

## On the Sensitivity Equations of Four-Dimensional Variational (4D-Var) Data Assimilation

DACIAN N. DAESCU

*Portland State University, Portland, Oregon*

(Manuscript received 18 September 2007, in final form 16 December 2007)

### ABSTRACT

The equations of the forecast sensitivity to observations and to the background estimate in a four-dimensional variational data assimilation system (4D-Var DAS) are derived from the first-order optimality condition in unconstrained minimization. Estimation of the impact of uncertainties in the specification of the error statistics is considered by evaluating the sensitivity to the observation and background error covariance matrices. The information provided by the error covariance sensitivity analysis is used to identify the input components for which improved estimates of the statistical properties of the errors are of most benefit to the analysis and forecast. A close relationship is established between the sensitivities within each input pair data/error covariance such that once the observation and background sensitivities are available the evaluation of the sensitivity to the specification of the corresponding error statistics requires little additional computational effort. The relevance of the 4D-Var sensitivity equations to assess the data impact in practical applications is discussed. Computational issues are addressed and idealized 4D-Var experiments are set up with a finite-volume shallow-water model to illustrate the theoretical concepts. Time-dependent observation sensitivity and potential applications to improve the model forecast are presented. Guidance provided by the sensitivity fields is used to adjust a 4D-Var DAS to achieve forecast error reduction through assimilation of supplementary data and through an accurate specification of a few of the background error variances.

### 1. Introduction

The estimation of the observational data impact on analysis and subsequent model forecasts and the assessment of the contribution/benefit of each individual observation to the forecast error reduction are among the most important and most challenging problems in data assimilation and numerical weather prediction (NWP). Adjoint-based observation sensitivity techniques may be used to identify datasets that are most valuable to the assimilation procedure, to conduct optimal data thinning, and to design cost-effective field experiments for collecting targeted observations (Langland 2005a). A theoretical framework to sensitivity analysis in variational data assimilation is presented by Le Dimet et al. (1997) in the context of optimal control. Baker and Daley (2000) derived the equations of the forecast sensitivity to observations and to the background estimate

using the adjoint of a three-dimensional variational (3D-Var) analysis scheme and explained the need to account for the characteristics of the data assimilation system (DAS) in the design of adaptive observation strategies. Sensitivity-to-observations techniques have been used to assess the impact of satellite data (Fourrié et al. 2002) and of targeted observations and their interaction with the routine observational network (Doerenbecher and Bergot 2001; Bergot and Doerenbecher 2002). An efficient adjoint-based method for assessing the impact of observations on a scalar measure of the short-range forecast error using the adjoint of the Naval Research Laboratory 3D-Var DAS is presented in the work of Langland and Baker (2004) and Langland (2005b). Cardinali et al. (2004) used an influence-matrix diagnostic to assess the information content and data impact in a four-dimensional variational (4D-Var) DAS. Recently, Zhu and Gelaro (2008) developed the adjoint of the Gridpoint Statistical Interpolation analysis scheme (Wu et al. 2002) to obtain observation sensitivities that are fully consistent with the data assimilation procedure.

Theoretically, the adjoint model methodology may

---

*Corresponding author address:* Dr. Dacian N. Daescu, Department of Mathematics and Statistics, Portland State University, P.O. Box 751, Portland, OR 97207.  
E-mail: daescu@pdx.edu

be used to provide sensitivity information with respect to all parameters in the DAS. Previous studies have been mainly concerned with the forecast sensitivity to the observational data  $\mathbf{y}$  and to the background estimate  $\mathbf{x}_b$  to initial conditions and a large amount of additional parameters/input is involved in the data assimilation process. In an operational DAS several simplifying assumptions are made on the errors statistics, and only estimates  $\mathbf{R}$  and  $\mathbf{B}$  of the observation and background error covariances, respectively, may be provided because the true error covariance matrices  $\mathbf{R}_i$  and  $\mathbf{B}_i$  are unknown. A significant amount of research in NWP is focused on modeling the error covariance matrices (Janjić and Cohn 2006; Frehlich 2006; Kucukkaraca and Fisher 2006; Etherton and Bishop 2004; Lorenc 2003; Hamill and Snyder 2002). Diagnosis and tuning of error variances used in variational data assimilation are discussed in the work of Desroziers and Ivanov (2001) and Chapnik et al. (2006). A better understanding of how uncertainties in the specification of the error statistics will impact the analysis and forecast may be achieved by extending the sensitivity analysis to the DAS input  $(\mathbf{y}, \mathbf{R}, \mathbf{x}_b, \mathbf{B})$ .

The model fit to each observational dataset is weighted in the assimilation process by the inverse observation error covariance matrix  $\mathbf{R}^{-1}$  and valuable information may be lost if the observation errors are misrepresented. Given the increased amount of data available from the global observing systems, fine-tuning the observation error covariances by trial-and-error methods is not feasible (Rabier 2005) and guidance provided by the sensitivity analysis to the observation error covariance may be used to identify the data subsets where improvements in the observation error representation would be of most benefit.

Specification of the statistical properties of the background errors is a key ingredient in data assimilation and accurate error estimates are difficult to provide (Parrish and Derber 1992; Lorenc 2003). Sensitivity experiments of Baker and Daley (2000) indicate that proper input of the background error covariance in data assimilation may have a much greater impact on the forecast than the selection of the prior estimate to the initial conditions. In this context, information provided by the forecast sensitivity to the background error covariance matrix may be used to identify regions where improved background error estimates are of significant benefit to the forecast.

Observation sensitivity techniques have been mainly implemented in the 3D-Var context and an optimal use of the time-distributed observational data may be achieved with a 4D-Var DAS. In this work we provide an analytic derivation of the sensitivity equations of a

4D-Var DAS from the first-order necessary condition in unconstrained optimization. In section 2 the equations of the forecast sensitivity with respect to the time series of data/error covariance pairs  $(\mathbf{y}_i, \mathbf{R}_i)$ ,  $i = 0, 1, \dots, N$ , and to the background input pair  $(\mathbf{x}_b, \mathbf{B})$  are derived in a general 4D-Var context. A close relationship between the sensitivities within each pair is established. A table of sensitivity equations is provided and their relevance to assess the data impact in practical applications is discussed. Computational aspects are addressed in section 3 and idealized numerical experiments are used to illustrate the theoretical concepts. Potential applications of the sensitivity analysis and adjustment of a 4D-Var DAS to improve the model forecast are presented. A summary and concluding remarks are in section 4.

## 2. Sensitivity equations of 4D-Var

An in-depth presentation of the data assimilation techniques used in NWP may be found in the books of Jazwinski (1970), Daley (1991), Kalnay (2002), and Lewis et al. (2006). In 4D-Var data assimilation (Le Dimet and Talagrand 1986), optimal initial conditions (analysis)  $\mathbf{x}_0^a \in \mathbb{R}^n$  to a forecast model are obtained by minimizing a cost functional

$$\begin{aligned}
 \mathcal{J}(\mathbf{x}_0) &= \frac{1}{2} (\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_b) \\
 &+ \frac{1}{2} \sum_{i=0}^N [\mathbf{H}_i(\mathbf{x}_i) - \mathbf{y}_i]^T \mathbf{R}_i^{-1} [\mathbf{H}_i(\mathbf{x}_i) - \mathbf{y}_i], \\
 \mathbf{x}_0^a &= \arg \min \mathcal{J}, \tag{1}
 \end{aligned}$$

where  $\mathbf{x}_b$  is a prior (background) estimate to the initial state,  $\mathbf{y}_i \in \mathbb{R}^{k_i}$ ,  $i = 0, 1, \dots, N$ , is the vector of observational data at instant  $t_i$  of the analysis time interval  $[t_0, t_N]$ , and  $\mathbf{x}_i = \mathcal{M}_{t_0 \rightarrow t_i}(\mathbf{x}_0)$  is the model state at  $t_i$ . The model  $\mathcal{M}$  is nonlinear and  $H_i: \mathbb{R}^n \rightarrow \mathbb{R}^{k_i}$  is the observation operator that maps the state space into the observation space at time  $t_i$ . Statistical information on the errors in the background and data is used to define appropriate weights:  $\mathbf{B}$  is the covariance matrix of the background errors and  $\mathbf{R}_i$  is the covariance matrix of the observational errors. By imposing the model equations  $\mathbf{x}_{i+1} = \mathcal{M}_{t_i \rightarrow t_{i+1}}(\mathbf{x}_i)$  as a strong constraint, the 4D-Var data assimilation is formulated as a large-scale unconstrained optimization problem with the initial conditions as control parameters. A weak-constraint 4D-Var data assimilation may be formulated to account for the model error as part of the 4D-Var control variable (Trémolet 2007). The 3D-Var data assimilation may be viewed as the particular case of the 4D-Var data assimilation when only the data available at  $t_0$  are as-

simulated. The dimension of the discrete state vector in numerical weather prediction and climate models is of order  $n \sim 10^7$  and in practice several simplifications are required to reduce the computational burden associated with the minimization problem (1). The incremental 4D-Var scheme (Courtier et al. 1994) approximates the nonlinear problem (1) by a sequence (outer loop) of quadratic minimization problems (inner loop) to obtain an analysis increment  $\delta \mathbf{x}_0^a$  and it is currently the only approach proven to be feasible for operational implementation of the 4D-Var scheme (Rabier 2005). Implementation of the incremental 4D-Var algorithm and convergence issues are discussed by Trémolet (2004, 2005) and Lawless et al. (2005).

Mathematical foundations of the adjoint sensitivity for nonlinear dynamical systems are provided in the books of Cacuci (2003) and Marchuk et al. (1996). In this section the sensitivity equations of the 4D-Var problem (1) with respect to the DAS input  $(\mathbf{y}_i, \mathbf{R}_i, \mathbf{x}_b, \mathbf{B})$  are derived in the theoretical framework of sensitivity to parameters in unconstrained optimization.

#### a. Sensitivity to parameters in unconstrained optimization

Given a function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ , the gradient matrix  $\nabla f$  is the transpose of the Jacobian matrix:  $\nabla f(\mathbf{x}) = (\partial f_i / \partial x_j)^T \in \mathbb{R}^{n \times m}$ . If  $m = 1$ ,  $\nabla f$  denotes the vector of partial derivatives  $\partial f / \partial x_j$  in column format.

In a general framework of unconstrained optimization (Bertsekas 1995), consider the problem of finding the vector  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$  that minimizes the cost

$$\min_{\mathbf{x} \in \mathbb{R}^n} J(\mathbf{x}, \mathbf{u}), \quad (2)$$

where  $J: \mathbb{R}^{n+p} \rightarrow \mathbb{R}$  is a twice continuously differentiable function involving a parameter vector (input data)  $\mathbf{u} \in \mathbb{R}^p$ . Let  $\bar{\mathbf{x}}$  be a local minimum corresponding to the nominal parameter value  $\bar{\mathbf{u}}$  and assume that the Hessian matrix  $\nabla_{\mathbf{xx}}^2 J(\bar{\mathbf{x}}, \bar{\mathbf{u}}) \in \mathbb{R}^{n \times n}$  is positive definite. The implicit function theorem applied to the first-order optimality condition

$$\nabla_{\mathbf{x}} J(\bar{\mathbf{x}}, \bar{\mathbf{u}}) = \mathbf{0} \in \mathbb{R}^n \quad (3)$$

guarantees the existence of a vicinity of  $\bar{\mathbf{u}}$  where the optimal solution is a function of data  $\mathbf{x} = \mathbf{x}(\mathbf{u})$  and the gradient matrix

$$\nabla_{\mathbf{u}} \mathbf{x} = (\nabla_{\mathbf{u}} x_1, \nabla_{\mathbf{u}} x_2, \dots, \nabla_{\mathbf{u}} x_n) \in \mathbb{R}^{p \times n}$$

is expressed as

$$\nabla_{\mathbf{u}} \mathbf{x}(\mathbf{u}) = -\nabla_{\mathbf{ux}}^2 J[\mathbf{x}(\mathbf{u}), \mathbf{x}]\{\nabla_{\mathbf{xx}}^2 J[\mathbf{x}(\mathbf{u}), \mathbf{x}]\}^{-1}. \quad (4)$$

In (4),  $\nabla_{\mathbf{ux}}^2 J$  denotes the  $(p, n)$  dimensional matrix  $\nabla_{\mathbf{u}}(\nabla_{\mathbf{x}} J)$ .

The sensitivity (gradient) to parameters of a scalar-valued function  $J(\mathbf{x})$  of the optimal solution  $\mathbf{x} = \mathbf{x}(\mathbf{u})$  is obtained using chain rule differentiation

$$\nabla_{\mathbf{u}} J = \nabla_{\mathbf{u}} \mathbf{x} \nabla_{\mathbf{x}} J = -\nabla_{\mathbf{ux}}^2 J (\nabla_{\mathbf{xx}}^2 J)^{-1} \nabla_{\mathbf{x}} J \in \mathbb{R}^p. \quad (5)$$

#### b. The 4D-Var observation sensitivity equations

Corresponding to the 4D-Var cost (1), the first-order necessary condition (3) at the analysis state  $\mathbf{x}_0^a$  is

$$\nabla_{\mathbf{x}_0} J(\mathbf{x}_0^a) = \mathbf{B}^{-1}(\mathbf{x}_0^a - \mathbf{x}_b) + \sum_{i=0}^N \mathbf{M}_{0,i}^T \mathbf{H}_i^T \mathbf{R}_i^{-1} [H_i(\mathbf{x}_i) - \mathbf{y}_i] = 0, \quad (6)$$

where  $\mathbf{M}_{0,i}(\mathbf{x}_0^a)$  is the tangent linear model associated with the nonlinear model integration  $\mathbf{x}_i = \mathcal{M}_{t_0 \rightarrow t_i}(\mathbf{x}_0^a)$  and  $\mathbf{M}_{0,i}^T(\mathbf{x}_0^a)$  denotes its adjoint (transpose);  $\mathbf{H}_i(\mathbf{x}_i) \in \mathbb{R}^{k_i \times n}$  is the Jacobian matrix of the observation operator  $H_i$  evaluated at  $\mathbf{x}_i$  and  $\mathbf{H}_i^T(\mathbf{x}_i)$  denotes its transpose.

The significance of the inverse Hessian matrix  $[\nabla_{\mathbf{x}_0}^2 J(\mathbf{x}_0^a)]^{-1}$  is well established in the data assimilation literature since it provides an approximation to the covariance matrix of the errors in the analysis  $\mathbf{x}_0^a$ . The relationship is exact if the cost functional is quadratic (e.g., in the incremental 4D-Var framework) (Fisher and Courtier 1995) and the minimum reached  $\nabla_{\mathbf{x}_0} J(\mathbf{x}_0^a) = 0$ . Henceforth in this work, for notational convenience only, we denote

$$\mathbf{A} \stackrel{\text{def}}{=} [\nabla_{\mathbf{x}_0}^2 J(\mathbf{x}_0^a)]^{-1} \in \mathbb{R}^{n \times n}. \quad (7)$$

Differentiating (6) with respect to  $\mathbf{y}_i$

$$\nabla_{\mathbf{y}_i}^2 J(\mathbf{x}_0^a) = -\mathbf{R}_i^{-1} \mathbf{H}_i \mathbf{M}_{0,i} \in \mathbb{R}^{k_i \times n}, \quad (8)$$

and from (4), (7), and (8) the analysis sensitivity to observations  $\mathbf{y}_i$  is

$$\nabla_{\mathbf{y}_i} \mathbf{x}_0^a = \mathbf{R}_i^{-1} \mathbf{H}_i \mathbf{M}_{0,i} \mathbf{A} \in \mathbb{R}^{k_i \times n}. \quad (9)$$

In practice it is of interest to assess the observation impact on a scalar aspect  $J^v(\mathbf{x}_v)$  of the forecast  $\mathbf{x}_v = \mathcal{M}_{t_0 \rightarrow t_v}(\mathbf{x}_0^a)$  at a future time  $t_v > t_N$ , such as an appropriate energy norm of the forecast error over a certain geographical region (verification domain). Using chain rule differentiation

$$\nabla_{\mathbf{y}_i} J^v = \nabla_{\mathbf{y}_i} \mathbf{x}_0^a \nabla_{\mathbf{x}_0} J^v \in \mathbb{R}^{k_i}, \quad (10)$$

where

$$\nabla_{\mathbf{x}_0} J^v = \mathbf{M}_{0,v}^T \nabla_{\mathbf{x}_v} J^v \in \mathbb{R}^n \quad (11)$$

is the forecast sensitivity to analysis. Equations (9) and (10) provide the expression of the forecast sensitivity to observations

$$\nabla_{\mathbf{y}_i} J^v = \mathbf{R}_i^{-1} \mathbf{H}_i \mathbf{M}_{0,i} \mathbf{A} \nabla_{\mathbf{x}_0^a} J^v \in \mathbb{R}^{k_i}. \quad (12)$$

An equivalent expression may be obtained by denoting

$$\mathbf{A}_i = \mathbf{M}_{0,i} \mathbf{A} \mathbf{M}_{0,i}^T \in \mathbb{R}^{n \times n} \quad (13)$$

and noticing that the forecast sensitivity to analysis (11) may be written

$$\nabla_{\mathbf{x}_0^a} J^v = \mathbf{M}_{0,i}^T \mathbf{M}_{i,v}^T \nabla_{\mathbf{x}_i} J^v = \mathbf{M}_{0,i}^T \nabla_{\mathbf{x}_i} J^v. \quad (14)$$

After replacing (14) and (13) into (12)

$$\nabla_{\mathbf{y}_i} J^v = \mathbf{R}_i^{-1} \mathbf{H}_i \mathbf{A}_i \nabla_{\mathbf{x}_i} J^v \in \mathbb{R}^{k_i}. \quad (15)$$

The matrix  $\mathbf{A}_i$  is an approximation of the covariance matrix of the errors in  $\mathbf{x}_i$  (see, e.g., Cohn 1997). If  $\mathbf{H}_i^T \nabla_{\mathbf{y}_i} J^v$  denotes the analysis space projection of the observation sensitivity vector (Baker and Daley 2000), then

$$\mathbf{H}_i^T \nabla_{\mathbf{y}_i} J^v = \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{H}_i \mathbf{A}_i \nabla_{\mathbf{x}_i} J^v \in \mathbb{R}^n. \quad (16)$$

*c. Sensitivity to the specification of the observation error covariance*

The observation error includes the instrument error and the error of representativeness (Daley 1991; Cohn 1997), and we refer to the work of Janjić and Cohn (2006), Frehlich (2006), and Chapnik et al. (2006) for recent efforts to properly incorporate the observation error into the DAS. In this section we provide the equations of the forecast sensitivity to the specification of the observation error covariance matrix  $\mathbf{R}_i$ ,  $i = 0, 1, \dots, N$ . The derivation involves Kronecker (tensor) products and we refer to Horn and Johnson (1991), Neudecker (1969), and Magnus and Neudecker (1999) for the matrix calculus and the matrix differentiation rules. It is convenient to introduce the vectorization notation: for an arbitrary  $(p, q)$  dimensional matrix  $\mathbf{X} \in \mathbb{R}^{p \times q}$  denote  $\mathbf{X} \cdot : \in \mathbb{R}^{pq}$  the column vector formed by stacking the columns of the matrix  $\mathbf{X}$  one underneath the other (see Magnus and Neudecker 1999, chapter 2). If  $\mathbf{X}(\cdot, j)$  denotes the  $j$ th column of the matrix  $\mathbf{X}$ , then

$$\mathbf{X} \cdot : \stackrel{\text{def}}{=} \begin{bmatrix} \mathbf{X}(\cdot, 1) \\ \mathbf{X}(\cdot, 2) \\ \vdots \\ \mathbf{X}(\cdot, q) \end{bmatrix},$$

such that componentwise  $\mathbf{X}_{k,j} = \mathbf{X} \cdot : [k + (j - 1)p]$  for  $1 \leq k \leq p, 1 \leq j \leq q$ .

Differentiating (6) with respect to  $(\mathbf{R}_i^{-1} \cdot)$

$$\nabla_{\mathbf{R}_i^{-1} \cdot : \mathbf{x}_0}^2 J = [H_i(\mathbf{x}_i) - \mathbf{y}_i] \otimes \mathbf{H}_i \mathbf{M}_{0,i} \in \mathbb{R}^{k_i^2 \times n}, \quad (17)$$

where the operator  $\otimes$  denotes the Kronecker product of two matrices. We recall that for two matrices  $\mathbf{X} \in \mathbb{R}^{p \times q}$  and  $\mathbf{Y} \in \mathbb{R}^{m \times n}$ , the Kronecker product  $\mathbf{X} \otimes \mathbf{Y}$  is defined as the  $(pm, qn)$  dimensional matrix

$$\mathbf{X} \otimes \mathbf{Y} = \begin{bmatrix} x_{11} \mathbf{Y} & \dots & x_{1q} \mathbf{Y} \\ \vdots & & \vdots \\ x_{p1} \mathbf{Y} & \dots & x_{pq} \mathbf{Y} \end{bmatrix} \in \mathbb{R}^{pm \times qn}.$$

Equations (5), (7), and (17) provide the forecast sensitivity to the inverse observation error covariance matrix

$$\begin{aligned} \nabla_{\mathbf{R}_i^{-1} \cdot :} J^v &= -[H_i(\mathbf{x}_i) - \mathbf{y}_i] \otimes \mathbf{H}_i \mathbf{M}_{0,i} \mathbf{A} \nabla_{\mathbf{x}_0^a} J^v \\ &= -[H_i(\mathbf{x}_i) - \mathbf{y}_i] \otimes \mathbf{R}_i \nabla_{\mathbf{y}_i} J^v \in \mathbb{R}^{k_i^2}. \end{aligned} \quad (18)$$

The observation sensitivity (12) is used to establish the last equality in (18). Using the differential of the inverse formula  $d(\mathbf{X}^{-1} \cdot) = -(\mathbf{X}^{-T} \otimes \mathbf{X}^{-1})d(\mathbf{X} \cdot)$  (see Magnus and Neudecker 1999, chapter 9) and the symmetry of  $\mathbf{R}_i^{-1}$ , it follows that

$$\nabla_{\mathbf{R}_i}(\mathbf{R}_i^{-1} \cdot) = -\mathbf{R}_i^{-1} \otimes \mathbf{R}_i^{-1}, \quad (19)$$

such that the forecast sensitivity to the observation error covariance is

$$\begin{aligned} \nabla_{\mathbf{R}_i} J^v &= \nabla_{\mathbf{R}_i}(\mathbf{R}_i^{-1} \cdot) \nabla_{\mathbf{R}_i^{-1} \cdot :} J^v \\ &= (\mathbf{R}_i^{-1} \otimes \mathbf{R}_i^{-1}) \{ [H_i(\mathbf{x}_i) - \mathbf{y}_i] \otimes \mathbf{R}_i \nabla_{\mathbf{y}_i} J^v \} \in \mathbb{R}^{k_i^2}. \end{aligned} \quad (20)$$

The mixed-product rule  $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$  simplifies (20) to

$$\nabla_{\mathbf{R}_i} J^v = \{ \mathbf{R}_i^{-1} [H_i(\mathbf{x}_i) - \mathbf{y}_i] \} \otimes \nabla_{\mathbf{y}_i} J^v \in \mathbb{R}^{k_i^2}. \quad (21)$$

It is noticed that the sensitivities (21) are obtained with a modest additional computational effort after the evaluation of the sensitivities to observations. Denoting  $\sigma_i^2 \in \mathbb{R}^{k_i}$  the vector of observation error variances at time  $t_i$ , from (21) it follows that

$$\nabla_{\sigma_i^2} J^v = \{ \mathbf{R}_i^{-1} [H_i(\mathbf{x}_i) - \mathbf{y}_i] \} \circ \nabla_{\mathbf{y}_i} J^v \in \mathbb{R}^{k_i}, \quad (22)$$

where the operator  $\circ$  represents the Hadamard (elementwise) product of two vectors. The equations of the forecast sensitivity to the specification of the observation error statistics and the close relationship to the observation sensitivity are thus provided. Often in practice the observation errors are assumed to be uncorrelated such that  $\mathbf{R}_i$  is taken to be a diagonal matrix,  $\mathbf{R}_i =$

$\text{diag}(\sigma_{i,1}^2, \dots, \sigma_{i,k_i}^2)$ . In this case (22) is written componentwise:

$$\frac{\partial J^v}{\partial \sigma_{i,j}^2} = \frac{1}{\sigma_{i,j}^2} [H_i(\mathbf{x}_i) - \mathbf{y}_i]_j \frac{\partial J^v}{\partial (\mathbf{y}_i)_j}, \quad j = 1, 2, \dots, k_i. \quad (23)$$

For each component  $(\mathbf{y}_i)_j$  of the observational data vector  $\mathbf{y}_i$  at time  $t_i$ , the sensitivity to the observation error variance  $\sigma_{i,j}^2$  is related to the observation sensitivity by the proportionality coefficient  $[H_i(\mathbf{x}_i) - \mathbf{y}_i]_j / \sigma_{i,j}^2$ —the analysis fit to data divided by the observation error variance. While important, this ratio does not directly provide sensitivity information.

#### d. Sensitivity to background and background error covariance

Differentiating (6) with respect to  $\mathbf{x}_b$ ,

$$\nabla_{\mathbf{x}_b \mathbf{x}_0}^2 J = -\mathbf{B}^{-1}. \quad (24)$$

Equations (4), (7), and (24) provide the analysis sensitivity to background,

$$\nabla_{\mathbf{x}_b} \mathbf{x}_0^a = \mathbf{B}^{-1} \mathbf{A} \in \mathbb{R}^{n \times n}. \quad (25)$$

Chain rule differentiation,

$$\nabla_{\mathbf{x}_b} J^v = \nabla_{\mathbf{x}_b} \mathbf{x}_0^a \nabla_{\mathbf{x}_0^a} J^v,$$

and (25) provide the expression of the forecast sensitivity to background,

$$\nabla_{\mathbf{x}_b} J^v = \mathbf{B}^{-1} \mathbf{A} \nabla_{\mathbf{x}_0^a} J^v \in \mathbb{R}^n. \quad (26)$$

Differentiating (6) with respect to  $(\mathbf{B}^{-1})$ ,

$$\nabla_{\mathbf{B}^{-1}; \mathbf{x}_0}^2 J = (\mathbf{x}_0^a - \mathbf{x}_b) \otimes \mathbf{I} \in \mathbb{R}^{n^2 \times n}, \quad (27)$$

where  $\mathbf{I} \in \mathbb{R}^{n \times n}$  is the identity matrix. Equations (5), (7), and (27) provide the forecast sensitivity to the inverse background error covariance:

$$\nabla_{\mathbf{B}^{-1}} J^v = -(\mathbf{x}_0^a - \mathbf{x}_b) \otimes \mathbf{A} \nabla_{\mathbf{x}_0^a} J^v \in \mathbb{R}^{n^2}, \quad (28)$$

and using chain rule differentiation it follows that

$$\begin{aligned} \nabla_{\mathbf{B}} J^v &= \nabla_{\mathbf{B}} (\mathbf{B}^{-1}) \nabla_{\mathbf{B}^{-1}} J^v \\ &= (\mathbf{B}^{-1} \otimes \mathbf{B}^{-1}) [(\mathbf{x}_0^a - \mathbf{x}_b) \otimes \mathbf{A} \nabla_{\mathbf{x}_0^a} J^v] \\ &= [\mathbf{B}^{-1} (\mathbf{x}_0^a - \mathbf{x}_b)] \otimes (\mathbf{B}^{-1} \mathbf{A} \nabla_{\mathbf{x}_0^a} J^v) \in \mathbb{R}^{n^2}, \end{aligned} \quad (29)$$

which is the equation of the forecast sensitivity to the background error covariance. A close relationship to the sensitivity to background is established from (26) and (29):

TABLE 1. The equations of the forecast sensitivity to 4D-Var DAS input parameters.

DAS input	Forecast sensitivity
Observations	$\nabla_{\mathbf{y}_i} J^v = \mathbf{R}_i^{-1} \mathbf{H}_i \mathbf{M}_{0,i} \mathbf{A} \nabla_{\mathbf{x}_0^a} J^v$
Observation error variance	$\nabla_{\sigma_i^2} J^v = \{\mathbf{R}_i^{-1} [H_i(\mathbf{x}_i) - \mathbf{y}_i]\} \circ \nabla_{\mathbf{y}_i} J^v$
Observation error covariance	$\nabla_{\mathbf{R}_i} J^v = \{\mathbf{R}_i^{-1} [H_i(\mathbf{x}_i) - \mathbf{y}_i]\} \otimes \nabla_{\mathbf{y}_i} J^v$
Background estimate	$\nabla_{\mathbf{x}_b} J^v = \mathbf{B}^{-1} \mathbf{A} \nabla_{\mathbf{x}_0^a} J^v$
Background error variance	$\nabla_{\sigma_b^2} J^v = [\mathbf{B}^{-1} (\mathbf{x}_0^a - \mathbf{x}_b)] \circ \nabla_{\mathbf{x}_b} J^v$
Background error covariance	$\nabla_{\mathbf{B}} J^v = [\mathbf{B}^{-1} (\mathbf{x}_0^a - \mathbf{x}_b)] \otimes \nabla_{\mathbf{x}_b} J^v$

$$\nabla_{\mathbf{B}} J^v = [\mathbf{B}^{-1} (\mathbf{x}_0^a - \mathbf{x}_b)] \otimes \nabla_{\mathbf{x}_b} J^v \in \mathbb{R}^{n^2}. \quad (30)$$

The sensitivity with respect to any of the entries of  $\mathbf{B}$  is thus easily available once the sensitivity to the background is computed. In particular, the forecast sensitivity with respect to the vector of background error variances  $\sigma_b^2 \in \mathbb{R}^n$  is

$$\nabla_{\sigma_b^2} J^v = [\mathbf{B}^{-1} (\mathbf{x}_0^a - \mathbf{x}_b)] \circ \nabla_{\mathbf{x}_b} J^v \in \mathbb{R}^n. \quad (31)$$

A summary of the forecast sensitivity equations to various input parameters of a 4D-Var DAS [(12), (21), (22), (26), (30), and (31)] is provided in Table 1.

#### e. The 3D-Var sensitivity equations

An extensive amount of research on observation sensitivity has been done in the context of 3D-Var data assimilation. A comprehensive analysis and applications to targeted observations may be found in the work of Baker (2000) and Doerenbecher (2002).

The 3D-Var cost corresponds to  $N = 0$  in (1) and thus the time index  $i$  is omitted. Under the assumption that the observation operator is linear, the 3D-Var inverse Hessian matrix is the state-independent matrix

$$\mathbf{A} = (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1}. \quad (32)$$

Equations (12) and (26) of the sensitivity to observations and background become, respectively,

$$\nabla_{\mathbf{y}} J^v = \mathbf{R}^{-1} \mathbf{H} (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \nabla_{\mathbf{x}_0^a} J^v \quad (33)$$

$$\nabla_{\mathbf{x}_b} J^v = \mathbf{B}^{-1} (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \nabla_{\mathbf{x}_0^a} J^v. \quad (34)$$

Introducing the Kalman gain matrix  $\mathbf{K}$ ,

$$\mathbf{K} = (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1}, \quad (35)$$

and noticing that

$$\mathbf{K}^T = \mathbf{R}^{-1} \mathbf{H} (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1},$$

$$\mathbf{B}^{-1} (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} = \mathbf{I} - \mathbf{H}^T \mathbf{K}^T, \quad (36)$$

where  $\mathbf{I} \in \mathbb{R}^{n \times n}$  is the identity matrix, the sensitivity Eqs. (33) and (34) may be written

$$\nabla_{\mathbf{y}} J^v = \mathbf{K}^T \nabla_{\mathbf{x}_0^a} J^v \quad \text{and} \quad (37)$$

$$\nabla_{\mathbf{x}_b} J^v = (\mathbf{I} - \mathbf{H}^T \mathbf{K}^T) \nabla_{\mathbf{x}_0^a} J^v. \quad (38)$$

Consistency with the sensitivity equations derived in the original work of Baker and Daley (2000) from the analysis equation  $\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$  is thus proven. It is noticed that in this context the sensitivity to the observation error covariance (21) becomes

$$\nabla_{\mathbf{R}} J^v = [\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}_0^a - \mathbf{y})] \otimes (\mathbf{K}^T \nabla_{\mathbf{x}_0^a} J^v) \quad (39)$$

and the sensitivity to the observation error variance (22) becomes

$$\nabla_{\sigma^2} J^v = [\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}_0^a - \mathbf{y})] \circ (\mathbf{K}^T \nabla_{\mathbf{x}_0^a} J^v). \quad (40)$$

Replacing (38) into (30), the forecast sensitivity to the background error covariance is

$$\nabla_{\mathbf{B}} J^v = [\mathbf{B}^{-1}(\mathbf{x}_0^a - \mathbf{x}_b)] \otimes [(\mathbf{I} - \mathbf{H}^T \mathbf{K}^T) \nabla_{\mathbf{x}_0^a} J^v] \quad (41)$$

and the sensitivity to the background error variance is

$$\nabla_{\sigma_b^2} J^v = [\mathbf{B}^{-1}(\mathbf{x}_0^a - \mathbf{x}_b)] \circ [(\mathbf{I} - \mathbf{H}^T \mathbf{K}^T) \nabla_{\mathbf{x}_0^a} J^v]. \quad (42)$$

*f. On the relevance of the sensitivity equations*

For practical applications, the forecast sensitivity equations derived in this study should be cautiously used to assess the observation impact, and systematic observing system experiments are required to evaluate the impact of specific sets of observations on analysis and forecasts (Atlas 1997). Some of the issues that need to be addressed are discussed here. The actual forecast impact of a specific measurement depends not only on the magnitude of the corresponding observation sensitivity but also on the innovation component (Doerenbecher and Bergot 2001; Langland and Baker 2004; Zhu and Gelaro 2008). Therefore, as pointed out by an anonymous reviewer, large sensitivities do not necessarily imply a large forecast impact. The nonlinear relationship between the forecast error and observations (or other DAS input) must be considered since the forecast aspect  $J^v$  at the verification time is a nonlinear function of the analysis  $\mathbf{x}_0^a$  and, implicitly, of the observations. The sensitivities  $\nabla_{\mathbf{y}_i} J^v$  may be used to provide only a *first-order* measure of the observation impact and thus their relevance is determined by the higher-order terms in the Taylor series. A reliable estimate of

the observation impact requires an analysis beyond first order, as explained by Errico (2007). The approach of Langland and Baker (2004) provides a higher-order accuracy in the estimation of the observation impact on the forecast error by combining adjoint sensitivity gradients from two trajectories (background and analysis). Another issue is related to the fact that the sensitivity equations are derived from the optimality condition (6) whereas in the practical implementation the minimization process is terminated when the gradient satisfies a certain convergence criteria or simply after a prescribed number of iterations. To obtain an estimate of the observation sensitivity that is consistent to the data assimilation process, Zhu and Gelaro (2008) implemented the adjoint of the minimization algorithm in the grid-point statistical interpolation analysis scheme (Wu et al. 2002). In this way it is acknowledged that data assimilation is the sequence of computational steps performed to obtain the analysis from the input. The adjoint of the computation from input to analysis to forecast provides sensitivity estimates fully consistent with the data assimilation procedure.

**3. Numerical aspects and illustrative experiments**

Evaluation of the forecast sensitivity to observations in a 4D-Var DAS requires a significant computational effort in addition to the data assimilation, and the main stages involved are outlined below:

- S1— $\nabla_{\mathbf{x}_v} J^v(\mathbf{x}_v)$ : initialization of the adjoint model,
- S2— $\nabla_{\mathbf{x}_0^a} J^v(\mathbf{x}_v) = \mathbf{M}_{0,v}^T \nabla_{\mathbf{x}_v} J^v(\mathbf{x}_v)$ : adjoint model integration from  $t_v$  to  $t_0$ ,
- S3— $\mathbf{A}^{-1} \boldsymbol{\mu}_0 = \nabla_{\mathbf{x}_0^a} J^v(\mathbf{x}_v)$ : linear system for  $\boldsymbol{\mu}_0$  with matrix being the 4D-Var Hessian,
- S4— $\boldsymbol{\mu}_i = \mathbf{M}_{0,i} \boldsymbol{\mu}_0$ : tangent linear model integration from  $t_0$  to  $t_i$ , and
- S5— $\nabla_{\mathbf{y}_i} J^v(\mathbf{x}_v) = \mathbf{R}_i^{-1} \mathbf{H}_i \boldsymbol{\mu}_i$ : mapping on observation space, weighting.

It is noticed that after stage S3 is completed a single tangent linear model integration (S4) in the data assimilation window  $[t_0, t_N]$  is required to provide the sensitivities with respect to all available observational data. Stage S5 is performed at each observational time  $t_i$  and its computational demands are modest when the number of observations is small or the observational errors are uncorrelated—an assumption often made in practice. The sensitivity to background is obtained after the completion of stage S3,  $\nabla_{\mathbf{x}_b} J^v = \mathbf{B}^{-1} \boldsymbol{\mu}_0$ . The tangent linear and adjoint model integrations involved in stages S4 and S2, respectively, are fairly affordable and the required software tools are already in place in a

4D-Var DAS. Stage S3 dominates the computational load since an iterative procedure such as the conjugate gradient (CG) method must be used to *approximate* the solution  $\boldsymbol{\mu}_0$  of the linear system. The errors in the estimation of  $\boldsymbol{\mu}_0$  are further propagated into the observation sensitivity computations by the tangent linear model at stage S4; therefore, an increased accuracy in the solution is necessary for the sensitivity estimates to be reliable. An exact evaluation of the required Hessian vector products may be obtained with a second-order adjoint model (SOA; Le Dimet et al. 2002). Given that the software development and the computational demands of an SOA model are significant, gradient differences may be used to approximate  $\mathbf{A}^{-1}\mathbf{v} \approx [\nabla\mathcal{J}(\mathbf{x}_0 + \varepsilon\mathbf{v}) - \nabla\mathcal{J}(\mathbf{x}_0)]/\varepsilon$ . The relationship is exact for a quadratic 4D-Var cost functional (e.g., in the incremental 4D-Var scheme with only one outer loop). Preconditioning is required for the iterative method to be effective and information accumulated while solving the data assimilation problem (1) may be used to construct an appropriate preconditioner. The use of limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) preconditioners in conjunction with the CG method is particularly effective in problems where matrix–vector products are expensive to compute and when a low accuracy in the solution is required, as shown in the work of Morales and Nocedal (2000). BFGS and L-BFGS approximations to the inverse Hessian matrix have been proven to provide efficient preconditioners in data assimilation (Fisher and Courtier 1995; Veersé et al. 2000).

#### a. Experimental setup

Numerical experiments are set up with a finite-volume global shallow-water (SW) model of Lin and Rood (1997) at a resolution of  $2.5^\circ \times 2.5^\circ$  and with a time step  $\Delta t = 600$  s. The discrete state vector is of dimension  $n \sim 3 \times 10^4$ ,  $\mathbf{x} = (h, u, v)$ , where  $h$  is the geopotential height and  $u$  and  $v$  are the zonal and meridional wind velocities, respectively. Given the simplicity of our model, these experiments are merely used to illustrate the sensitivity concepts presented in this work and some of the potential applications.

An idealized 4D-Var DAS is considered in the twin experiments framework: a reference initial state  $\mathbf{x}_0^t$  (“the truth”) is taken from the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (ERA-40) 500-hPa data valid for 0600 UTC 15 March 2002; the background estimate  $\mathbf{x}_b$  to  $\mathbf{x}_0^t$  is obtained from a 6-h integration of the SW model initialized at  $t_0 - 6$  h with ERA-40 500-hPa data valid for 0000 UTC 15 March 2002. Isopleths of the errors in the

background estimate  $\mathbf{x}_0^t - \mathbf{x}_b$  are shown in Fig. 1 using a total energy norm to obtain gridpoint values ( $\text{m}^2 \text{s}^{-2}$ ),

$$\|\delta\mathbf{x}\|_{\mathbf{E}}^2 = \frac{1}{2} \left( |\delta u|^2 + |\delta v|^2 + \frac{g}{h} |\delta h|^2 \right), \quad (43)$$

where  $g$  is the gravitational constant and  $\bar{h}$  is the mean height of the reference data. “Observational data” for the assimilation procedure are generated from a model trajectory initialized with  $\mathbf{x}_0^t$  and corrupted with random errors from a normal distribution  $N(0, \sigma^2)$ . The standard deviation is chosen as  $\sigma_h = 5$  m for the height and  $\sigma_u = \sigma_v = 0.5 \text{ m s}^{-1}$  for the velocities, and the observation errors are assumed uncorrelated (diagonal  $\mathbf{R}_i$ ). The background errors are assumed uncorrelated (diagonal  $\mathbf{B}$ ) and are specified at a ratio  $\sigma_b^2/\sigma^2 = 4$  to the observations. An inspection of the contours in Fig. 1 reveals that the background errors are in general misrepresented in the DAS and we will return to this aspect later in this section. A data assimilation time interval  $[t_0, t_0 + 6\text{h}]$  is considered with data provided at each time step ( $N = 36$ ) and uniformly distributed on a  $10^\circ \times 10^\circ$  model subgrid. The observation operator  $H$  is thus represented by a matrix with entries 0 and 1 only and there are 648 observations for each  $h, u$ , and  $v$  per time step,  $k_i = 1944, i = 0, 1, \dots, 36$ . The L-BFGS algorithm (Liu and Nocedal 1989) was used to solve the nonlinear 4D-Var minimization problem (1) with a convergence criteria  $(1/n) \|\nabla_{\mathbf{x}_v} \mathcal{J}\| \leq 10^{-3}$ , where  $\|\cdot\|$  denotes the Euclidean norm in  $\mathbb{R}^n$ . The evolution of the cost functional and its gradient during the 4D-Var minimization process are displayed in Fig. 2. At the analysis  $\mathbf{x}_0^a$ , the norm of the gradient was found to be  $(1/n) \|\nabla_{\mathbf{x}_v} \mathcal{J}(\mathbf{x}_0^a)\| \sim 0.88 \times 10^{-3}$ . The errors in the background estimate  $\mathbf{x}_0^t - \mathbf{x}_b$  and the analysis errors  $\mathbf{x}_0^t - \mathbf{x}_0^a$ , averaged over the longitudinal direction, are displayed in Fig. 3, and it is noticed that a significant reduction in the initial-condition error was achieved through the assimilation procedure.

#### b. Forecast error sensitivity analysis

At the verification time  $t_v = t_0 + 30$  h we consider as the reference state  $\mathbf{x}_v^t = \mathcal{M}_{t_0 \rightarrow t_v}(\mathbf{x}_0^t)$  and the forecast  $\mathbf{x}_v^f = \mathcal{M}_{t_0 \rightarrow t_v}(\mathbf{x}_0^a)$ . The region  $\mathcal{D}_v = 50^\circ\text{--}65^\circ\text{N}, 60^\circ\text{--}30^\circ\text{W}$  is taken as the verification domain and the forecast error  $\|\mathbf{x}_v^t - \mathbf{x}_v^f\|_{\mathbf{E}}^2$  over  $\mathcal{D}_v$  is displayed in Fig. 4. The functional  $\mathcal{J}^v$  is defined as a scalar measure of the forecast error over  $\mathcal{D}_v$ ,

$$\mathcal{J}^v = (\mathbf{x}_v^f - \mathbf{x}_v^t)^T \mathbf{P}^T \mathbf{E} \mathbf{P} (\mathbf{x}_v^f - \mathbf{x}_v^t), \quad (44)$$

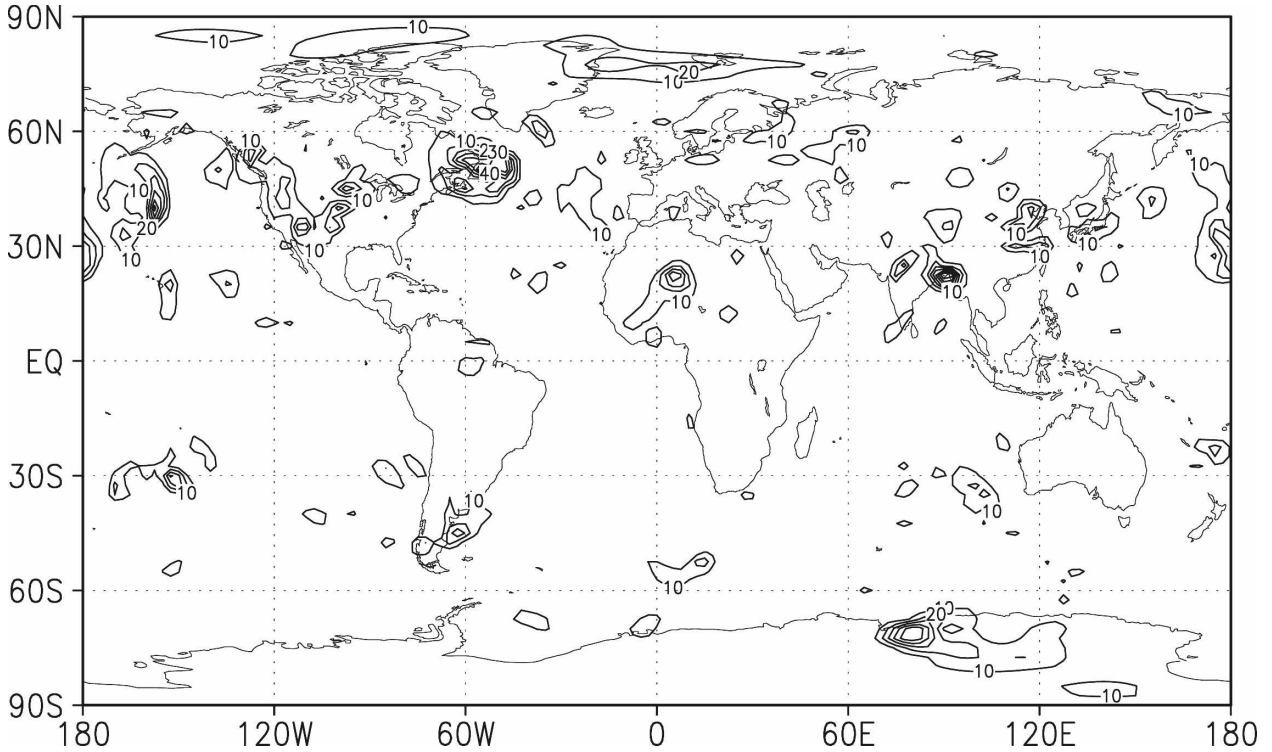


FIG. 1. Total energy errors ( $\text{m}^2 \text{s}^{-2}$ ) in the background estimate to the initial conditions. The contour interval is  $10 \text{ m}^2 \text{s}^{-2}$ .

where  $\mathbf{P}$  is the projection operator on  $\mathcal{D}_v$  and  $\mathbf{E}$  is a diagonal matrix of the total energy norm. The sensitivities of  $J^v$  with respect to the ensemble DAS input  $[\mathbf{y}_i, \mathbf{R}_i, \mathbf{x}_b, \mathbf{B}]$  were evaluated by performing the computational stages S1–S5. The first-order adjoint of the finite volume SW model used in this study was developed in the work of Akella and Navon (2006). The simplicity of the SW model allowed us to implement an SOA model with the aid of automatic differentiation software (Giering and Kaminski 1998). The verification of the SOA model included a Taylor series test (Daescu and Navon 2007) and a Hessian symmetry test using pairs of random vectors. An approximate solution to the linear system at stage S3 was computed by imposing the convergence criteria  $(1/n)\|\mathbf{A}^{-1}\boldsymbol{\mu}_0 - \nabla_{\mathbf{x}_0} J^v\| < 10^{-4}$ .

1) OBSERVATION SENSITIVITY

The sensitivity of the forecast error with respect to time–space-distributed observational data  $\mathbf{y}_i$  and to the specification of the corresponding observation error variances  $\boldsymbol{\sigma}_i^2$  is evaluated in the 4D-Var context according to (12) and (23), respectively, for  $i = 0, 1, \dots, 36$ . The time dimension of the sensitivity to data is characteristic of the 4D-Var framework and in Fig. 5a we display the observation space Euclidean norm of the forecast error gradients  $\nabla_{h_i^o} J^v, \nabla_{u_i^o} J^v, \nabla_{v_i^o} J^v$  correspond-

ing to the observational data components  $\mathbf{y}_i = (h_i^o, u_i^o, v_i^o)$  as a function of the observation time  $t_i - t_0$ . The Euclidean norm of the forecast error sensitivity to the specification of the observation error variances  $\nabla_{\boldsymbol{\sigma}_{i,h}^2} J^v, \nabla_{\boldsymbol{\sigma}_{i,u}^2} J^v, \nabla_{\boldsymbol{\sigma}_{i,v}^2} J^v$  is also displayed in Fig. 5b. An increased sensitivity is noticed with respect to data toward the end of the assimilation window indicating that the time distribution of the observations plays an essential role

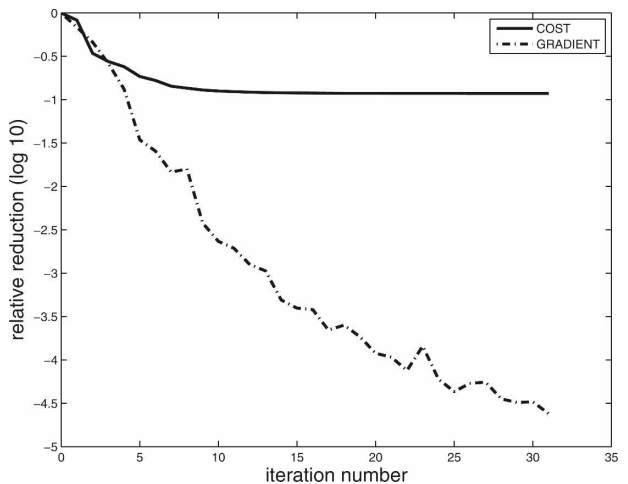


FIG. 2. Relative reduction in the cost function and the gradient norm during the 4D-Var minimization process.



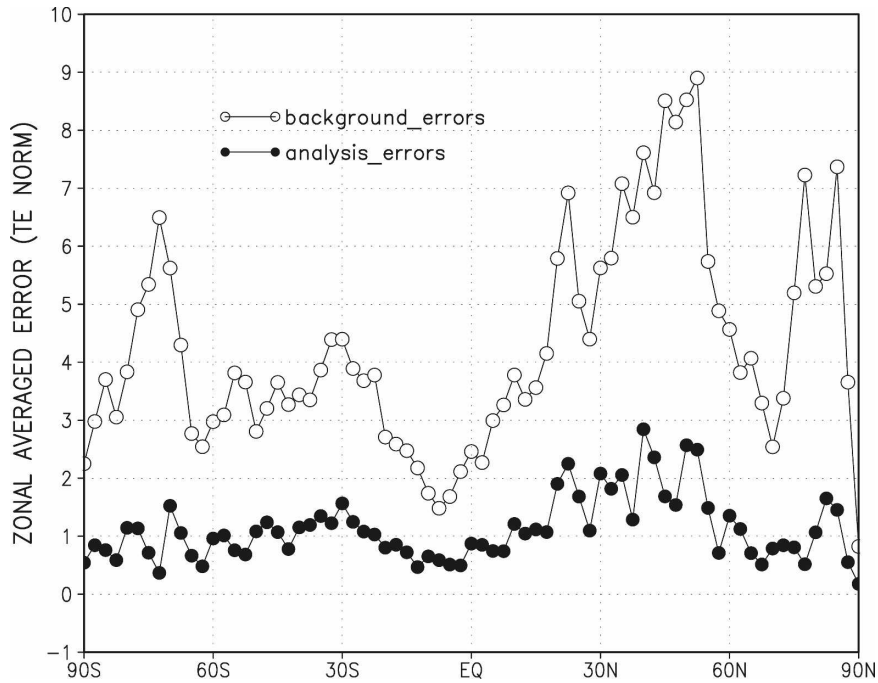


FIG. 3. Zonally averaged total energy errors ( $\text{m}^2 \text{s}^{-2}$ ) in the prior estimate to the initial conditions and after the data assimilation process.

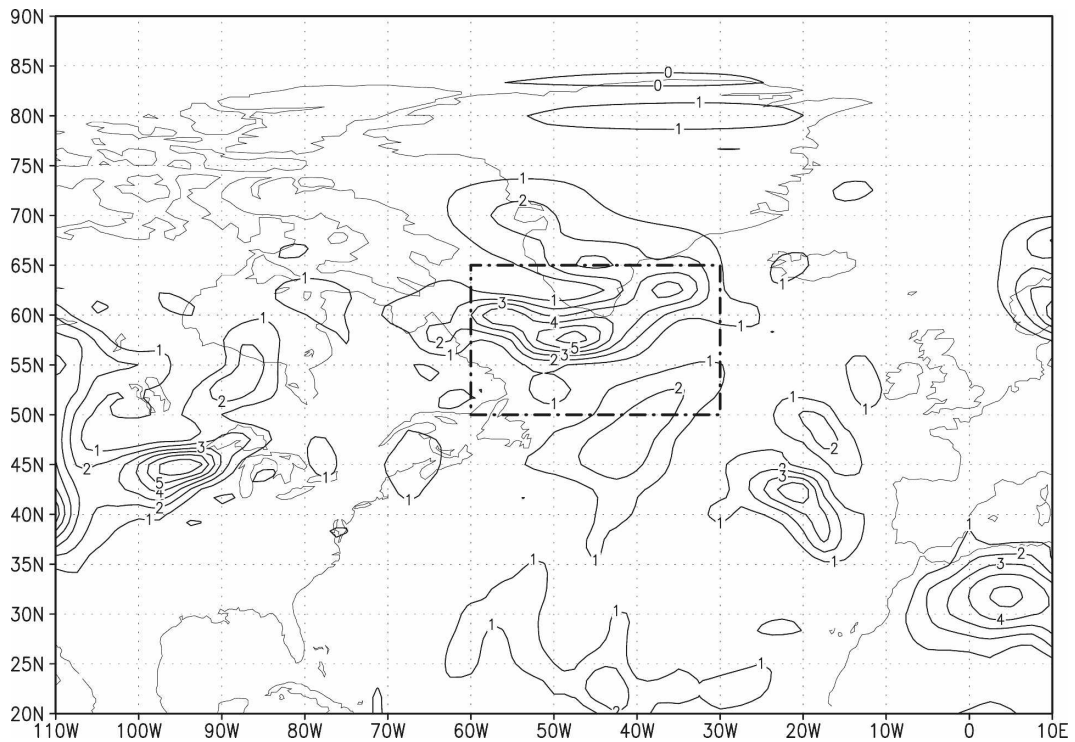


FIG. 4. Total energy forecast error ( $\text{m}^2 \text{s}^{-2}$ ) at the verification time  $t_v = t_0 + 30 \text{ h}$  and selection of the verification domain. The contour interval is  $1 \text{ m}^2 \text{s}^{-2}$ .

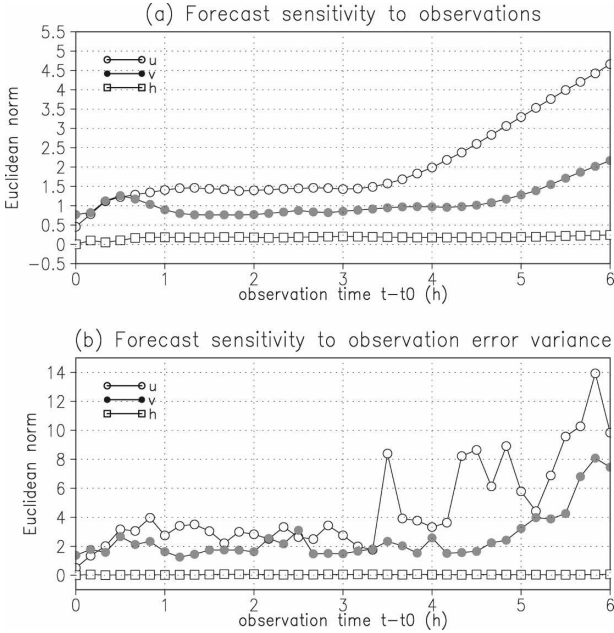


FIG. 5. (a) Forecast sensitivity to observational data  $h_i^o$ ,  $u_i^o$ ,  $v_i^o$  as a function of the observation time  $t_i - t_0$ . For each data component the corresponding Euclidean norm of the gradient is evaluated in the observation space:  $\nabla_{h_i^o} J^v$  ( $\text{m s}^{-2}$ ),  $\nabla_{u_i^o} J^v$  ( $\text{m s}^{-1}$ ), and  $\nabla_{v_i^o} J^v$  ( $\text{m s}^{-1}$ ). (b) Forecast sensitivity with respect to the observation error variances  $\sigma_{i,h}^2$ ,  $\sigma_{i,u}^2$ ,  $\sigma_{i,v}^2$  as a function of the observation time  $t_i - t_0$ .

on the observation impact. For instance, the sensitivity to the  $u$ -velocity data at the end of the assimilation window is higher by as much as a factor of 5 as compared to the data near the initial time. The sensitivity to velocity data  $u_i^o$ ,  $v_i^o$  was found to be an order of magnitude higher than the sensitivity to the height data  $h_i^o$ , whereas the sensitivity to  $\sigma_{i,h}^2$  was over two orders of magnitude lower than the sensitivity to  $\sigma_{i,u}^2$  and  $\sigma_{i,v}^2$ . From (12) it is noticed that for uncorrelated observational errors the observation sensitivity is inversely proportional to the observation error variance such that an increased error variance may result in a low magnitude observation sensitivity (in our experiments  $\sigma_h^2 = 25 \text{ m}^2$  for the height data versus  $\sigma_u^2 = \sigma_v^2 = 0.25 \text{ m}^2 \text{ s}^{-2}$  for the velocities). This aspect must be taken into account in practical applications when evaluating the forecast sensitivity with respect to various data components. For the velocity data the sensitivity to the observation error variance has a larger magnitude than the observation sensitivity (by as much as a factor of 3–4), indicating that proper specification of the observation error statistics has a significant impact on the forecast error. The proportionality coefficient in (23) depends not only on the error variance but also on the analysis fit to data. The increased variability noticed in the time evolution

of the sensitivity to the error variances, as compared to the observation sensitivity, indicates a nonuniform fit to data over the assimilation time interval.

An analysis of the distribution of the observation sensitivity in the observation space reveals that the specified forecast aspect  $J^v$  exhibits a large sensitivity with respect to only a few of the observations in the DAS. As a measure of the forecast sensitivity to observation and error variance at each data location over the assimilation time interval we consider the time cumulative magnitude of the sensitivities  $\sum_{i=0}^N |\nabla_{y_i} J^v|$  and  $\sum_{i=0}^N |\nabla_{\sigma_i^2} J^v|$ , respectively. The locations of the observations and error variances of the largest forecast sensitivity are displayed in Fig. 6 for each of the  $u$ ,  $v$ , and  $h$  data. A distinct configuration is noticed for each data component, indicating that the location of observations and error variances that provide a potentially large forecast impact depends on the data type. In the 4D-Var context the forecast sensitivities to data are time-space-varying fields. To illustrate this aspect, in Fig. 7 observation locations with sensitivities of largest magnitude at  $t_i - t_0 = 3 \text{ h}$  and  $t_i - t_0 = 6 \text{ h}$  are displayed for each of the  $u$ ,  $v$ , and  $h$  data.

## 2) BACKGROUND SENSITIVITY

Forecast sensitivity to the background estimate  $\mathbf{x}_b$  and to the specification of the background error variance  $\sigma_b^2$  is evaluated according to (26) and (31), respectively. In our setup,  $\mathbf{B}$  is a diagonal matrix,  $\mathbf{B} = \text{diag}(\sigma_b^2)$ ,  $\sigma_b^2 = (\sigma_{b,h}^2, \sigma_{b,u}^2, \sigma_{b,v}^2)^T$ . Forecast sensitivities to the background velocities  $\nabla_{u_b} J^v$ ,  $\nabla_{v_b} J^v$  and to the specification of the corresponding error variances  $\nabla_{\sigma_{b,u}^2} J^v$ ,  $\nabla_{\sigma_{b,v}^2} J^v$  are shown in Fig. 8. For comparison, the sensitivities to the analysis velocities  $\nabla_{u_0^a} J^v$ ,  $\nabla_{v_0^a} J^v$  are also displayed in Fig. 8. It is noticed that the analysis sensitivity is greater in magnitude than the background sensitivity and therefore *supersensitivity* (Baker and Daley 2000) is not observed in these experiments. The sensitivities  $\nabla_{h_b} J^v$  and  $\nabla_{\sigma_{b,h}^2} J^v$  (not shown) were found to be a few orders of magnitude lower than the sensitivities to the background velocities and to the velocities error variances, respectively.

The sensitivity to the background error variances identifies the state components (regions) where the specification of the error statistics has a potentially high impact on the forecast. In particular, misspecification of the errors in the background velocities in the southwest region of the verification domain may contribute significantly to the forecast errors. The sensitivity to the background variances  $\sigma_{b,u}^2$ ,  $\sigma_{b,v}^2$  has a much larger magnitude than the sensitivity to the background  $u_b$ ,  $v_b$  (by as much as a factor of 10 for the  $u$  velocity), indicating

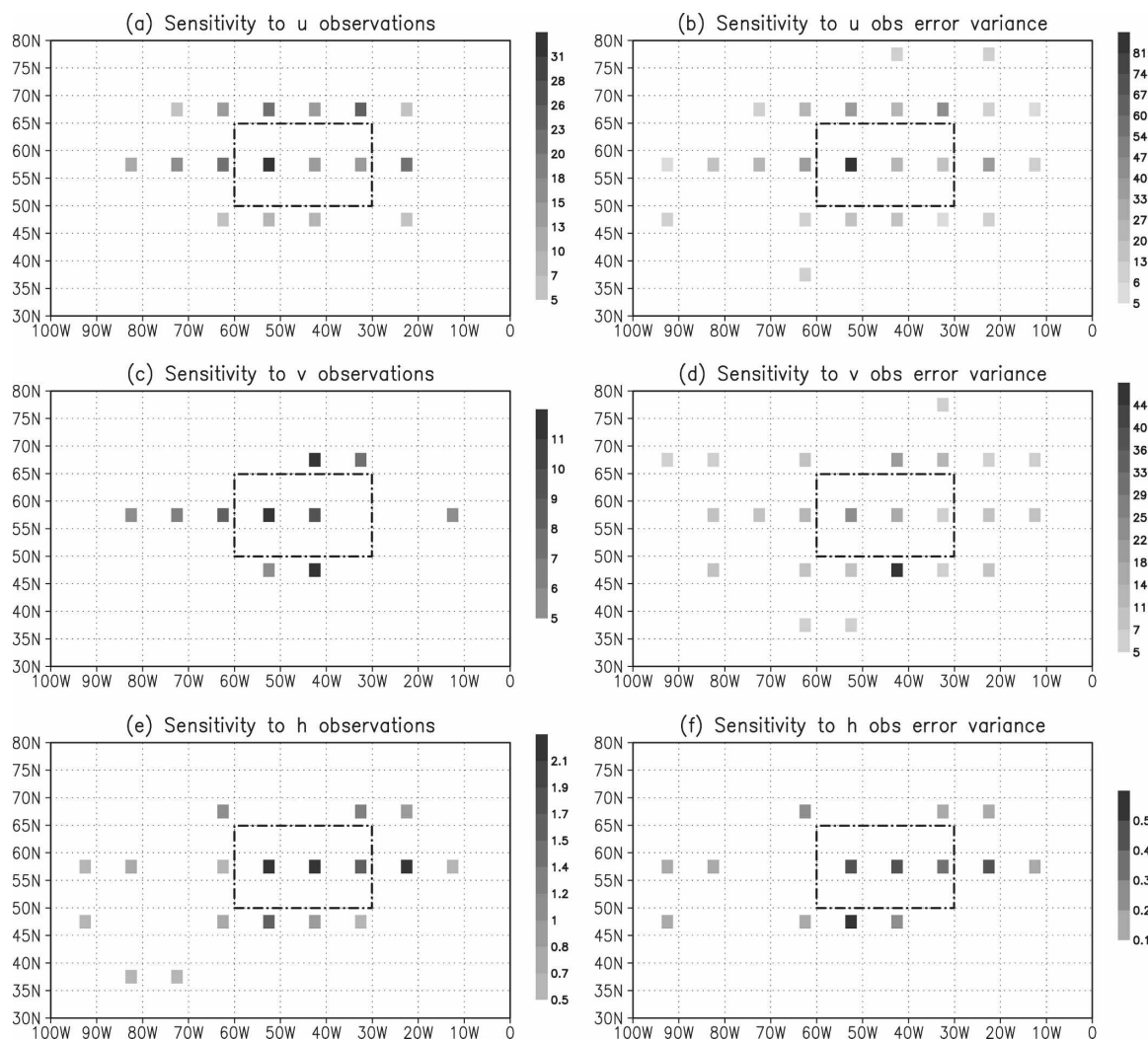


FIG. 6. Time cumulative magnitude of the forecast sensitivity to each of the data components  $u$ ,  $v$ , and  $h$ , and the sensitivity to the corresponding observation error variances  $\sigma_u^2$ ,  $\sigma_v^2$ ,  $\sigma_h^2$ .

that proper input of the background error statistics may be of much greater importance than the specification of the initial guess state.

### c. Improving the forecast

The information provided by the forecast sensitivity to the DAS input ( $\mathbf{y}_i$ ,  $\mathbf{R}_i$ ,  $\mathbf{x}_b$ ,  $\mathbf{B}$ ) may be used to identify the input components that have a significant impact on the forecast aspect  $J^n$ . We will refer to the previous 4D-Var experimental setup as DAS-1. Additional 4D-Var experiments are used in this section to illustrate some of the potential applications of the sensitivity analysis: forecast error reduction through collection of additional data (DAS-2) and through an improved representation of the background error (DAS-3). A 4D-Var experiment that combines the features of

DAS-2 and DAS-3 (i.e., both supplementary data and appropriate background weights are prescribed to the data assimilation procedure) is also considered (DAS-4). Since a model-generated reference state  $\mathbf{x}_0^t$  is used to define the forecast error (44), the issue of dealing with the SW model error to atmospheric dynamics is avoided; however, overly optimistic results may be obtained [the “identical twin” problem as discussed by Atlas (1997)].

The DAS-2 setup differs from the DAS-1 setup in one aspect only: the assimilation procedure includes supplementary ( $h$ ,  $u$ ,  $v$ ) data collected at all time steps at a few additional grid points, to be determined. The DAS-3 setup also differs from the DAS-1 setup in one aspect only: the error in the background estimate is prescribed accurately as  $\mathbf{x}_0^t - \mathbf{x}_b$  for a few state com-

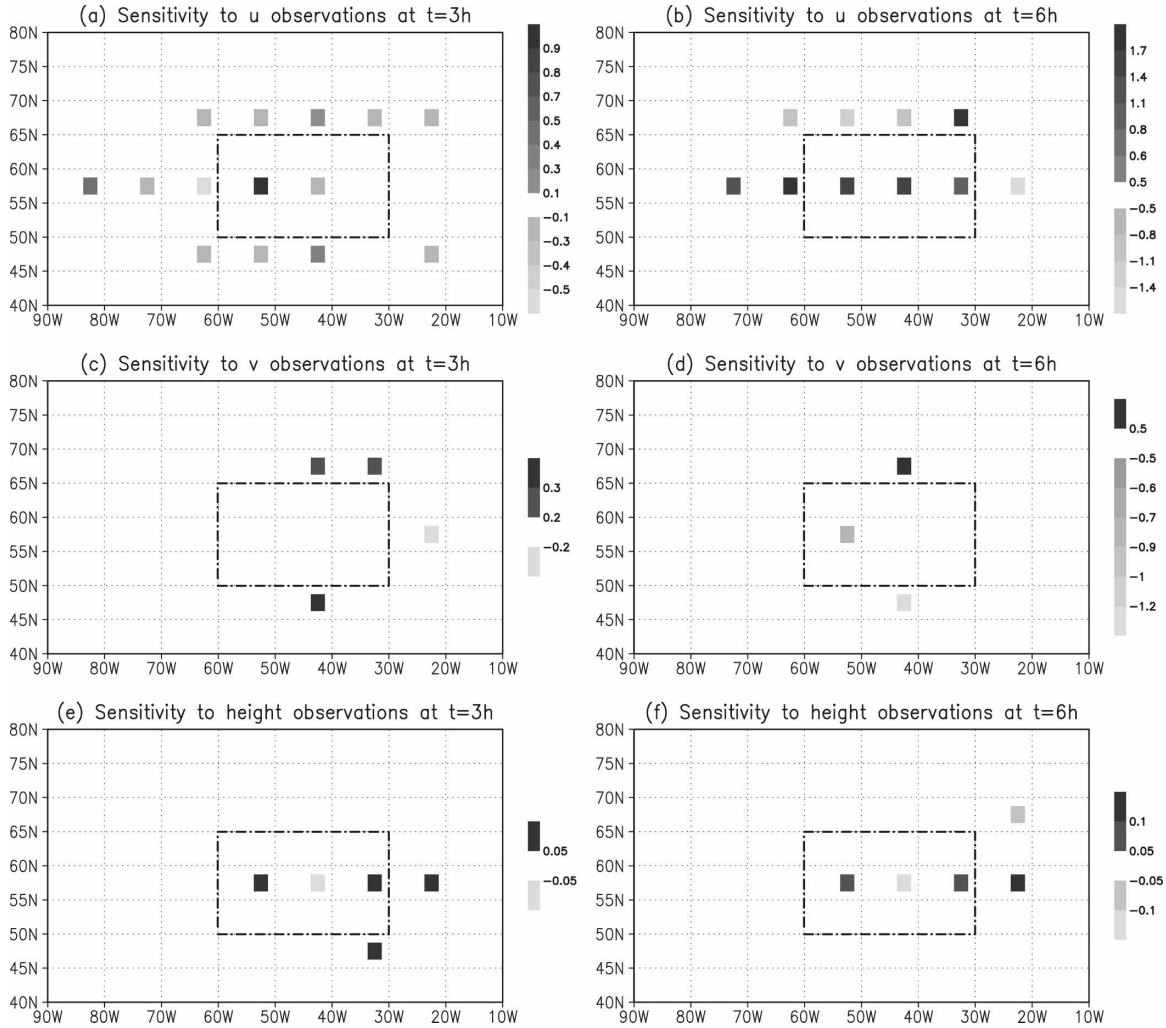


FIG. 7. Location of data with largest observation sensitivity magnitude at time  $t = 3$  h and  $t = 6$  h of the assimilation window  $[0, 6]$  h. The units are meters per second in (a)–(d) and meters per second squared in (e) and (f).

ponents, to be determined. Guidance for adjusting DAS-2, DAS-3, and DAS-4 is provided by the DAS-I forecast error sensitivity fields to velocity observations  $\sum_{i=0}^N (|\nabla_{u_i^p} J^p| + |\nabla_{v_i^p} J^p|)$  and to the specification of the background error variances  $|\nabla_{\sigma_{a,b}^2} J^p| + |\nabla_{\sigma_{u,b}^2} J^p|$ , respectively. Locations of the largest sensitivity magnitude are displayed in Fig. 9 and it is noticed that observations with a potentially high forecast impact are located in the central and northern region of the verification domain, whereas the specification of the background error variances is particularly important for the state components in the southwest region of the verification domain. In DAS-2, five additional observations are inserted at locations marked by open circles in Fig. 9. In DAS-3, the background weights are accurately prescribed in the region 50°–55°N, 60°–50°W and the DAS-4 setup incorporates both additional data and ac-

curate background weights, as indicated in Fig. 9. For each 4D-Var experiment the corresponding forecast errors over the verification domain are presented in Fig. 10 using a total energy norm. It is noticed that both DAS-2 and DAS-3 experiments resulted in a significant reduction in the forecast error as compared with DAS-1. The mean forecast error over  $\mathcal{D}_v$ , expressed as the sum of the gridpoint errors divided by the number of grid points was  $1.81 \text{ m}^2 \text{ s}^{-2}$  in DAS-1,  $0.88 \text{ m}^2 \text{ s}^{-2}$  in DAS-2,  $1.22 \text{ m}^2 \text{ s}^{-2}$  in DAS-3, and  $0.88 \text{ m}^2 \text{ s}^{-2}$  in DAS-4. The DAS-4 setup proved to be of benefit over DAS-3, but not over DAS-2, which indicates that strategies to optimally adjust the DAS to improve the model forecast must carefully consider the interaction between the DAS input parameters. The forecast sensitivity depends on the nominal values of the data assimilation input and any changes in various input components re-

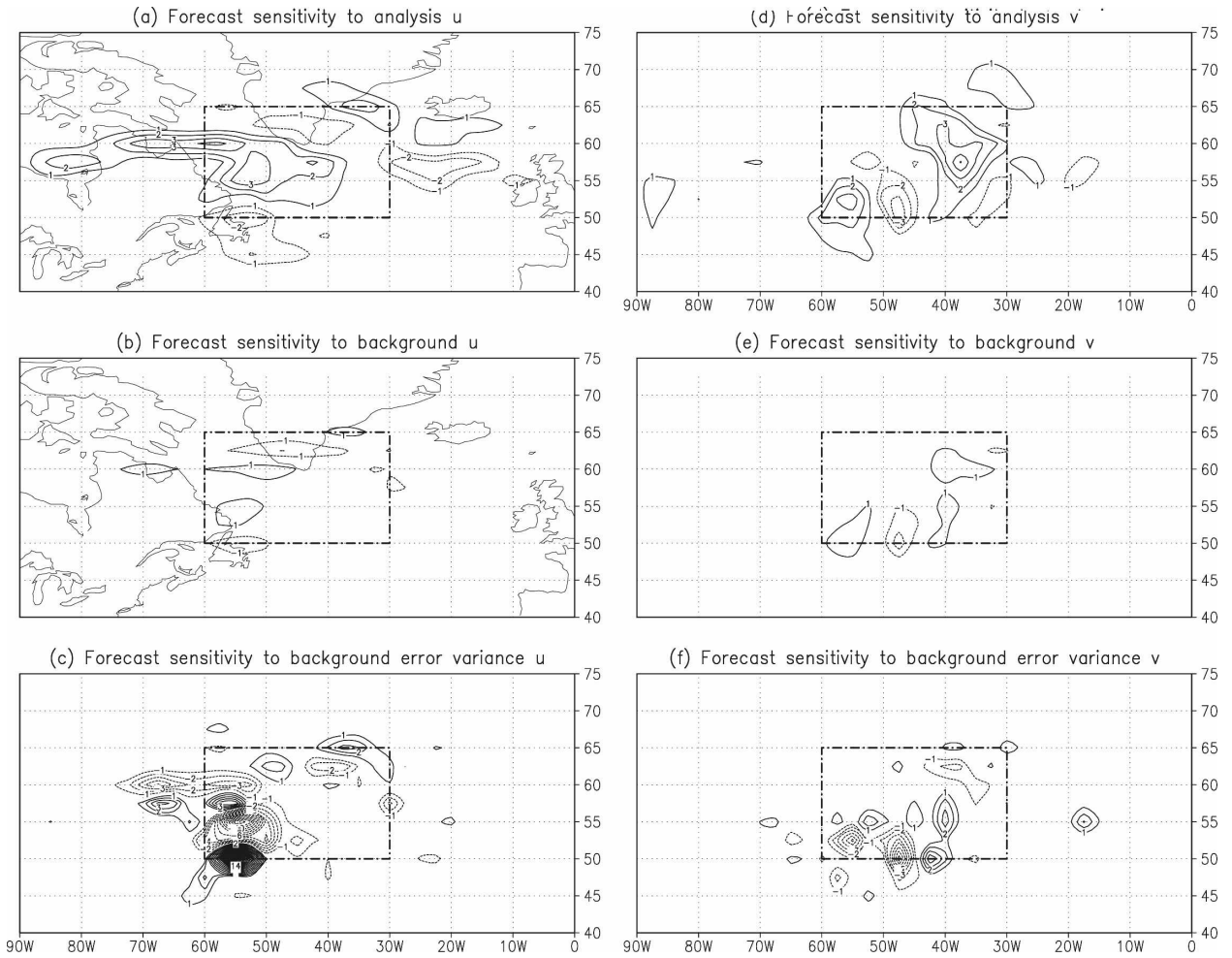


FIG. 8. Isopleths of the sensitivity of the forecast error to the analysis velocities (a)  $\nabla_{u_0^a} J^u$  and (d)  $\nabla_{v_0^a} J^v$ , background velocities (b)  $\nabla_{u_b} J^u$  and (e)  $\nabla_{v_b} J^v$ , and background error variance specification (c)  $\nabla_{\sigma_{b,u}^2} J^u$  and (f)  $\nabla_{\sigma_{b,v}^2} J^v$ . The contour interval is  $1 \text{ m s}^{-1}$  in (a), (b), (d), and (e) and 1 (nondimensional) in (c) and (f).

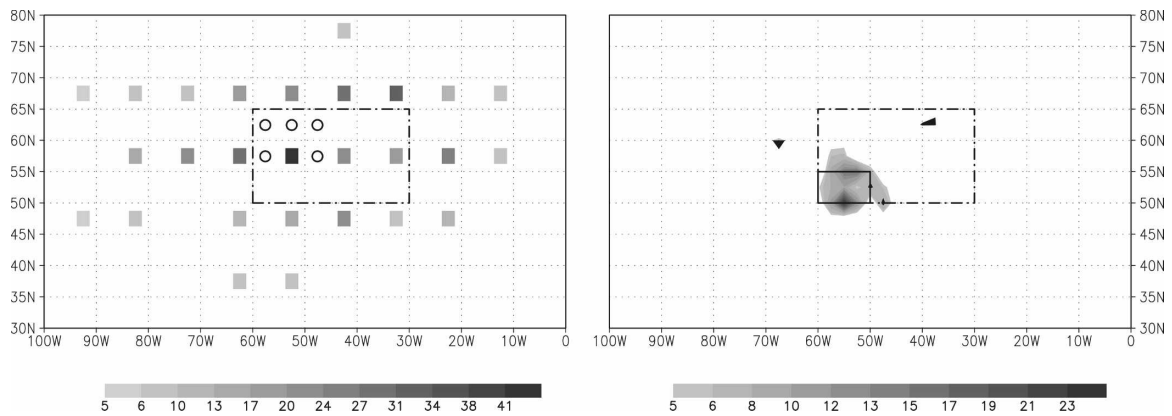


FIG. 9. (left) Time cumulative magnitude of the forecast sensitivity to velocity data  $|\nabla_{u_i^o} J^u| + |\nabla_{v_i^o} J^v| (\text{m s}^{-1})$  and location of additional observations in DAS-2 and DAS-4. (right) Magnitude of the forecast sensitivity to the specified background error variances  $|\nabla_{\sigma_{b,u}^2} J^u| + |\nabla_{\sigma_{b,v}^2} J^v|$  (nondimensional) and selection of the high-impact background error variance region in DAS-3 and DAS-4.

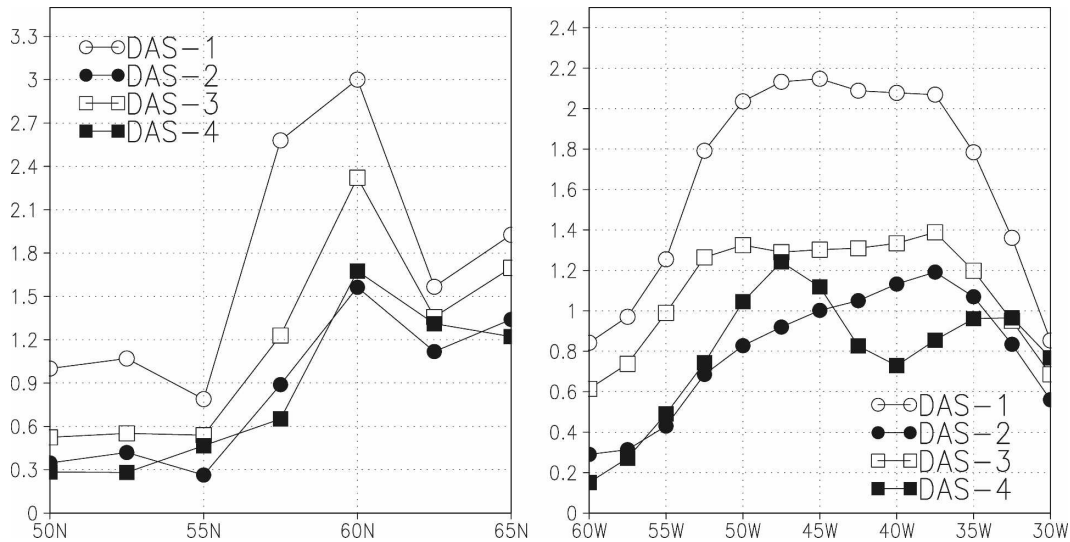


FIG. 10. Zonal- and meridional-averaged total energy forecast error ( $\text{m}^2 \text{s}^{-2}$ ) at the verification time over the verification domain for each of the 4D-Var experiments.

quire a periodic reevaluation of the sensitivity fields to account for the updated DAS configuration.

#### 4. Conclusions and further research

Observation sensitivity techniques have been initially considered for applications to targeted observations and are currently implemented in NWP to provide an effective global monitor of the short-range forecast performance of the observations. The adjoint DAS approach may be used to estimate the observation impact (or more general other-input parameters) on the analysis and subsequent forecasts. In this study the sensitivity equations of the nonlinear 4D-Var scheme were derived in the context of sensitivity to parameters in unconstrained minimization. The theoretical framework of the implicit function theorem allowed a compact derivation of the sensitivity equations with respect to the extended set of input parameters ( $\mathbf{y}_i$ ,  $\mathbf{R}_i$ ,  $\mathbf{x}_b$ ,  $\mathbf{B}$ ). Numerical experiments were performed to illustrate the ability of the adjoint DAS approach to provide forecast sensitivity, not only with respect to the background and time-distributed observational data but also with respect to the specification of the background and observation error statistics. Modeling the input error covariance matrices  $\mathbf{R}$  and  $\mathbf{B}$  is an area of intensive research in NWP and in practice only few of the error statistics may be properly estimated. The error covariance sensitivity analysis is particularly useful since information provided by the sensitivity fields may be used to identify the input components where improved statistical information would be of most benefit to the analysis

and forecast. Further applications to adaptive data thinning and targeted observations are envisaged and valuable insight may be gained through observing system simulation experiments. The 4D-Var framework allows a sensitivity analysis with respect to time-space-distributed data and it is well suited for applications that involve multiple observation targeting instants in the assimilation window (e.g., flight path design).

The 4D-Var sensitivity analysis involves a significant software development and several simplifying assumptions are required in NWP applications to reduce the computational burden. The simplicity of the SW model used in this study allowed the implementation of a second-order adjoint model associated with the nonlinear 4D-Var formulation and an increased numerical accuracy in the sensitivity computations. For practical purposes, inconsistencies between the theoretical formulation (such as the analytic sensitivity equations derived in this study) and the operational implementation of the data assimilation (involving several approximations) must be carefully considered. Consistent sensitivity estimates may be obtained by using the adjoint of the sequence of computations performed by the data assimilation procedure. Aside from the initial code development effort, a periodic software update is required as advanced data assimilation schemes and optimization algorithms are implemented.

*Acknowledgments.* This research was supported by the NASA Modeling, Analysis and Prediction Program under Award NNG06GC67G. We thank two anony-

mous reviewers whose thoughtful comments and suggestions helped to improve the manuscript.

## REFERENCES

- Akella, S., and I. M. Navon, 2006: A comparative study of the performance of high-resolution advection schemes in the context of data assimilation. *Int. J. Numer. Methods Fluids*, **51**, 719–748.
- Atlas, R., 1997: Atmospheric observations and experiments to assess their usefulness in data assimilation. *J. Meteor. Soc. Japan*, **75**, 111–130.
- Baker, N. L., 2000: Observation adjoint sensitivity and the adaptive observations targeting problem. Ph.D. thesis, Naval Postgraduate School, 238 pp.
- , 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., and A. Doerenbecher, 2002: A study on the optimization of the deployment of targeted observations using adjoint-based methods. *Quart. J. Roy. Meteor. Soc.*, **128**, 1689–1712.
- Bertsekas, D. P., 1995: *Nonlinear Programming*. Athena Scientific, 646 pp.
- Cacuci, D. G., 2003: *Theory*. Vol. 1, *Sensitivity and Uncertainty Analysis*, Chapman and Hall/CRC, 304 pp.
- Cardinali, C., S. Pezzulli, and E. Andersson, 2004: Influence-matrix diagnostic of a data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **130**, 2767–2786.
- Chapnik, B., G. Desroziers, F. Rabier, and O. Talagrand, 2006: Diagnosis and tuning of observational error in a quasi-operational data assimilation setting. *Quart. J. Roy. Meteor. Soc.*, **132**, 543–565.
- Cohn, S. E., 1997: An introduction to estimation theory. *J. Meteor. Soc. Japan*, **75**, 257–288.
- Courtier, P., J. N. Thépaut, and A. Hollingsworth, 1994: A strategy of operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1388.
- Daescu, D. N., and I. M. Navon, 2007: Efficiency of a POD-based reduced second-order adjoint model in 4D-Var data assimilation. *Int. J. Numer. Methods Fluids*, **53**, 985–1004.
- Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.
- Desroziers, G., and S. Ivanov, 2001: Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. *Quart. J. Roy. Meteor. Soc.*, **127**, 1433–1452.
- Doerenbecher, A., 2002: Étude de l'optimisation d'un système d'observation adaptatif pour améliorer la prévision des dépressions météorologiques. Ph.D. thesis, Paul Sabatier University, 264 pp.
- , and T. Bergot, 2001: Sensitivity to observations applied to FASTEX cases. *Nonlinear Proc. Geophys.*, **8**, 467–481.
- Errico, R. M., 2007: Interpretations of an adjoint-derived observational impact measure. *Tellus*, **59A**, 273–276.
- Etherton, B. J., and C. H. Bishop, 2004: Resilience of hybrid ensemble/3DVAR analysis schemes to model error and ensemble covariance error. *Mon. Wea. Rev.*, **132**, 1065–1080.
- Fisher, M., and P. Courtier, 1995: Estimating the covariance matrices of analysis and forecast error in variational data assimilation. ECMWF Tech. Memo. 220, 28 pp.
- Fourrié, N., A. Doerenbecher, T. Bergot, and A. Joly, 2002: Adjoint sensitivity of the forecast to TOVS observations. *Quart. J. Roy. Meteor. Soc.*, **128**, 2759–2777.
- Frehlich, R., 2006: Adaptive data assimilation including the effect of spatial variations in observation error. *Quart. J. Roy. Meteor. Soc.*, **132**, 1225–1257.
- Giering, R., and T. Kaminski, 1998: Recipes for adjoint code construction. *ACM Trans. Math. Software*, **24**, 437–474.
- Hamill, T. M., and C. Snyder, 2002: Using improved background-error covariances from an ensemble Kalman filter for adaptive observations. *Mon. Wea. Rev.*, **130**, 1552–1572.
- Horn, R. A., and C. R. Johnson, 1991: *Topics in Matrix Analysis*. Cambridge University Press, 607 pp.
- Janjić, T., and S. E. Cohn, 2006: Treatment of observation error due to unresolved scales in atmospheric data assimilation. *Mon. Wea. Rev.*, **134**, 2900–2915.
- Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory*. Academic Press, 376 pp.
- Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge University Press, 364 pp.
- Kucukkaraca, E., and M. Fisher, 2006: Use of analysis ensembles in estimating flow-dependent background error variances. ECMWF Tech. Memo. 492, 16 pp.
- Langland, R. H., 2005a: Issues in targeted observing. *Quart. J. Roy. Meteor. Soc.*, **131**, 3409–3425.
- , 2005b: Observation impact during the North Atlantic TReC-2003. *Mon. Wea. Rev.*, **133**, 2297–2309.
- , and N. L. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56A**, 189–201.
- Lawless, A. S., S. Gratton, and N. K. Nichols, 2005: An investigation of incremental 4D-Var using non-tangent linear models. *Quart. J. Roy. Meteor. Soc.*, **131**, 459–476.
- Le Dimet, F. X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A**, 97–110.
- , H.-E. Ngodock, B. Luong, and J. Verron, 1997: Sensitivity analysis in variational data assimilation. *J. Meteor. Soc. Japan*, **75**, 245–255.
- , I. M. Navon, and D. N. Daescu, 2002: Second-order information in data assimilation. *Mon. Wea. Rev.*, **130**, 629–648.
- Lewis, J., S. Lakshminarayanan, and S. Dhall, 2006: *Dynamic Data Assimilation: A Least Squares Approach*. Cambridge University Press, 654 pp.
- Lin, S.-J., and R. B. Rood, 1997: An explicit flux-form semi-Lagrangian shallow-water model on the sphere. *Quart. J. Roy. Meteor. Soc.*, **123**, 2477–2498.
- Liu, D. C., and J. Nocedal, 1989: On the limited memory BFGS method for large-scale optimization. *Math. Program.*, **45**, 503–528.
- Lorenc, A. C., 2003: Modelling of error covariances by 4D-Var data assimilation. *Quart. J. Roy. Meteor. Soc.*, **129**, 3167–3182.
- Magnus, J. R., and H. Neudecker, 1999: *Matrix Differential Calculus with Applications in Statistics and Econometrics*. rev. ed. Wiley, 395 pp.
- Marchuk, G. I., V. I. Agoshkov, and V. P. Shutyaev, 1996: *Adjoint Equations and Perturbation Algorithms in Nonlinear Problems*. CRC Press, 275 pp.
- Morales, J. L., and J. Nocedal, 2000: Automatic preconditioning by limited memory quasi-Newton updating. *SIAM J. Optim.*, **10** (4), 1079–1096.
- Neudecker, H., 1969: Some theorems on matrix differentiation with special reference to Kronecker matrix products. *J. Amer. Stat. Assoc.*, **64**, 953–963.

- Parrish, D. F., and J. C. Derber, 1992: National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.
- Rabier, F., 2005: Overview of global data assimilation developments in numerical weather-prediction centres. *Quart. J. Roy. Meteor. Soc.*, **131**, 3215–3233.
- Trémolet, Y., 2004: Diagnostics of linear and incremental approximations in 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **130**, 2233–2251.
- , 2005: Incremental 4D-Var convergence study. ECMWF Tech. Memo. 469, 34 pp.
- , 2007: Model-error estimation in 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **133**, 1267–1280.
- Veersé, F., D. Auroux, and M. Fisher, 2000: Limited-memory BFGS diagonal preconditioners for a data assimilation problem in meteorology. *Optim. Eng.*, **1**, 323–339.
- Wu, W., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916.
- Zhu, Y., and R. Gelaro, 2008: Observation sensitivity calculations using the adjoint of the Gridpoint Statistical Interpolation (GSI) analysis system. *Mon. Wea. Rev.*, **136**, 335–351.