sensitivity matrix w. R. T. Parameters
$D_{x(0)}(x) = [\partial x_i / \partial x_j(0)] \in R^{n \times n}$	Sensitivity matrix w. R. T. Initial condition
$D_x(h) \in R^{m \times n}$	Jacobian of the forward operator w. R. T. X
$H_1(t) = D_x(h)D_{x(0)}(x) \in R^{m \times n}$	Sensitivity matrix accounting for $h(x(t))$
$H_2(t) = D_x(h)D_\alpha(x) \in R^{m \times p}$	Sensitivity matrix accounting for $h(x(t))$
$\langle a, b \rangle$	Inner product
$J(c), J_1(c)$	Objective or cost functions
$\delta J, \delta J_1, \delta x(t), \delta \alpha$	First variations
$\nabla J, \nabla J_1$	Gradients of cost functions
$M(t, s)$	Model state transition matrix (Appendix A)
$L(t, s)$	Matrix that determines particular solution (Appendix A)

Consider a dynamical system described by a system of ordinary nonlinear differential equations of the form

$$\frac{dx}{dt} = f(x, \alpha, t), \quad (1a)$$

or in component form

$$\frac{dx_i}{dt} = f_i(x, \alpha, t), \quad (1b)$$

where dx/dt denotes the time derivative of the state $x(t)$, with $x(0) \in R^n$ the given initial condition. The control vector for the model is given by $c = (x(0), \alpha) \in R^n \times R^p$, the combination of initial condition and parameters referred to as the control space. It is tacitly assumed that the map of f in (1a) and (1b) is such that the solution $x(t) = x(t, x(0), \alpha) = x(t, c)$ exists and is unique. It is further assumed that $x(t)$ has a smooth dependence on the control vector c such that the first k (≥ 1) partial derivatives of $x(t)$ with respect to the components of c also exist. The solution $x(t)$ of (1a) and (1b) is known as the deterministic forecast of the state of the system at time $t > 0$. If the map $f(\cdot)$ in (1a) and (1b) depends explicitly on t , then this system is called a time varying or nonautonomous system; if $f(\cdot)$ does not depend on t , then the system is known as a time invariant or autonomous system.

Let $z(t) \in R^m$ be the observation vector obtained from the field measurements at time $t \geq 0$. Let $h : R^n \rightarrow R^m$ be the mapping from the model space R^n to the observation space R^m .

If $\hat{x}(t)$ denotes the (unknown) true state, then we assume that $z(t)$ is given by

$$z(t) = h(\hat{x}(t)) + v(t), \quad (2)$$

where $v(t) \in R^m$ is the additive (unobservable and unavoidable) noise. The mapping $h(\cdot)$ is known as the forward operator or the observation operator. It is further assumed that $v(t)$ is a white Gaussian noise with mean zero possessing a known covariance matrix $\mathbf{R}(t) \in R^{m \times m}$. That is, $v(t) \sim N(0, \mathbf{R}(t))$.

2.2. A Classification of Forecast Errors. The forecast error $e_F(t) \in R^m$ is defined as follows:

$$e_F(t) \equiv z(t) - h(x(t)) = b(t) + v(t), \quad (3)$$

the sum of a deterministic part

$$b(t) = h(\hat{x}(t)) - h(x(t)), \quad (4)$$

and the random part $v(t)$ induced by the observation noise. Our immediate goal is to analyze and isolate sources and types of forecast errors.

First, if the model map $f(\cdot)$ and the forward operator $h(\cdot)$ are without error, that is, exact, and if the control vector c is also error free, then the deterministic forecast $x(t)$ must be correct in the sense that $x(t) = \hat{x}(t)$, the true state. Then from (3), the forecast error is purely random or white Gaussian noise. That is,

$$e_F(t) = v(t). \quad (5)$$

Second, if $f(\cdot)$ and $h(\cdot)$ are known perfectly but c has an error, then the forecast $x(t)$ will have a deterministic error

induced by the incorrect control vector. In such a case, we can, in principle, decompose the forecast error as a sum

$$e_F(t) = b(c, t) + v(t), \quad (6)$$

where the deterministic part, $b(c, t) = h(\hat{x}(t)) - h(x(t))$, is purely a function of the control vector error.

Third, if $f(\cdot)$, $h(\cdot)$, and c are in error, then the forecast error can be represented as

$$e_F(t) = b(c, f, h, t) + v(t), \quad (7)$$

where the deterministic part $b(c, f, h, t)$ may have a complex (additive and/or multiplicative) dependence on errors in c , $f(\cdot)$, and $h(\cdot)$.

The following assumption is key to our analysis that follows. The model of choice is faithful to the phenomenon under study. The system is predicted with fidelity—the forecasted state is creditable and useful in understanding the dynamical processes that underpin the phenomenon. Certainly, the forecast will generally exhibit some error, but the primary physical processes are included; that is, the vector field f includes the pertinent physical processes. In this situation the forecast error stems from erroneously specified elements of control. Thus, in our study the forecast error assumes the form shown in (6). Dee's work [5] contains a very good discussion of the estimation of the bias b in (7) arising from errors in the model and/or observations.

2.3. Dynamics of First-Order Sensitivity Function Evolution. Since our approach is centered on sensitivity functions, we develop the dynamics of evolution of the forward sensitivities in this section.

Differentiating both sides of (1b) with respect to α_j , interchanging the order of differentiation on the left-hand side, we obtain

$$\frac{\partial(dx_i/dt)}{\partial\alpha_j} = \frac{d}{dt} \left(\frac{\partial x_i}{\partial \alpha_j} \right) = \sum_{k=1}^n \left[\left(\frac{\partial f_i}{\partial x_k} \right) \left(\frac{\partial x_k}{\partial \alpha_j} \right) + \frac{\partial f_i}{\partial \alpha_j} \right] \quad (8)$$

for $1 \leq i \leq n$ and $1 \leq j \leq p$ with $\partial x_i(0)/\partial \alpha_j = 0$ as the initial condition.

These np equations can be succinctly written in matrix form as

$$\frac{d}{dt} [D_\alpha(x)] = D_x(f) D_\alpha(x) + D_\alpha(f) \quad (9)$$

with $D_\alpha(x(0)) = 0$ as initial condition. This system of linear time-varying ordinary nonhomogeneous differential equations describe the evolution of the elements of $D_\alpha(x) = [\partial x_i / \partial \alpha_j] \in R^{n \times p}$, where the Jacobian matrices $D_x(f) \in R^{n \times n}$ and $D_\alpha(f) \in R^{n \times p}$ are given by

$$D_x(f) = \left[\frac{\partial f_i}{\partial x_j} \right], \quad (10)$$

$$D_\alpha(f) = \left[\frac{\partial f_i}{\partial \alpha_j} \right]. \quad (11)$$

Similarly, by differentiating both sides of (1b) with respect to $x_j(0)$, we obtain

$$\frac{\partial(dx_i/dt)}{\partial x_j(0)} = \frac{d}{dt} \left(\frac{\partial x_i}{\partial x_j(0)} \right) = \sum_{k=1}^n \left(\frac{\partial f_i}{\partial x_k} \right) \left(\frac{\partial x_k}{\partial x_j(0)} \right) \quad (12)$$

for $1 \leq i, j \leq n$, with $\partial x_i(0)/\partial x_j(0) = \delta_{ij}$, where δ_{ij} is the standard Kronecker delta. These n^2 equations can be succinctly represented as

$$\frac{d}{dt} [D_{x(0)}(x)] = D_x(f) D_{x(0)}(x) \quad (13)$$

with $D_{x(0)}(x(0)) = I$, the identity matrix. This system of linear, time-varying homogeneous equations governs the evolution of the elements of the matrix $D_{x(0)}(x) = [\partial x_i / \partial x_j(0)] \in R^{n \times p}$. Notice that (9) and (13) are independent of the observations and have the same system matrix $D_x(f)$ on the right-hand sides; thus, the homogeneous solutions to (9) and (13) have the same structure.

The evolution of the sensitivities (solution to (9) and (13)) is dependent on the solution to the governing dynamical equations ((1a) and (1b)). Generally, these equations are solved numerically using the standard fourth-order Runge-Kutta method. Rabitz et al. work [1] contains more details relating to solutions of (9) and (13). In special cases such as in air quality modeling, the sensitivity equations (9) and (13) exhibit extreme stiffness. Special methods are needed to handle the inherent stiffness of these equations. Seefeld and Stockwell work [2] includes a discussion of these issues. Gear's work [6] is a good reference for a general discussion of stiff equations.

3. Duality between the FSM and 4D-Var Based on Adjoint Method

Let $z(t_1), z(t_2), \dots, z(t_N)$ be a sequence of N observation vectors at times $t_0 = 0 < t_1 < t_2 \dots < t_N$. The goal is to use these observations to improve the estimate of the control vector c . This estimation problem is recast as a constrained minimization of an objective function $J : R^n \times R^p \rightarrow R$ given by

$$J(c) = \frac{1}{2} \sum_{i=1}^N [z(t_i) - h(x(t_i))]^T \mathbf{R}^{-1}(t_i) [z(t_i) - h(x(t_i))], \quad (14)$$

where the model state $x(t)$ evolves according to (1a), (1b), and $\mathbf{R}(t_i) \in R^{m \times m}$ is the known covariance of the observational errors $v(t_i)$ at time t_i , $1 \leq i \leq N$.

Fundamental to minimizing (14) is the computation of the gradient of $J(c)$, denoted by $\nabla_c J(c)$. In the following we describe two ways of characterizing

$$\nabla_c J(c) = \begin{bmatrix} \nabla_{x(0)} J \\ - - - \\ \nabla_\alpha J \end{bmatrix}, \quad (15)$$

where $\nabla_{x(0)} J \in R^n$ and $\nabla_\alpha J \in R^p$.

3.1. The Adjoint Method. This method is based on the basic principle that if δJ is the first variation of $J(c)$ induced by the perturbation δc in c , then

$$\delta J = \langle \nabla_c J, \delta c \rangle, \quad (16)$$

where $\langle a, b \rangle$ denotes the standard inner product of two vectors a and b of the same dimension. Once the first variation δJ is determined, the gradient can be found. In the following we exploit two basic properties of inner product:

(i) linearity

$$\langle a + b, x \rangle = \langle a, x \rangle + \langle b, x \rangle, \quad (17)$$

(ii) adjoint property

$$\langle a, Gb \rangle = \langle G^T a, b \rangle, \quad (18)$$

where G^T is the transpose or the adjoint of the matrix G .

From first principles (Chapter 24 in Lewis et al. [7]), it can be verified that the first variation δJ of J in (14) is given by

$$\delta J = - \sum_{k=1}^N \langle \mathbf{R}^{-1}(t_k) e_F(t_k), D_x(h) \delta x(t_k) \rangle, \quad (19)$$

where the forecast error given by

$$\begin{aligned} e_F(t_k) &= z(t_k) - h(x(t_k)), \\ D_x(h) &= \left[\frac{\partial h_i}{\partial x_j} \right] \in R^{m \times n} \end{aligned} \quad (20)$$

is the Jacobian of forward operator $h(x)$ with respect to x .

By invoking the adjoint property in (17), (19) becomes

$$\delta J = - \sum_{k=1}^N \langle \eta(t_k), \delta x(t_k) \rangle, \quad (21)$$

where

$$\eta(t_k) = D_x^T(h) \mathbf{R}^{-1}(t_k) e_F(t_k). \quad (22)$$

The first variation $\delta x(t_k)$ in $x(t)$ at $t = t_k$ resulting from the perturbation δc in c is given by (A.4) in Appendix A. Using (A.4) and the linearity of the inner product it follows that

$$\begin{aligned} \delta J &= - \sum_{k=1}^N \langle \eta(t_k), M(t_k, 0) \delta x(0) \rangle \\ &\quad - \sum_{k=1}^N \langle \eta(t_k), M(t_k, 0) L(t_k, 0) \delta \alpha \rangle, \end{aligned} \quad (23)$$

where we will refer to the first term on the right-hand side of (23) as ‘‘Term I’’ and the second as ‘‘Term II.’’ Using the adjoint and linearity property in (17), we get

$$\text{Term I} = - \left\langle \sum_{k=1}^N M^T(t_k, 0) \eta(t_k), \delta x(0) \right\rangle \quad (24)$$

from which we obtain

$$\nabla_{x(0)} J = - \sum_{k=1}^N M^T(t_k, 0) \eta(t_k). \quad (25)$$

Similarly, rewriting Term II as

$$\text{Term II} = - \left\langle \sum_{k=1}^N L^T(t_k, 0) M^T(t_k, 0) \eta(t_k), \delta \alpha \right\rangle, \quad (26)$$

we get

$$\nabla_{\alpha} J = - \sum_{k=1}^N L^T(t_k, 0) M^T(t_k, 0) \eta(t_k). \quad (27)$$

Hence, we obtain the components of $\nabla_c J$ which are used in conjunction with the minimization algorithm to find the optimal c^* that minimizes $J(c)$ in (14).

We conclude this discussion with an efficient recursive method for evaluating the expressions in (25) and (27). Define

$$\lambda_N = \eta(t_N), \quad (28)$$

and for $k = N - 1, N - 2, \dots, 2, 1, 0$,

$$\lambda_k = M^T(t_{k+1}, t_k) \lambda_{k+1} + \eta(t_k). \quad (29)$$

It can be verified that $\nabla_{x(0)} J = M^T(t_1, 0) \lambda_1$ in (25).

Details on the recursive computation of (27) are given in Appendix C.

3.2. Sensitivity-Based Approach. Let us first consider the special case when $N = 1$. Then (14) becomes

$$J_1(c^-) = \frac{1}{2} [z - h(x)]^T \mathbf{R}^{-1}(t) [z - h(x)], \quad (30)$$

where we recall that $c \in R^n \times R^p$ and $x = x(t, c)$ is the solution of the model equations (1a) and (1b) at time t .

Setting $q = (n + p)$ and $A = \mathbf{R}^{-1}(t)$, the expression $J_1(c)$ in (30) becomes identical to $Q(c)$ in (B.16) (Appendix B). Hence, by using (B.20) it follows that

$$\nabla_c J_1(c) = -D_c^T(x) \eta(t), \quad (31)$$

where $\eta(t)$ is given by (22) and

$$D_c^T(x) = \begin{bmatrix} D_{x(0)}^T(x) \\ \text{-----} \\ D_{\alpha}^T(x) \end{bmatrix} \in R^{(n+p) \times n}. \quad (32)$$

Now comparing (32) with (25)–(27), we obtain the duality relation:

$$\begin{aligned} D_{x(0)}(x) &\longleftrightarrow M(t, 0), \\ D_{\alpha}(x) &\longleftrightarrow M(t, 0) L(t, 0). \end{aligned} \quad (33)$$

The sensitivity matrices on the left-hand side of (33) are obtained by solving the forward sensitivity equations, whereas the matrix products $M(t, 0)$ and $L(t, 0)$ are obtained by integrating along the path as given by (A.3) and (A.5) (Appendix A).

The primary advantage of the sensitivity-based approach is that it provides a natural interpretation of the expression for the gradient in (31). Recall from (22) that $\eta(t) \in R^n$ is the weighted version of the forecast error $\mathbf{R}^{-1}(t)e_F(t) \in R^m$ mapped onto the model space by the linear operator $D_x^T(h)$. Also the i th column of $D_c^T(x)$ is the sensitivity of the i th component $x_i(t)$ of the state vector $x(t) \in R^n$ with respect to the control vector $c \in R^q$ given by $(\partial x_i/\partial c_1, \partial x_i/\partial c_2, \dots, \partial x_i/\partial c_q)$. Thus, it follows from (31) that $\nabla_c J_1$ is a linear combination of the columns of $D_c^T(x)$ which are the sensitivity vectors, where the coefficients of the linear terms are the components of $\eta(t)$. Thus, columns of $D_c^T(x)$ with large norms that are associated with large forecast errors will be dominant in determination of the components of $\nabla J_1(c)$. In other words, we gain some insight into the interplay between corrections to control and forecast errors—something that can be seen through a careful examination of the sensitivity vector at various times from initial state to forecast horizon. (The illustrative example in Section 5 further explores this diagnostic function.) Expression (31) also enables us to isolate the effect of different components x_i of $x(t)$ on the performance index $J_1(c)$.

For the general case of observations at multiple times, (31) assumes the following form:

$$\nabla J = - \sum_{i=1}^N D_c^T(x(t_i))\eta(t_i). \quad (34)$$

This gradient is the sum of linear combinations of the columns of $D_c^T(x(t_i))$ at various time instances. With so many directions (the directions associated with the columns of $D_c^T(x(t_i))$) contributing to the components of $\nabla_c J$, the connection between sensitivity, and forecast error is obscured.

4. Data Assimilation Using Sensitivity

We seek to find the solution to the following problem using the FSM. Given $f(\cdot)$ and $h(\cdot)$, the control vector c , the observation $z(t)$, and the error covariance of observations $\mathbf{R}(t)$, find a correction δc to the control vector such that the new model forecast starting from $(c + \delta c)$ will render the forecast error $e_F(t)$ purely random; that is, the systematic forecast error is removed and accordingly $E(e_F(t)) = 0$.

We start by quantifying the change Δx in the solution $x(t) = x(t, c)$ resulting from a change δc in c . Invoking the standard Taylor series expansion, we obtain

$$\Delta x = x(t, c + \delta c) - x(t, c) = \sum_j \delta^j x, \quad (35)$$

where $\delta^k x$ is the k th variation of $x(t)$, the fraction of the total change that can be accounted by the k th partial derivatives of $x(t)$ with respect to c and the perturbation δc . Since practical considerations dictate that the total number of correction

terms on the right-hand side of (35) be finite, we often settle for an approximation of only k terms (k generally ≤ 2). This is a tradeoff between improved accuracy resulting from a large value of k and the complexity of the resulting inverse problem. Although we have developed the methodology for second-order analysis ($k = 2$, where Δx is approximated by the sum $\delta x + \delta^2 x$) (Lakshmivarahan and Lewis [8]), our development will follow the first-order analysis ($k = 1$, where Δx is given by the first variation δx). Second-order analysis is justified when $\delta^2 x$ is a significant fraction of δx —this occurs when $f(x)$ and/or $h(x)$ exhibit strong nonlinearity. It is further shown in Section 5 that iterative application of the first-order method often leads to improved results.

4.1. First-Order Analysis. From first principles and (35) we obtain

$$\Delta x \approx \delta x = D_{x(0)}(x)\delta x(0) + D_\alpha(x)\delta\alpha, \quad (36)$$

where $D_{x(0)}(x) \in R^{n \times n}$ is the Jacobian of $x(t)$ with respect to $x(0)$, and $D_\alpha(x) \in R^{n \times p}$ is the Jacobian of $x(t)$ with respect to α . The matrices $D_{x(0)}(x)$ and $D_\alpha(x)$, found as solutions of (13) and (9), respectively, are known as the first-order sensitivity of the solution $x(t)$ with respect to $x(0)$ and α , respectively, and the elements of these matrices are called sensitivity functions.

4.1.1. Observations at One Time Only. We first consider the case where observation $z(t) \in R^m$ is available at one time t . The first variation δx in $x(t)$ induces a variation Δh in the forward operator $h(x(t))$. Again, by approximating Δh by the first variation, we get

$$\Delta h \approx \delta h = D_x(h)\delta x, \quad (37)$$

where $D_x(h) \in R^{m \times n}$ is the Jacobian of $h(\cdot)$ with respect to x and is given by

$$D_x(h) = \left[\frac{\partial h_i}{\partial x_j} \right], \quad 1 \leq i \leq m, \quad 1 \leq j \leq n. \quad (38)$$

substituting (36) into (37), we obtain

$$dh = H_1 dx(0) + H_2 d\alpha, \quad (39)$$

where $H_1(t) = D_x(h)D_{x(0)}(x) \in R^{m \times n}$ and $H_2(t) = D_x(h)D_\alpha(x) \in R^{m \times p}$. Setting $H(t) = [H_1(t), H_2(t)] \in R^{m \times (n+p)}$ and $\zeta = (\zeta_1, \zeta_2) \in R^{n+p}$, where $\zeta_1 = \delta x(0)$ and $\zeta_2 = \delta\alpha$, (39) becomes

$$dh = H(t)\zeta. \quad (40)$$

Given the operating point c , our goal is to find the perturbation δc such that the observation is equal to the model counterpart, that is,

$$z(t) = h(x(t, c + \delta c)) \approx h(x(t)) + \delta h, \quad (41)$$

or

$$e_F(t) \equiv z(t) - h(x(t)) = \delta h. \quad (42)$$

From (43), it follows that the required perturbation $\zeta \in R^{n+p}$ is obtained by solving the linear inverse problem

$$H(t)\zeta = e_F(t), \quad (43)$$

where $H(t) \in R^{m \times (n+p)}$ and $e_F(t) \in R^m$.

4.1.2. Observations at Multiple Times. The above analysis can be readily extended to the case where observations are available at N times. We denote these sets of observation vectors by $z(t_1), z(t_2), \dots, z(t_N)$, where $0 < t_1 < t_2 < \dots < t_N$. The forecast error $e_F(t_i)$ is given by

$$e_F(t_i) = z(t_i) - h(x(t_i)), \quad 1 \leq i \leq N. \quad (44)$$

Define

$$\begin{aligned} H_1(t_i) &= D_{x(t_i)}(h)D_{x(0)}(x(t_i)) \in R^{m \times n}, \\ H_2(t_i) &= D_{x(t_i)}(h)D_\alpha(x(t_i)) \in R^{m \times p}. \end{aligned} \quad (45)$$

Then at time t_i we have

$$H(t_i)\zeta = e_F(t_i), \quad 1 \leq i \leq N, \quad (46)$$

where

$$H(t_i) = [H_1(t_i), H_2(t_i)] \in R^{m \times (n+p)}. \quad (47)$$

Now define a matrix $H \in R^{Nm \times (n+p)}$ and a vector $e_F \in R^{Nm}$ as

$$H = \begin{bmatrix} H(t_1) \\ H(t_2) \\ \vdots \\ H(t_N) \end{bmatrix}, \quad e_F = \begin{bmatrix} e_F(t_1) \\ e_F(t_2) \\ \vdots \\ e_F(t_N) \end{bmatrix}. \quad (48)$$

Then, the N relations in (46) can be succinctly denoted by

$$H\zeta = e_F. \quad (49)$$

A number of special cases arise depending on (a) the value of Nm relative to $(n+p)$, namely, over (under) determined cases when $Nm > (<)(n+p)$ and (b) the rank of the matrix $H(t)$, namely, $H(t)$ is of full rank or rank deficient. In all these cases, the linear inverse problem (43) is recast as a minimization problem using the standard least squares framework (Lawson and Hanson [9]). The resulting minimization problem can then be solved using one of many standard methods, for example, the conjugate gradient method (Lewis et al. [7]; Nash and Sofer [10]).

As an illustration, consider the case when $Nm > (n+p)$ and that the rank of H is $(n+p)$, that is, full rank. The solution ζ is then obtained by minimizing the weighted least squares criterion

$$\begin{aligned} J_N(\zeta) &= \frac{1}{2} \sum_{i=1}^N (e_F(t_i) - H(t_i)\zeta)^T \mathbf{R}^{-1}(t_i) (e_F(t_i) - H(t_i)\zeta) \\ &= \frac{1}{2} (e_F - H\zeta)^T \mathbf{R}^{-1} (e_F - H\zeta), \end{aligned} \quad (50)$$

where

$$\mathbf{R} = \text{Diag}[\mathbf{R}(t_1), \mathbf{R}(t_2), \dots, \mathbf{R}(t_N)] \quad (51)$$

is an $Nm \times Nm$ diagonal matrix with $\mathbf{R}(t_i)$ as its i th diagonal block.

Although it is computationally efficient to minimize (50) by using a method like conjugate gradient, there is an advantage to analyze the properties of the optimal solution via the classical approach, that is, by setting the gradient of $J_N(\zeta)$ to zero. It can be verified that the minimizing J_N is found by solving the symmetric linear system

$$\left[\sum_{i=1}^N H^T(t_i) \mathbf{R}^{-1}(t_i) H(t_i) \right] \zeta = \sum_{i=1}^N \left(H^T(t_i) \mathbf{R}^{-1}(t_i) e_F(t_i) \right) \quad (52)$$

or succinctly as

$$\zeta_{\text{LS}} = \left(H^T \mathbf{R}^{-1} H \right)^{-1} H^T \mathbf{R}^{-1} e_F, \quad (53)$$

where H, e_F , and \mathbf{R} are defined in (48) and (51), and subscript ‘‘LS’’ refers to the least squares solution.

From the discussion relating to the classification of forecast errors, recall that the forecast error inherits its randomness from the (unobservable) observation noise. The vector e_F on the right hand side of (53) is random and hence the solution ζ of (53) is also random.

Since we are interested in forecast errors in response to incorrect control, we have

$$e_F(t_i) = b(c, t_i) + v(t_i) \quad \text{for } 1 \leq i \leq N. \quad (54)$$

Now define

$$b = \begin{bmatrix} b(c, t_1) \\ b(c, t_2) \\ \vdots \\ b(c, t_N) \end{bmatrix}, \quad v = \begin{bmatrix} v(t_1) \\ v(t_2) \\ \vdots \\ v(t_N) \end{bmatrix}. \quad (55)$$

Hence, the vector e_F in (48) can be expressed as

$$e_F = b + v \quad (56)$$

with $E(e_F) = b$ since $E(v) = 0$. Substituting (56) into (53) and taking the expectation give

$$E(\zeta_{\text{LS}}) = \left[\left(H^T \mathbf{R}^{-1} H \right)^{-1} H^T \mathbf{R}^{-1} \right] b. \quad (57)$$

Thus, the expected value of the correction to control is indeed a linear function of the forecast error b itself. It can be verified (Lewis et al. [7]) that the covariance of the least squares estimate [(53)] is given by

$$\text{cov}(\zeta_{\text{LS}}) = \left(H^T \mathbf{R}^{-1} H \right)^{-1} = \left[\nabla^2 J_N(\zeta) \right]^{-1}, \quad (58)$$

where $\nabla^2 J_N(\zeta)$ is the Hessian of $J_N(\zeta)$ in (50).

5. A Practical Example: Air/Sea Interaction

We choose a simple but nontrivial differential equation to demonstrate the applicability of the forward-sensitivity method to identification of error in a dynamical model. We break this discussion into three parts as follows: (1) the model, (2) discussion of the diagnostic value of FSM, and (3) numerical experiments with data assimilation using FSM.

5.1. The Model. Consider the case where cold continental air moves out over an ocean of constant surface temperature. We follow a column of air in a Lagrangian frame; that is, the column of air moves with the prevailing low-level wind speed. Turbulent transfer of heat from the ocean to air warms the column. The governing equation is taken to be

$$\frac{d\theta}{d\tau} = \frac{C_T V}{H}(\theta_s - \theta), \quad (59)$$

where

- θ : temperature of the air column ($^{\circ}\text{C}$),
- θ_s : temperature of the sea surface ($^{\circ}\text{C}$),
- C_T : turbulent heat exchange coefficient (nondimensional),
- V : speed of air column (ms^{-1}),
- H : height of the column (mixed layer)(m),
- τ : time (h). Equation (59) is nondimensionalized by the following scaling:

$$\begin{aligned} \theta &= \pi \cdot x \quad (\pi = 1.0^{\circ}\text{C}), \\ \theta_s &= \pi \cdot x_s, \\ \tau &= T \cdot t \quad (T = 1.0 \text{ h}). \end{aligned} \quad (60)$$

The governing equation then takes the form

$$\frac{dx(t)}{dt} = \frac{C_T V T}{H}(x_s - x(t)). \quad (61)$$

Assuming $H \sim 150 \text{ m}$, $V \sim 10 \text{ ms}^{-1}$, $C_T \sim 10^{-3}$, then

$$k = \frac{C_T V T}{H} \approx \frac{1}{4}. \quad (62)$$

Thus, we take our governing equation to be

$$\frac{dx(t)}{dt} = f(x, \alpha) = k(x_s - x(t)), \quad (63)$$

where $k = 0.25$. The solution to (63) is

$$x(t, x(0), \alpha) = [x(0) - x_s]e^{-kt} + x_s \quad (64)$$

with $c = (x(0), \alpha) \in R \times R^2$, and $\alpha = (x_s, k)^T \in R^2$, that is, $n = 1$ and $p = 2$. The solution depends linearly on $x(0)$ and x_s but nonlinearly on k .

There are three elements of control: initial condition, $x(0)$, boundary condition, x_s , and parameter, k .

5.2. Diagnostic Aspects of FSM. The Jacobians of f with respect to x and α are given by

$$\begin{aligned} D_x(f) &= -k \in R, \\ D_\alpha(f) &= [k, x_s - x(t)] \in R^{1 \times 2}, \end{aligned} \quad (65)$$

and the Jacobians of the solution $x(t)$ with respect to α and $x(0)$ are given by

$$\begin{aligned} D_\alpha(x(t)) &= \left[\frac{\partial x(t)}{\partial x_s}, \frac{\partial x(t)}{\partial k} \right] \in R^{1 \times 2}, \\ D_{x(0)}(x(t)) &= \left[\frac{\partial x(t)}{\partial x(0)} \right] \in R. \end{aligned} \quad (66)$$

From (9) and (13) the evolution of the forward sensitivities is given by

$$\frac{d}{dt} \left(\frac{\partial x(t)}{\partial x_s} \right) = -k \left(\frac{\partial x(t)}{\partial x_s} \right) + k, \quad (67)$$

$$\frac{d}{dt} \left(\frac{\partial x(t)}{\partial k} \right) = -k \left(\frac{\partial x(t)}{\partial k} \right) + (x_s - x(t)), \quad (68)$$

$$\frac{d}{dt} \left(\frac{\partial x(t)}{\partial x(0)} \right) = -k \left(\frac{\partial x(t)}{\partial x(0)} \right), \quad (69)$$

where

$$\left[\frac{\partial x(t)}{\partial x_s} \right]_{t=0} = 0 = \left[\frac{\partial x(t)}{\partial k} \right]_{t=0}, \quad \left[\frac{\partial x(t)}{\partial x(0)} \right]_{t=0} = 1. \quad (70)$$

Either by solving (67)–(70) or by computing directly from (64), it can be verified that the required sensitivities evolve according to

$$\begin{aligned} \frac{\partial x(t)}{\partial x(0)} &= e^{-kt}, & \frac{\partial x(t)}{\partial x_s} &= 1 - e^{-kt}, \\ \frac{\partial x(t)}{\partial k} &= [x_s - x(0)]te^{-kt}. \end{aligned} \quad (71)$$

The plots of the solution and the three sensitivities are given in Figure 1.

Let $z(t)$ be the direct observation of the state $x(t)$, namely,

$$z(t) = x(t) + v(t). \quad (72)$$

In this case, $h(t)$ is the forecast variable and therefore $D_x(h) = 1$. Then

$$e_F(t) = z(t) - x(t) \quad (73)$$

is the forecast error.

Following the developments in Section 4 for the case of a single observation described by (39)–(46), we obtain the analog of (46) as

$$H\zeta = e_F(t), \quad (74)$$

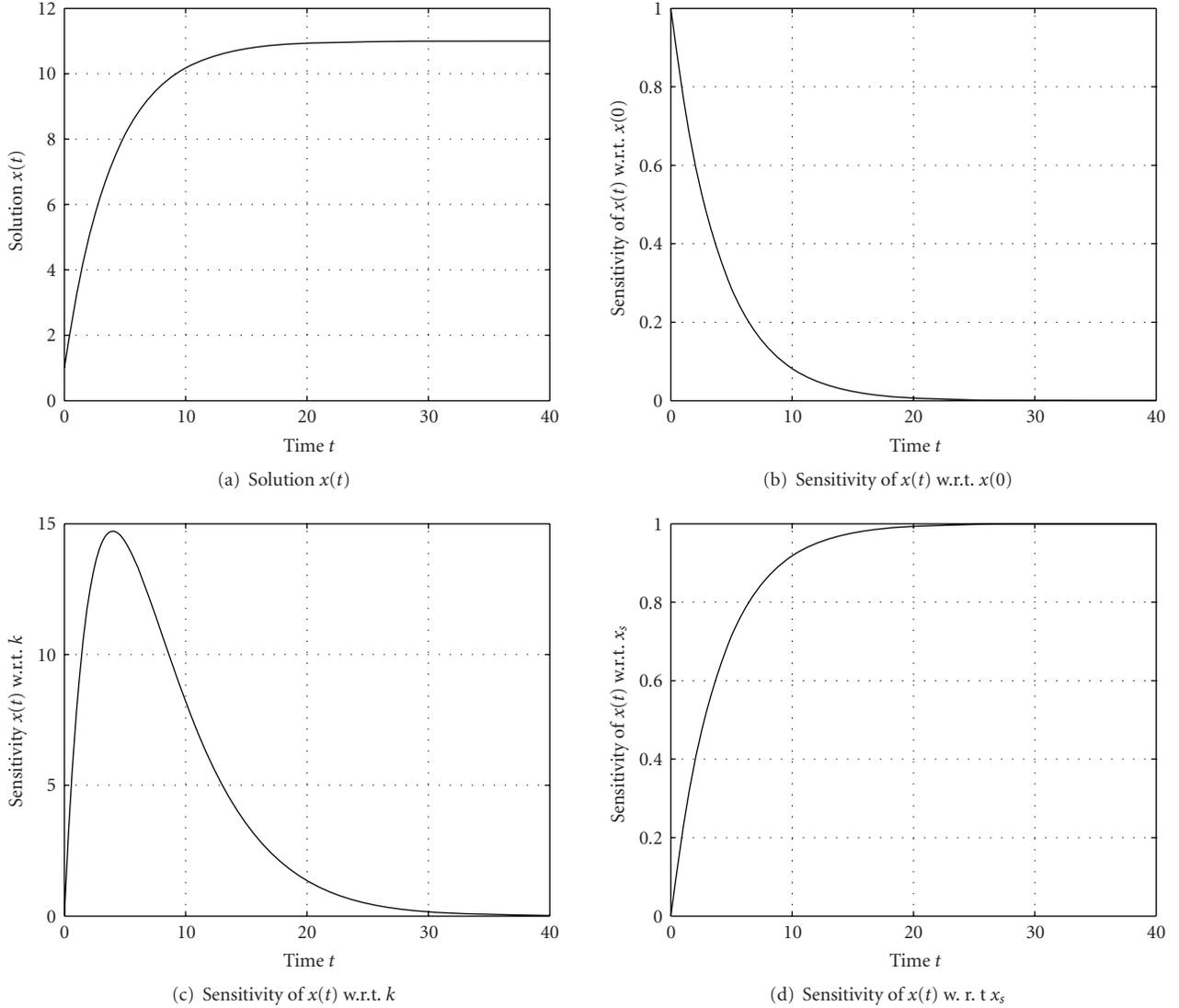


FIGURE 1: Evolution of the solution of $x(t)$ and its sensitivities to $c = (x(0), x_s, k) = (1.0, 11.0, 0.25)$.

where $H = H(t) = [\partial x(t)/\partial x(0), \partial x(t)/\partial x_s, \partial x(t)/\partial k]$ is the forward sensitivity vector and $\zeta = [\delta x(0), \delta x_s, \delta k]^T$. Clearly, (74) corresponds to an under-determined linear least squares problem, whose optimal solution is given by [7, chapter 5]

$$\zeta_{LS} = H^T (HH^T)^{-1} e_F(t) = \frac{H^T}{\|H\|} \frac{e_F(t)}{\|H\|}, \quad (75)$$

where $H^T/\|H\|$ is the unit forward sensitivity vector at time t and $e_F(t)/\|H\|$ is a scalar that is the forecast error normalized by the norm of the forward sensitivity vector H . In other words, the direction of the optimal corrections to the control that annihilate the forecast error is a constant multiple of the unit forward sensitivity vector.

If we assume the following control vector: $c = (1.0, 11.0, 0.25)$, that is, $[x(0) = 1^\circ\text{C}, x_s = 11^\circ\text{C}, \text{ and } k = 0.25 \text{ (nondimensional)}]$, we get

$$H^T = \begin{bmatrix} e^{-0.25t} \\ 1 - e^{-0.25t} \\ 10te^{-0.25t} \end{bmatrix} = \begin{bmatrix} \frac{\partial x(t)}{\partial x(0)} \\ \frac{\partial x(t)}{\partial x_s} \\ \frac{\partial x(t)}{\partial k} \end{bmatrix}. \quad (76)$$

The time variations of elements of H are given in Table 2 (also refer to Figure 1). From this table, it is clear that the direction of corrections to control varies as t increases. At $t = 0$, the corrections lie in the direction $(1, 0, 0)^T$, where $x(t)$ is only sensitive to the initial condition $x(0)$. For large

TABLE 2: Sensitivities of model state to initial conditions and parameters for the sea/air turbulent transfer model.

Time (hours)	$t = 0$	$t = 1$	$t = 5$	$t = 10$	$t = 15$	$t = 20$
$\partial x(t)/\partial x(0)$	1.0	0.7788	0.2865	0.0821	0.0235	0.0007
$\partial x(t)/\partial x_s$	0.0	0.2212	0.7135	0.9179	0.9765	0.9993
$\partial x(t)/\partial k$	0.0	7.788	14.325	8.2100	3.525	0.1400

t , the corrections lie in the direction $(0, 1, 0)^T$, where $x(t)$ is only sensitive to the boundary condition x_s (the sea-surface temperature). For intermediate times, all the components of control have nonzero sensitivities. $\partial x(t)/\partial k$ reaches its maximum at $t = 4.0$.

5.3. Numerical Experiments. We assume that the incorrect control vector is $c' = (x'(0), x'_s, k') = (2.0, 10.0, 0.30)$; in dimensional form, $x'(0) = 2^\circ\text{C}$, $x'_s = 10^\circ\text{C}$, and $k' = 0.30$ (non-dimensional). Thus, for an ideal correction to control,

$$\delta c = (-1.0, +1.0, -0.05). \quad (77)$$

To mimic reality, the correction process uses sensitivity functions that stem from the erroneous solution, that is, where the incorrect control is used.

We have explored both the goodness and failure of recovery of control under two different scenarios, where either 3 or 6 observations are used to recover the control vector. Since there are 3 unknowns, the case for 6 observations is an over-determined system.

We execute numerical experiments where the observations are spread over different segments of the temporal range—generally divided into an “early stage” and a “saturated stage.” By saturated stage we refer to that segment where the solution becomes close to the asymptotic state, that is, $x \rightarrow x_s$. The dividing time between these segments is arbitrary; but generally, based on experimental results to follow, we divide the segments at $t = 24$ where $\hat{x}(24) = 10.975$, 0.025 less than $\hat{x}_s = 11.0$ (see Figure 1).

The following general statement can be made. If more than one of the observations falls in the saturated zone, the matrix becomes illconditioned. As can be seen from (39) and the plots of sensitivity functions in Figure 1, $\partial x/\partial x_s \rightarrow 1$ and $\partial x/\partial x(0)$ and $\partial x/\partial k \rightarrow 0$ as $t \rightarrow \infty$. Accordingly, if two of the observations are made in the saturated zone, this induces linear dependency between the associated rows of the H matrix and in turn leads to ill-conditioning. This illconditioning is exhibited by a large value of the condition number, the ratio of the largest to the smallest eigenvalue of the matrix $H^T H$. The inversion of this matrix is central to the optimal adjustments of control (see (55)).

Illconditioning can also occur as a function of the observation spacing in certain temporal segments. This is linked to lack of variability or lack of change of sensitivity from one observation time to another. And, as can be seen in Figure 1, the absolute value of the slope in sensitivity function curves is generally large at the early stages of evolution and small at later stages. As an example, we find satisfactory recovery, $\delta c = (-0.882, -0.067, +0.922)$, when the observations are located at 5.0, 5.1, and 5.2 (a uniform

spacing of $\Delta t = 0.1$). Yet, near the saturated state, at $t = 20.0$, 20.1, and 20.2, again a spacing of 0.1, the recovery is poor with the result $\delta c = (+5.317, -0.142, +0.998)$. The associated condition numbers for these two experiments are 1.0×10^3 and 1.0×10^6 , respectively. Similar results follow from the case where 6 observations are taken. In all of these cases, the key factor is the condition number of $H^T H$. For our dynamical constraint, a condition number less than $\sim 10^4$ portends a good result.

For the case where we have 6 observations at $t = 2, 7, 12, 17, 22$, and 27, with a random error of 0.01 (standard deviation), we have executed an ensemble experiment with 100 members to recover control. In this case, the condition number is 2.4×10^3 . Results are plotted three dimensionally and in two-dimensional planes in the space of control, that is, plots of correction in the $x_s/x(0)$, x_s/k , and $x(0)/k$ planes. Results are shown in Figure 2.

Finally, we explore the iterative process of finding corrections to control. Here, the results from the 1st iteration are used to find the new control vector. This vector is then used to make another forecast and find a new set of sensitivity functions. The error of the forecast is obtained, and along with the new sensitivity functions, a second correction to control is found, and so forth. For the experiment with 6 observations that has been discussed in the previous paragraph, we apply this iterative methodology. As can be seen in Figure 3, the correct value of control is found in 3 iterations.

6. Concluding Remarks

The basic contributions of this paper may be stated as follows.

(1) While the 4D-Var has been the primary methodology for operational data assimilation in meteorology/oceanography (Lewis and Lakshmivaran [11]), and while forward sensitivity has been a primary tool for reaction kinetics and chemistry (Rabitz et al. [1]) and air quality modeling (Russell et al. [3]), to our knowledge these two methodologies have not been linked. We have now shown that the method of computing the gradient of $J(c)$ by these two approaches exhibits a duality hitherto unknown.

(2) By treating the forward sensitivity problem as an inverse problem in data assimilation, we are able to understand the fine structure of the forecast error. This is not possible with the standard 4D-Var formulation using adjoint equations.

(3) While it is true that computation of the evolution of the forward sensitivity involves computational demands beyond those required for solving the adjoint equations in

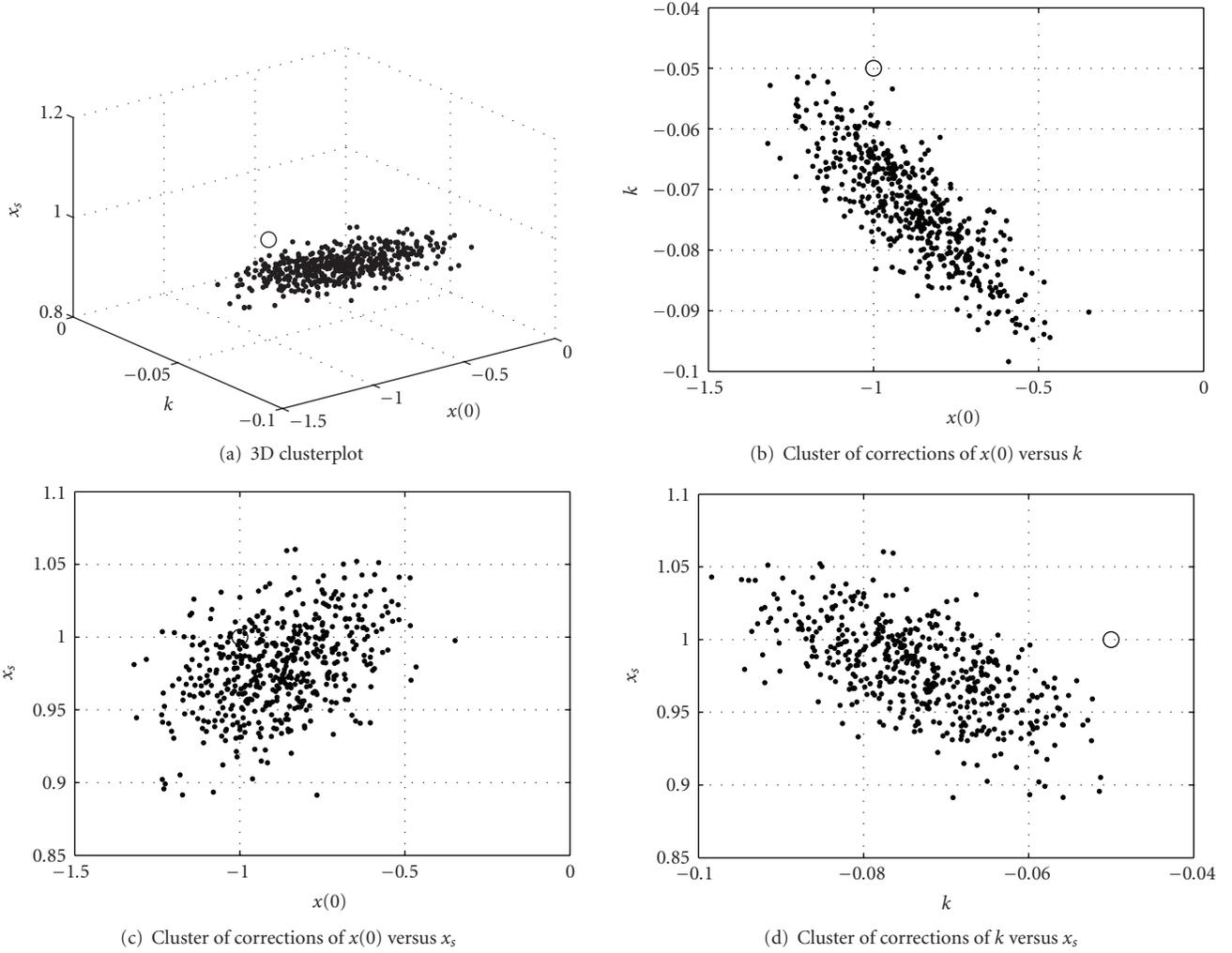


FIGURE 2: 3D cluster and its projections where corrections are computed using observations at the following times: $t = 2, 7, 12, 17, 22,$ and 27 . An ensemble of 100 members is used where the standard deviation of the observation noise is 0.01.

the standard 4D-Var methodology, there is a richness or augmentation to the information that comes with this added computational demand. In essence, it allows us to make judicious decisions on placement of observations through understanding of the time dependence of correction to control.

Appendices

A. Dynamics of Evolution of Perturbations

Let $\delta c = (\delta x^T(0), \delta \alpha^T)^T \in R^n \times R^p$ be the perturbation in the control vector c and $\delta x(t)$ the resulting perturbation in the state $x(t)$ induced by the dynamics (1a) and (1b). Our goal is to derive the dynamics of evolution of $\delta x(t)$.

From first principles, the evolutionary dynamics of $\delta x(t)$ are given by the variational equation (Hirsch et al. [12]) or the tangent linear model (Lewis et al. [7])

$$\frac{d}{dt} \delta x(t) = A(t) \delta x(t) + B(t) \delta \alpha, \quad (\text{A.1})$$

where the Jacobians $A(t)$ and $B(t)$ are given by

$$\begin{aligned} A(t) &= D_x(f) = \left[\frac{\partial f_i}{\partial x_j} \right] \in R^{n \times n}, \\ B(t) &= D_\alpha(f) = \left[\frac{\partial f_i}{\partial \alpha_j} \right] \in R^{n \times p}. \end{aligned} \quad (\text{A.2})$$

Define the integrating factor

$$M(t, s) = \exp \left[\int_s^t A(u) du \right] \in R^{n \times n}. \quad (\text{A.3})$$

premultiplying both sides of (A.1) by $M^{-1}(t, 0)$ and integrating, we get the solution of the linear nonhomogeneous and nonautonomous equation (A.1) as

$$\delta x(t) = M(t, 0) \delta x(0) + M(t, 0) L(t, 0) \delta \alpha, \quad (\text{A.4})$$

where

$$L(t, s) = \int_s^t M^{-1}(u, 0) B(u) du \in R^{n \times p}. \quad (\text{A.5})$$

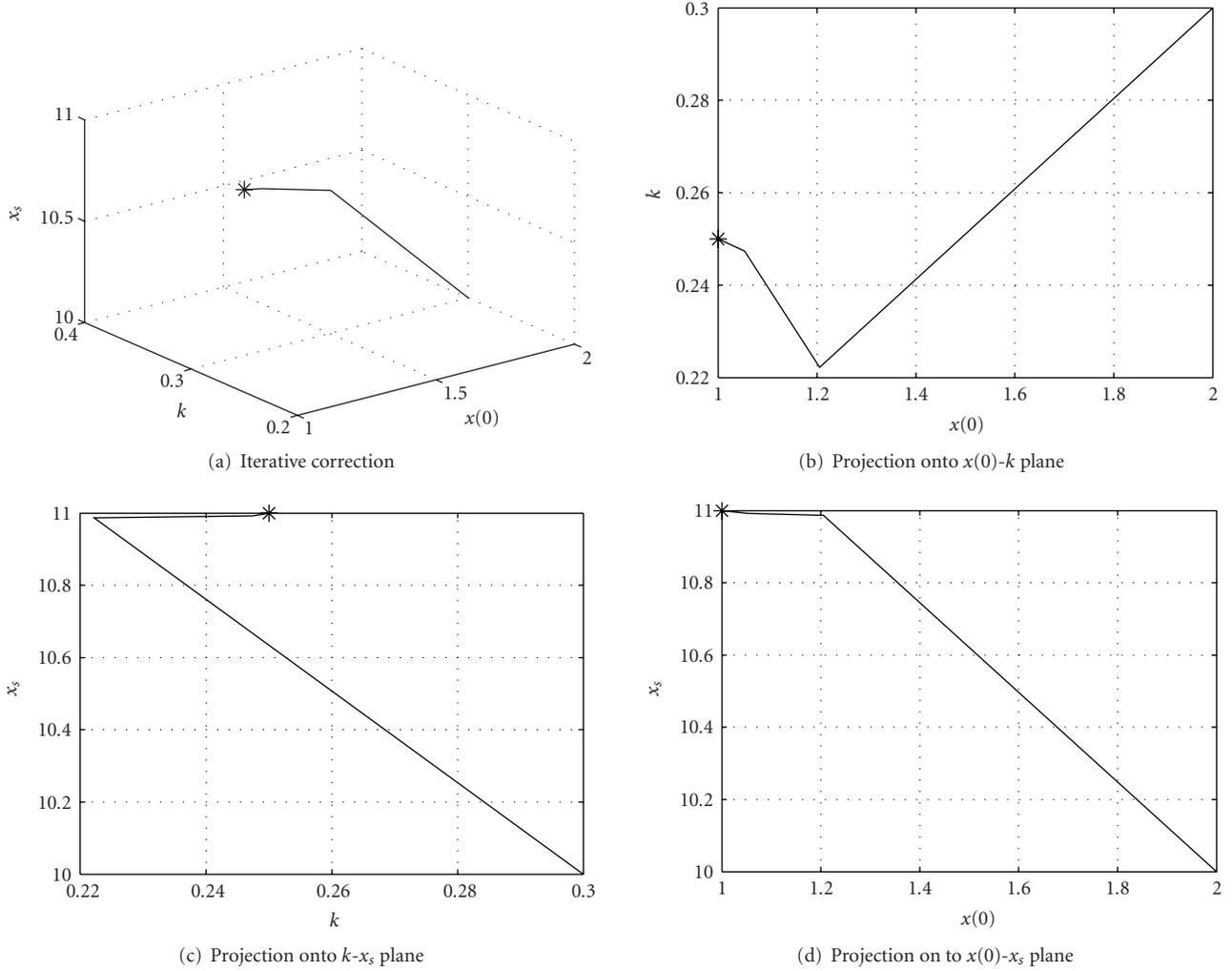


FIGURE 3: An illustration of the progression of first-order iterative corrections using six observations at $t = 2, 7, 12, 17, 22,$ and 27 .

From definitions (A.3)–(A.5), it can be verified that, for $u < s < t$,

$$\begin{aligned} M(t, u) &= M(t, s)M(s, u), \\ L(t, u) &= L(t, s) + L(s, u). \end{aligned} \quad (\text{A.6})$$

We now consider two special cases.

Case A. Let $\delta\alpha = 0$, that is, the initial perturbations are confined only to the initial condition, $x(0)$. Then setting $B(t) \equiv 0$, from (A.5) we see that $L(t, 0) \equiv 0$. From (A.4) we get

$$\delta x(t) = M(t, 0)\delta x(0). \quad (\text{A.7})$$

Case B. Let $\delta x(0) = 0$, that is, the initial perturbations are confined only to the parameter, α . Then setting $A(t) \equiv 0$, from (A.3), we see that $M(t, s) \equiv I$, the identity matrix. Now from (A.4) it follows that

$$\delta x(t) = L(t, 0)\delta\alpha. \quad (\text{A.8})$$

B. Computation of Sensitivity Functions

Let $c = (c_1, c_2, \dots, c_q)^T \in R^q$, where $x = x(t, c) = (x_1(t, c), x_2(t, c), \dots, x_n(t, c))^T \in R^n$ and $h(x) = (h_1(x), h_2(x), \dots, h_m(x))^T \in R^m$. Let $a = (a_1, a_2, \dots, a_m)^T \in R^m$ and consider

$$\phi_1(c) = a^T h(x) = \sum_{i=1}^m a_i h_i(x). \quad (\text{B.9})$$

By applying the standard chain rule it can be verified that the gradient $\nabla_c \phi_1(c)$ with respect to c is given by

$$\nabla_c \phi_1(c) = D_c^T(x) D_x^T(h) a, \quad (\text{B.10})$$

where

$$D_x(h) = \left[\frac{\partial h_i}{\partial x_j} \right] \in R^{m \times n} \quad (\text{B.11})$$

is the Jacobian of h and

$$D_c(x) = \left[\frac{\partial x_i}{\partial c_j} \right] \in R^{n \times q} \quad (\text{B.12})$$

```

DO j = N to 1
  DO i = N to j
     $\eta(t_i) = M^T(t_j, t_{j-1})\eta(t_i)$ 
  END
END
 $\mu(1) = L(t_1, 0)$ 
DO i = 2 to N
 $\mu(i) = \mu(i-1) + L(t_i, t_{i-1})$ 
END
Grad = 0
DO i = 1 to N
  Grad = Grad +  $\mu(i)\eta(t_i)$ 
END

```

ALGORITHM 1

is the (Jacobian) sensitivity of the vector x with respect to c at time t .

Now consider a quadratic form

$$\phi_2(c) = h^T(x)Ah(x), \quad (\text{B.13})$$

where $A \in R^{m \times m}$ is a symmetric and positive definite matrix.

Then by the product rule

$$\nabla_c \phi_2(c) = \nabla_c [h^T b] + \nabla_c [a^T h], \quad (\text{B.14})$$

where $b = Ah = a$. By applying (B.10) to each of the two terms on the right side of (B.14), it follows that

$$\nabla_c \phi_2(c) = 2D_c^T(x)D_x^T(h)Ah. \quad (\text{B.15})$$

Finally, if

$$Q(c) = [z - h(x)]^T A [z - h(x)], \quad (\text{B.16})$$

then expand

$$Q(c) = z^T A z - 2z^T A h(x) + h^T(x) A h(x) \quad (\text{B.17})$$

since z does not depend on c , by using (B.10) and (B.15), we get

$$Q(c) = -2D_c^T(x)D_x^T(h)A[z - h(x)]. \quad (\text{B.18})$$

Define

$$\xi = D_x^T(h)A[z - h(x)] \in R^n, \quad (\text{B.19})$$

then

$$\nabla_c Q(c) = -2D_c^T(x)\xi. \quad (\text{B.20})$$

That is, the gradient of Q with respect to c is the linear combination of the sensitivity vectors, that is, the columns of $D_c^T(x)$, where the coefficients in this linear combination are the elements of the vector ξ .

C. Computation of $\nabla_\alpha J$ in (27)

Given $\eta(t_i)$, $M^T(t_i, t_{i-1})$, $L^T(t_i, t_{i-1})$ for $1 \leq i \leq N$, the expression on the right-hand side of (27) can be efficiently computed as shown in Algorithm 1.

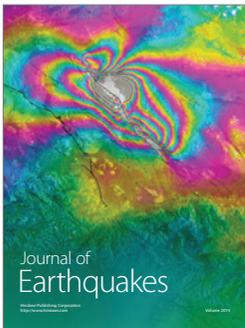
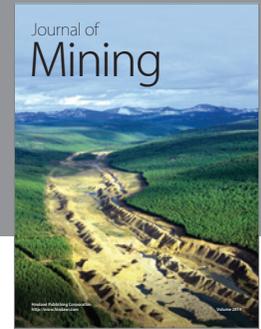
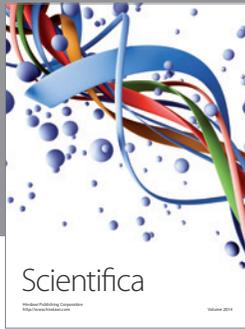
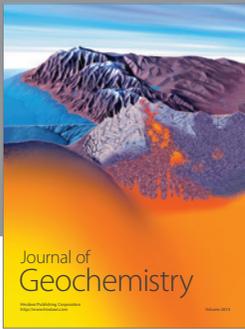
Then $\nabla_\alpha J = -\text{Grad}$. It is to be noticed that there is only matrix-vector multiplication involved in these operations and not matrix-matrix multiplication.

Acknowledgments

At an early stage of formulating our ideas on this method of data assimilation, Fedor Mesinger and Qin Xu offered advice and encouragement. Qin Xu and Jim Purser carefully checked the mathematical development, and suggestions from the following reviewers went far to improve the presentation: Yoshi Sasaki, Bill Stockwell, and anonymous formal reviewers of the manuscript. S. Lakshmivarahan's work is supported in part by NSF EPSCoR RII Track 2 Grant 105-155900 and by NSF Grant 105-15400, and J. Lewis acknowledges the Office of Naval Research (ONR), Grant No. N00014-08-1-0451, for research support on this project.

References

- [1] H. Rabitz, M. Kramer, and D. Dacol, "Sensitivity analysis in chemical kinetics," *Annual Review of Physical Chemistry*, vol. 334, pp. 419–461, 1983.
- [2] S. Seefeld and W. R. Stockwell, "First-order sensitivity analysis of models with time-dependent parameters: an application to PAN and ozone," *Atmospheric Environment*, vol. 33, no. 18, pp. 2941–2953, 1999.
- [3] A. Russell, J. Milford, M. S. Bergin, et al., "Urban ozone control and atmospheric reactivity of organic gases," *Science*, vol. 269, no. 5223, pp. 491–495, 1995.
- [4] F.-X. Le Dimet and O. Talagrand, "Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects," *Tellus A*, vol. 38, no. 2, pp. 97–110, 1986.
- [5] D. P. Dee, "Bias and data assimilation," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 613, pp. 3323–3343, 2006.
- [6] C. W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1971.
- [7] J. M. Lewis, S. Lakshmivarahan, and S. Dhall, *Dynamic Data Assimilation: A Least Squares Approach*, Cambridge University Press, Cambridge, UK, 2006.
- [8] S. Lakshmivarahan and J. M. Lewis, "Forecast error correction using forward sensitivity analysis: framework," Tech. Rep., School of Computer Sciences, University of Oklahoma, Norman, Okla, USA, 2009.
- [9] C. L. Lawson and R. J. Hanson, "Solving least squares problems," in *Classics in Applied Mathematics*, vol. 15, p. 337, SIAM, Philadelphia, Pa, USA, 1995.
- [10] S. G. Nash and A. Sofer, *Linear and Nonlinear Programming*, McGraw-Hill, New York, NY, USA, 1996.
- [11] J. M. Lewis and S. Lakshmivarahan, "Sasaki's pivotal contribution: calculus of variations applied to weather map analysis," *Monthly Weather Review*, vol. 136, no. 9, pp. 3553–3567, 2008.
- [12] M. W. Hirsch, S. Smale, and R. L. Devaney, *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press, New York, NY, USA, 2nd edition, 2004.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

