INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2012; **69**:110–123
Published online 2 March 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/fld.2545

Effect of random perturbations on adaptive observation techniques

M. J. Hossen[1], I. M. Navon[2,*,†] and D. N. Daescu[3]

[1]*Department of Mathematics and Natural Sciences, BRAC University, 66 Mohakhali, Dhaka-1212, Bangladesh*
[2]*Department of Scientific Computing, Florida State University, Tallahassee, FL 32306-4120, U.S.A.*
[3]*Fariborz Maseeh Department of Mathematics and Statistics, Portland State University, P.O. Box 751, Portland, OR 97207, U.S.A.*

SUMMARY

An observation sensitivity (OS) method to identify targeted observations is implemented in the context of four-dimensional variational (4D-Var) data assimilation. This methodology is compared with the well-established adjoint sensitivity (AS) method using a nonlinear Burgers equation as a test model. Automatic differentiation software is used to implement the first-order adjoint model (ADM) to calculate the gradient of the cost function required in the 4D -Var minimization algorithm and in the AS computations and the second-order ADM to obtain information on the Hessian matrix of the 4D-Var cost that is necessary in the OS computations. Numerical results indicate that the observation-targeting is particularly successful in reducing the forecast error for moderate Reynolds numbers. The potential benefits of the OS targeting approach over the AS are investigated. The effect of random perturbations on the performance of these adaptive observation techniques is also analyzed. Copyright © 2011 John Wiley & Sons, Ltd.

Received 15 February 2010; Revised 19 October 2010; Accepted 31 December 2010

KEY WORDS: forecast error; Hessian matrix; automatic differentiation; adjoint sensitivity; observation sensitivity; adaptive observations

1. INTRODUCTION

Variational data assimilation (VDA) methods combine the information provided by the forecast model equations and physical information given by the observations in a window of assimilation in order to retrieve an accurate flow. It is well established that initial condition errors have a major impact on the forecasts and sensitivity analysis techniques are used to provide an assessment of the relative importance and contribution of various model parameters and data assimilation system (DAS) components in reducing the forecast errors. A theoretical framework to sensitivity analysis in VDA was first presented by Le Dimet *et al.* [1] in the context of optimal control theory.

Recently, the four-dimensional VDA (4D-Var) method was implemented at various numerical weather prediction centers to improve the model forecast skill. The advantage of this method is that it can use time-distributed observations to define a cost function reconciling the forecast with observations. Besides the existing observational network, some additional observations placed at critical time and space locations can improve the model forecast. Adaptive (targeted) observation methods aim to identify optimal locations where a small number of additional observational resources must be deployed to improve a specific forecast aspect, see Langland [2].

*Correspondence to: I. M. Navon, Department of Scientific Computing, Florida State University, Tallahassee, FL 32306-4120, U.S.A.
†E-mail: inavon@fsu.edu

The design of cost-effective observation-targeting strategies relies on the ability to *a priori* identify optimal sites for collecting data of large impact on reducing forecast errors. Several techniques for adaptive observations have already been proposed and tested in practical applications, for instance, the adjoint sensitivity (AS) method of Langland *et al.* [3], singular vector methods including the total energy metric and the Hessian metric as discussed by Palmer *et al.* [4], Gelaro *et al.* [5], Daescu and Navon [6], and the sensitivity to observations method by Baker and Daley [7] and Daescu [8]. In the AS approach, the adjoint model (ADM) is used to evaluate the gradient with respect to the initial conditions of a scalar measure $\mathscr{J}^{\mathrm{v}}$ of the forecast of interest that is defined over a subregion (verification domain) of the global domain. The singular vector method is used to identify the error structure in the analysis field of the largest forecast impact. The singular vectors of the forward tangent linear propagator of a nonlinear dynamical system provide a mathematically rigorous and tractable approach to quantifying the perturbation growth over a finite time interval. The dominant singular vectors describe the most rapidly growing structures with respect to a given metric (norm) over this interval in the tangent linear model (TLM) space.

Baker and Daley [7] have shown that the adjoint of a DAS provides an efficient way to estimate the forecast sensitivity with respect to observations. The sensitivities may be computed with respect to any or all of the observations simultaneously based on a single execution of the adjoint DAS. The observation sensitivity (OS) method has also been used to assess the effectiveness of targeted observations by Doerenbecher and Bergot [9] and the impact of satellite data by Fourrié *et al.* [10]. OS estimation was initially considered in the 3D-Var context by Baker and Daley [7], Langland and Baker [11], Zhu and Gelaro [12] and extended to the 4D-Var context by Daescu [8] and Trémolet [13].

The key to the OS method is the ability to compute the transpose of the gain matrix that determines the weights given to the observation-minus-background residuals either explicitly or through a sequence of available operators. Daescu [8] shows that even without knowing the full gain matrix, it is possible to compute the OS vector if the Hessian matrix of the cost function with respect to the control variable is available in the operator format (matrix–vector products). This information may be provided by implementing a second-order ADM, as explained by Le Dimet *et al.* [14].

The plan of this paper is as follows.

In Section 2, we provide a brief description of the theory related to the AS and OS methods. Section 3 details the model used (Burgers'equation) and the experimental setup and provides details of the algorithm used for calculating the OS vector. Numerical tests are carried out in Section 4 and their outcomes are assessed and discussed. Section 5 provides an analysis of the impact of random perturbations (used to generate the background vector and observations) on the performance of the AS method and of the OS method. Summary and conclusions are in Section 6.

2. ADJOINT-BASED OBSERVATION-TARGETING METHODS

In 4D-Var data assimilation, an initial condition is sought so that the forecast best fits the observations within an assimilation time interval. 4D-Var provides an optimal estimate $x_0^{\mathrm{a}} \in \mathbb{R}^n$ to the initial condition of a nonlinear forecast model by minimizing the cost function defined as

$$\mathscr{J}(x_0) = \underbrace{\frac{1}{2}(x_0 - x_{\mathrm{b}})^{\mathrm{T}}\mathbf{B}^{-1}(x_0 - x_{\mathrm{b}})}_{\mathscr{J}_{\mathrm{b}}} + \underbrace{\frac{1}{2}\sum_{i=0}^{N}(y_i - H_i(x_i))^{\mathrm{T}}\mathbf{R}_i^{-1}(y_i - H_i(x_i))}_{\mathscr{J}_{\mathrm{o}}} \quad (1)$$

where $x_0 = x(t_0)$ denotes the initial state at the initial time t_0, x_{b} is a prior (background) estimate to the initial state, $y_i \in \mathbb{R}^{k_i}$, $i = 0, 1, 2, \ldots, N$, is the set of observations available at time t_i, $x_i = \mathscr{M}(t_0, t_i)(x_0)$ is the nonlinear model forecast state at t_i, and $H_i : \mathbb{R}^n \to \mathbb{R}^{k_i}$ is the observation operator that maps the state space into the observation space at t_i. $\mathbf{B}$ is the background error covariance matrix and $\mathbf{R}_i$ is the observational error covariance matrix at time t_i. In practice, the dimensionality of the state vector is in the range of $n \sim 10^7$–10^8 and simplifying assumptions are necessary to

DOI: 10.1002/fld

reduce the computational burden associated with the minimization problem. The model $\mathscr{M}$ is assumed to be perfect and by imposing the model equations as the strong constraint, the control variable in the minimization of the cost functional (1) is the initial state of the model x_0.

The adjoint method provides an efficient approach to evaluate the gradient of the cost functional with respect to the control variables, $\nabla_{x_0}\mathscr{J}$, that is necessary in the minimization process. The gradient is expressed as

$$\begin{aligned}\nabla_{x_0}\mathscr{J} &= \mathbf{B}^{-1}(x_0 - x_\mathrm{b}) - \sum_{i=0}^{N} \mathbf{M}^\mathrm{T}(t_0, t_i)\mathbf{H}_i^\mathrm{T}\mathbf{R}_i^{-1}(y_i - H_i[x_i]) \\ &= \mathbf{B}^{-1}(x_0 - x_\mathrm{b}) - \sum_{i=0}^{N} \mathbf{M}_{0,i}^\mathrm{T}\mathbf{H}_i^\mathrm{T}\mathbf{R}_i^{-1}(y_i - H_i[x_i]) \end{aligned} \tag{2}$$

where $\mathbf{H}_i$ is the linearized observational operator, $\mathbf{M}_{0,i} = \mathbf{M}(t_0, t_i)$ is the TLM and its transpose $\mathbf{M}_{0,i}^\mathrm{T}$ is the ADM. ADM is computationally obtained by transposing the TLM, integrated backward in time.

Each iteration of the 4D-Var minimization requires the evaluation of the gradient (2) i.e. computing the increment $y_i - H_i[x_i]$ at each observation time t_i during a forward integration, multiplying it by $\mathbf{H}_i^\mathrm{T}\mathbf{R}_i^{-1}$ and integrating the weighted increments back to the initial time using the ADM. Therefore, the implementation of the 4D-Var algorithm involves a sequence of calls to an unconstrained minimization algorithm that uses the functional and gradient information to provide an optimal initial condition (analysis) x_0^a. In practice, the implementation of the first- and second-order ADMs associated with nonlinear dynamics is non-trivial and automatic differentiation software packages that provide forward and reverse sweeps are used to facilitate the code development see e.g. Giering and Kaminski [15].

2.1. *The AS approach*

A first approach to identify the adaptive observations is the AS method. In practice, it is of interest to assess the observation impact on a forecast aspect $\mathscr{J}^\mathrm{v}(x_\mathrm{v})$ defined on the verification domain $\mathscr{D}_\mathrm{v}$ at the verification time t_v. The functional $\mathscr{J}^\mathrm{v}$ is typically defined as a scalar measure of the forecast error over $\mathscr{D}_\mathrm{v}$

$$\mathscr{J}^\mathrm{v} = \tfrac{1}{2}(x_\mathrm{v}^\mathrm{f} - x_\mathrm{v}^\mathrm{t})^T P^\mathrm{T} E P (x_\mathrm{v}^\mathrm{f} - x_\mathrm{v}^\mathrm{t}) \tag{3}$$

where x_v^f is the model forecast at the verification time initialized from the analysis x_0^a and x_v^t is a verification state at t_v that serves as a proxy to the true atmospheric state. P is the state projection operator on $\mathscr{D}_\mathrm{v}$ satisfying $P^*P = P^2 = P$ and the metric E is typically prescribed as a diagonal matrix associated with the total energy norm.

2.1.1. Location of adaptive observations by AS. The AS approach identifies targeted observations based on the information extracted from the gradient of the functional (3). The sensitivity of the forecast aspect with respect to the model state at each targeting instant t_i is

$$\nabla\mathscr{J}^\mathrm{v}(x_i) = \mathbf{M}_{i,v}^\mathrm{T} P^\mathrm{T} E P (x_\mathrm{v}^\mathrm{f} - x_\mathrm{v}^\mathrm{t}) \tag{4}$$

where $x_i = x(t_i)$. A sensitivity value of large magnitude indicates that small variations in the model state x_i will have a significant impact on the forecast at the verification time. The AS field at each targeting instant t_i is defined as

$$F_\mathrm{v}(t_i) = \|\mathrm{grad}\ \mathscr{J}^\mathrm{v}(x(t_i))\|_E \in \mathbb{R}^n \tag{5}$$

The E-norm is defined as

$$\|x\|_E^2 = x^\mathrm{T} E x \tag{6}$$

DOI: 10.1002/fld

and in the numerical experiments, E is taken as the Euclidean norm. We prescribe five targeting instants τ_k and at each τ_k, a number of $n_k=5$ adaptive observations are selected at the locations where the sensitivity field $F_{\mathrm{v}}(\tau_k)$ attains its largest values.

2.2. *The OS approach*

The theory presented in Daescu [8] shows that if the cost function $\mathscr{J}(x,u)$ is a twice continuously differentiable function involving parameter vector (input data) $u \in \mathbb{R}^p$, then the optimal solution $\bar{x}$ corresponding to the parameter $\bar{u}$ that minimizes $\mathscr{J}$ is obtained by satisfying the condition $\nabla_x \mathscr{J}(\bar{x},\bar{u})=0$. Assuming that the Hessian matrix $\nabla^2_{xx}\mathscr{J}(\bar{x},\bar{u})$ is positive definite, the implicit function theorem applied to the first-order optimality condition

$$\nabla_x \mathscr{J}(\bar{x},\bar{u})=0 \in \mathbb{R}^n$$

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

$$\nabla_u x = (\nabla_u x_1, \nabla_u x_2, \ldots, \nabla_u x_n) \in \mathbb{R}^{p\times n} \tag{7}$$

is expressed as

$$\nabla_u x(u) = -\nabla^2_{ux}\mathscr{J}[x(u),u]\{\nabla^2_{xx}\mathscr{J}[x(u),u]\}^{-1} \tag{8}$$

Rabier and Courtier [16] show that the inverse of the Hessian matrix $\nabla^2_{x_0x_0}\mathscr{J}$ is an approximation of the error covariance matrix $\mathbf{A}$ associated with the optimal analysis x_0^{a}. For notational convenience, we denote

$$\mathbf{A} = [\nabla^2_{x_0x_0}\mathscr{J}(x_0^{\mathrm{a}})]^{-1} \in \mathbb{R}^{n\times n} \tag{9}$$

By differentiating (2) with respect to y_i, we obtain

$$\nabla^2_{y_i x_0}\mathscr{J}(x_0^{\mathrm{a}}) = -\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{M}_{0,i} \tag{10}$$

and from Equations (8), (9), and (10) the analysis sensitivity to observations is expressed as

$$\begin{aligned}\nabla_{y_i} x_0^{\mathrm{a}} &= -[\nabla^2_{y_i x_0}\mathscr{J}(x_0^{\mathrm{a}})]\mathbf{A}\\ &= \mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{M}_{0,i}\mathbf{A} \in \mathbb{R}^{k_i\times n}\end{aligned} \tag{11}$$

The gradient of $\mathscr{J}^{\mathrm{v}}$ defined in (3) with respect to the forecast at verification time t_{v} is

$$\nabla_{x_{\mathrm{v}}}\mathscr{J}^{\mathrm{v}} = P^{\mathrm{T}} E P (x_{\mathrm{v}}^{\mathrm{f}} - x_{\mathrm{v}}^{\mathrm{t}}) \tag{12}$$

and using the chain rule we obtain

$$\nabla_{y_i}\mathscr{J}^{\mathrm{v}}(x_0^{\mathrm{a}}) = \nabla_{y_i} x_0^{\mathrm{a}} \nabla_{x_0}\mathscr{J}^{\mathrm{v}}(x_0^{\mathrm{a}}) \in \mathbb{R}^{k_i} \tag{13}$$

where the gradient $\nabla_{x_0}\mathscr{J}^{\mathrm{v}}(x_0^{\mathrm{a}})$ is the forecast sensitivity to the analysis at the initial time and it is obtained by integrating the ADM along the forecast model trajectory initialized from x_0^{a}

$$\nabla_{x_0}\mathscr{J}^{\mathrm{v}}(x_0^{\mathrm{a}}) = \mathbf{M}_{0,\mathrm{v}}^{\mathrm{T}} \nabla_{x_{\mathrm{v}}}\mathscr{J}^{\mathrm{v}} \in \mathbb{R}^n \tag{14}$$

By using (11), Equation (13) can be written as

$$\nabla_{y_i}\mathscr{J}^{\mathrm{v}}(x_0^{\mathrm{a}}) = \mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{M}_{0,i}\mathbf{A}\nabla_{x_0}\mathscr{J}^{\mathrm{v}}(x_0^{\mathrm{a}}) \in \mathbb{R}^{k_i} \tag{15}$$

which provides the forecast sensitivity to the observations.

DOI: 10.1002/fld

2.2.1. Location of adaptive observations by OS. The 4D-Var method allows an optimal use of time–space-distributed observational data. In the context of adaptive observations, we need to consider additional observations along with routine observations. To identify the location of adaptive observations that have a large impact on the forecast, we evaluate the OS with respect to adaptive observations by using Equation (15), but we use the analysis error covariance matrix **A** associated with routine observations only. For details, see Doerenbecher and Bergot [9]. The OS vectors are used to determine the location of targeted observations. By analogy to the AS setup, we specify five targeting instants τ_k and at each τ_k, a number of $n_k = 5$ adaptive observations are selected at the locations where the sensitivity field $|\nabla_{y_{\mathrm{a}}} \mathscr{J}^{\mathrm{v}}(\tau_k)|$ attains its maximum values, where the subscript a stands for adaptive observations.

3. THE DESIGN OF THE NUMERICAL EXPERIMENTS

The 4D-Var method requires a significant computational effort to evaluate the forecast sensitivity to observations when a large-scale model is used. Simple prototype models allow to provide a proof-of-concept and are used as a first step in testing new methodologies, see e.g. Xu *et al.* [17]. A nonlinear Burgers equation model is described below. The algorithm to compute the OS consists of:

- A model integration of a reference ('true') initial condition state x_0^{t} to obtain the verification state $x_{\mathrm{v}}^{\mathrm{t}}$ at t_{v}

$$x_{\mathrm{v}}^{\mathrm{t}} = \mathscr{M}_{0,\mathrm{v}}(x_0^{\mathrm{t}}) \tag{16}$$

- Obtain optimal initial condition x_0^{a} by minimizing the cost functional $\mathscr{J}$ defined in (1). Calculate model forecast

$$x_{\mathrm{v}}^{\mathrm{f}} = \mathscr{M}_{0,\mathrm{v}}(x_0^{\mathrm{a}}) \tag{17}$$

- Compute $\nabla_{x_{\mathrm{v}}} \mathscr{J}^{\mathrm{v}} = P^{\mathrm{T}} E P(x_{\mathrm{v}}^{\mathrm{f}} - x_{\mathrm{v}}^{\mathrm{t}})$ and use it to initialize the ADM.
- Integration of the ADM backward from t_{v} to t_0: $\nabla_{x_0} \mathscr{J}^{\mathrm{v}}(x_0^{\mathrm{a}}) = \mathbf{M}_{0,\mathrm{v}}^{\mathrm{T}} \nabla_{x_{\mathrm{v}}} \mathscr{J}^{\mathrm{v}}$.
- Solve the linear system for z_0:$\mathbf{A}^{-1} z_0 = \nabla_{x_0} \mathscr{J}^{\mathrm{v}}(x_0^{\mathrm{a}})$ using an iterative procedure with the Hessian matrix–vector products provided by the second-order ADM.
- Integrate the TLM from t_0 to t_{a} (observation-targeting time): $z_{\mathrm{a}} = \mathbf{M}_{0,\mathrm{a}} z_0$.
- Mapping on observation space, weighting: $\nabla_{y_{\mathrm{a}}} \mathscr{J}^{\mathrm{v}} = \mathbf{R}_{\mathrm{a}}^{-1} \mathbf{H}_{\mathrm{a}} z_{\mathrm{a}}$.

3.1. Experimental setup

Illustrative numerical experiments are set up with the one-dimensional nonlinear Burgers equation

$$\frac{\partial x}{\partial t} + x \frac{\partial x}{\partial \xi} = \frac{1}{Re} \frac{\partial^2 x}{\partial \xi^2} \tag{18}$$

where Re denotes the Reynolds number. The initial condition is taken as

$$x(\xi, 0) = \begin{cases} 1.0, & -2 \leqslant \xi \leqslant 0 \\ 0.0, & 0 < \xi \leqslant 2 \end{cases}$$

and boundary conditions are specified as

$$x(-2.0, t) = 1.0 \quad \text{and} \quad x(2.0, t) = 0.0$$

To discretize Equation (18), we use the forward in time and centered in space (FTCS) finite difference method with a time step $\Delta t = 0.01$. The numerical grid comprises 101 grid points in

DOI: 10.1002/fld

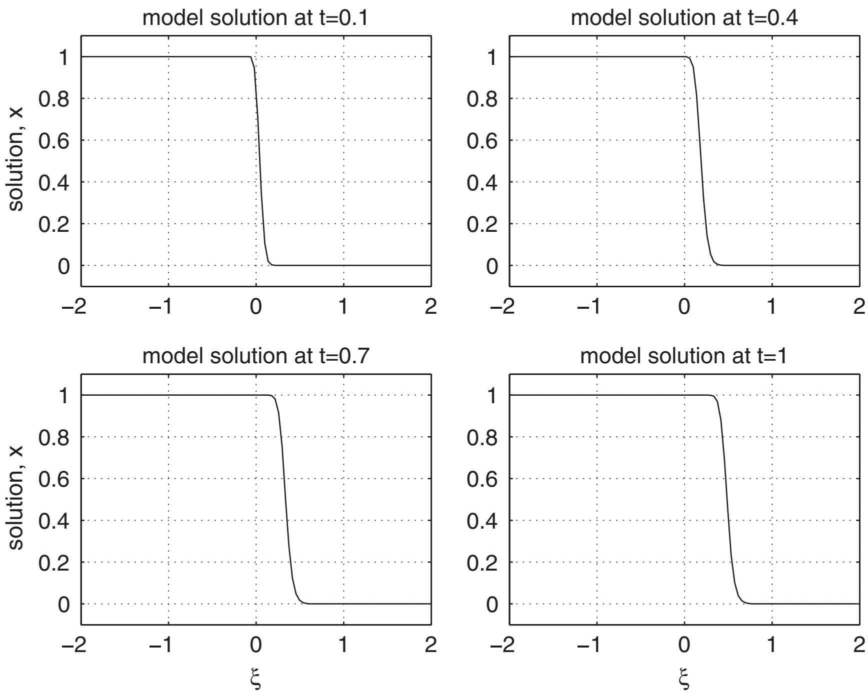


Figure 1. Burgers model solutions for Reynolds number $Re=50$ at different time steps.

space and 300 steps in time, such that $t_v=3$, and a Reynolds number $Re=50$ is used in all experiments. The equations of the discrete model are

$$x_j^{k+1}=x_j^k-\frac{\Delta t}{4\Delta\xi}\{(x_{j+1}^k)^2-(x_{j-1}^k)^2\}+\frac{1}{Re}\frac{\Delta t}{\Delta\xi^2}\{x_{j+1}^k-2x_j^k+x_{j-1}^k\} \tag{19}$$

where $\Delta\xi$ is the space grid increment. Figure 1 displays the model state at various time instants. In Equation (1), we see that the cost function requires background information as well as observations. A background state x_b and the observation vector y are generated from the model solution initialized with x_0^t by introducing random perturbations taken from normal distributions $N(0,\sigma_b^2)$ and $N(0,\sigma_o^2)$, respectively, and the standard deviations specified as $\sigma_b=0.05$ and $\sigma_o=0.04$. The observational errors are assumed to be uncorrelated, such that the observation error covariance matrices $\mathbf{R}_i$ are diagonal and time invariant with the values of the diagonal entries is set to σ_o^2. The routine observational data is assumed to be available at every five grid points in the space dimension and scattered in time at an increment of 20 time steps between $t_0=1$ and $t_N=100$, which is our 4D-Var data assimilation window. For simplicity, we assume that the background errors are uncorrelated and as a result the background error covariance is also a diagonal matrix with the diagonal entries set to σ_b^2.

The L-BFGS (limited memory Quasi-Newton) minimization algorithm [18] is used to solve the nonlinear 4D-Var minimization problem with a convergence criterion $\|\nabla_{x_0}\mathscr{J}\|\leqslant 10^{-5}$, where $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^n$. The reference solution x_v^t at t_v and the model forecast x_v^f from an optimal initial condition obtained by the assimilation of only routine observations are presented in Figure 2.

4. NUMERICAL RESULTS

Numerical experiments are performed to provide a comparative analysis between the AS and the OS methods to observation-targeting and to assess their impact on the forecast error reduction on the verification domain $\mathscr{D}_v$ at verification time t_v. In our experiments, we select the verification

DOI: 10.1002/fld

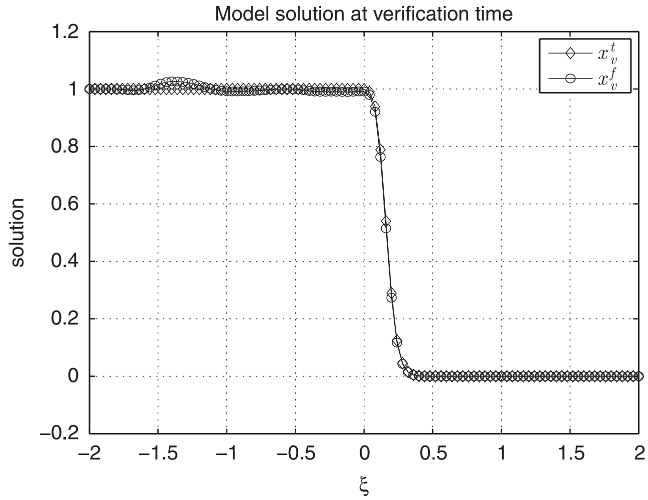


Figure 2. The reference state $x_{\rm v}^{\rm t}=\mathscr{M}_{0,\rm v}(x_0^{\rm t})$ and the model forecast state $x_{\rm v}^{\rm f}=\mathscr{M}_{0,\rm v}(x_0^{\rm a})$ at the verification time $t_{\rm v}$, where $x_0^{\rm a}$ denotes the optimal initial condition obtained from data assimilation of routine observations only.

domain as the spatial region $[-1.52, 0.20]$ where an increased forecast error was noticed. The adjoint method requires only the evaluation of the gradient of the functional $\mathscr{J}_{\rm v}$ that is defined in (4). On the other hand, the evaluation of the OS is computationally expensive. It requires the availability of the Hessian matrix of the 4D-Var cost function (1) in operator format and thus the development of the second-order ADM. The correctness of the first-order adjoint (FOA) and second-order adjoint (SOA) model implementation may be validated using standard procedures such as a gradient test, see Navon *et al.* [19]. A Taylor expansion in the direction of $\nabla\mathscr{J}$ leads to

$$\mathscr{J}(x+\alpha\nabla\mathscr{J})=\mathscr{J}(x)+\alpha\nabla\mathscr{J}^{\rm T}\nabla\mathscr{J}+\tfrac{1}{2}\alpha^2\nabla\mathscr{J}^{\rm T}[\nabla^2\mathscr{J}]\nabla\mathscr{J}+O(\alpha^3) \tag{20}$$

and allows to verify the FOA and SOA implementation according to the criteria

$$\phi(\alpha)=\frac{\mathscr{J}(x+\alpha\nabla\mathscr{J})-\mathscr{J}(x)}{\alpha\nabla\mathscr{J}^{\rm T}\nabla\mathscr{J}}=1+O(\alpha)$$

and respectively,

$$\psi(\alpha)=\frac{\mathscr{J}(x+\alpha\nabla\mathscr{J})-\mathscr{J}(x)-\alpha\nabla\mathscr{J}^{\rm T}\nabla\mathscr{J}}{\frac{1}{2}\alpha^2\nabla\mathscr{J}^{\rm T}[\nabla^2\mathscr{J}]\nabla\mathscr{J}}=1+O(\alpha)$$

The graphs of the FOA and SOA verification procedures are displayed in Figure 3. The results show that $\phi(\alpha)\approx 1$ as well as $\psi(\alpha)\approx 1$ with the increasing accuracy until the round-off error becomes dominant at the level of machine accuracy.

4.1. Observation-targeting using the AS

To evaluate the adjoint field (4), the verification state $x_{\rm v}^{\rm t}$ is calculated at $t_{\rm v}$ using the initial condition $x_0^{\rm t}$, $x_{\rm v}^{\rm t}=\mathscr{M}_{0,\rm v}(x_0^{\rm t})$, and the forecast state $x_{\rm v}^{\rm f}$ is taken as $x_{\rm v}^{\rm f}=\mathscr{M}_{0,\rm v}(x_0^{\rm a})$ with an optimal initial condition $x_0^{\rm a}$ obtained by minimizing the cost function (1) where only routine observations are assimilated. The vector $x_{\rm v}^{\rm f}-x_{\rm v}^{\rm t}$ is then used to initialize the backward integration of the ADM that provides the gradient (4) at every $t_i\in[t_0,t_{\rm f}]$. We obtain the AS vector and the optimal space locations for adaptive observations according to the method discussed in Section 2.1.1. The AS vector at various time instants is displayed in Figure 4. We choose five adaptive observations at each targeting instant τ_k and the locations of adaptive observations are shown in Figure 5.

DOI: 10.1002/fld

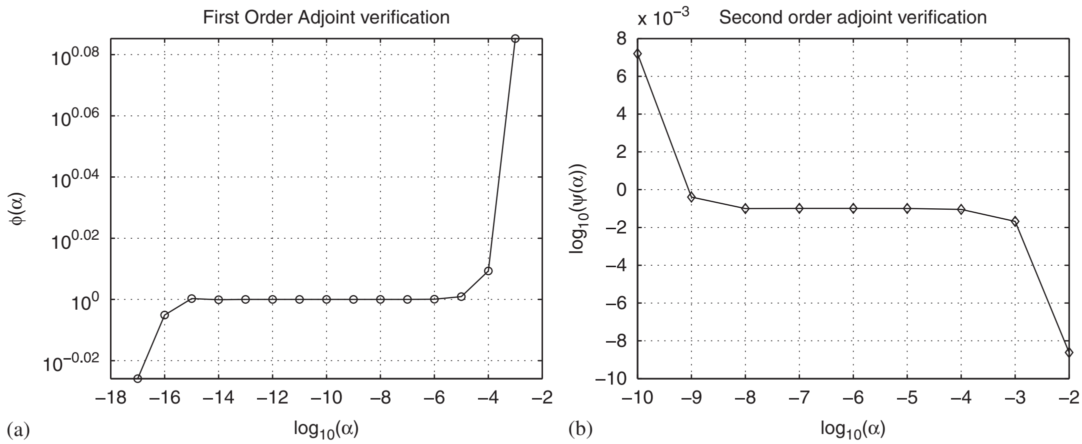


Figure 3. Verification of the correctness of the first-order adjoint model (a) and of the second-order adjoint model (b).

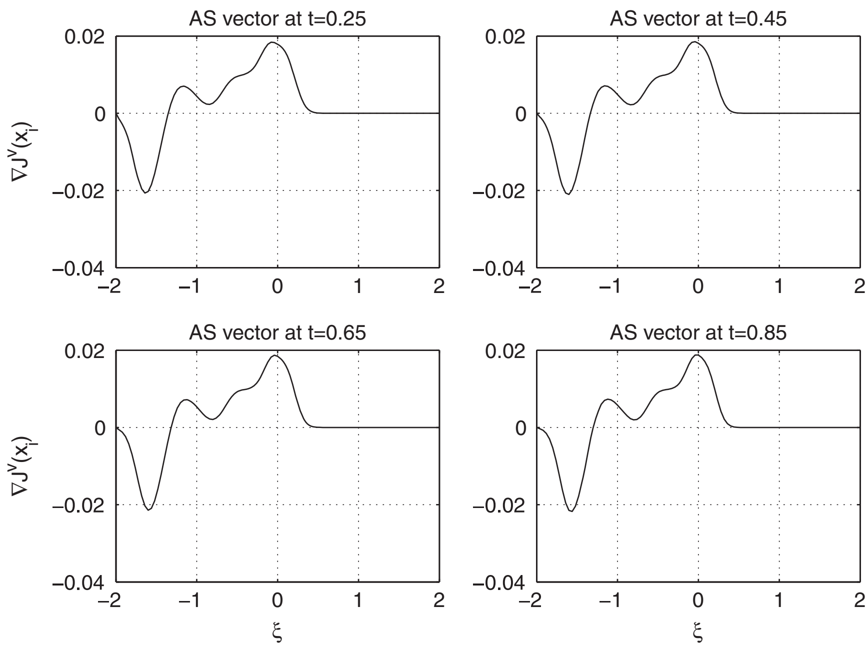


Figure 4. The spatial profile of the adjoint sensitivity vector at various time instants.

4.2. *Observation-targeting using the OS*

The OS vectors are evaluated using the algorithm presented in Section 3. At the first stage, we evaluate $\nabla_{x_v}\mathscr{J}^{\mathrm{v}}$ according to (12) and use it to initialize the ADM. The FOA model is integrated backward to obtain the gradient $\nabla_{x_0}\mathscr{J}^{\mathrm{v}}(x_0^{\mathrm{a}})$ evaluated at the analysis x_0^{a}. The linear system $\mathbf{A}^{-1}z_0 = \nabla_{x_0}\mathscr{J}^{\mathrm{v}}(x_0^{\mathrm{a}})$ is then solved where $\mathbf{A}^{-1}$ is the Hessian matrix of the cost function. The simplicity of our model and the use of a second-order ADM allowed us to obtain the full Hessian matrix, see Wang *et al.* [20], Le Dimet *et al.* [14]. We verified that this matrix is symmetric and we found that all its eigenvalues are positive. Therefore, the Hessian matrix is positive definite and a Cholesky decomposition method was used to solve the linear system. In our experimental setup, all the observational operators are linear and the observation error covariance matrices $\mathbf{R}_i$ are diagonal such that each of $\mathbf{R}_i^{-1}$ is also a diagonal matrix with diagonal entries σ_{o}^{-2}. OS vectors are displayed in Figure 6 and the adaptive observations are selected as discussed in Section 2.2.1. The locations of the adaptive observations are displayed in Figure 7.

DOI: 10.1002/fld

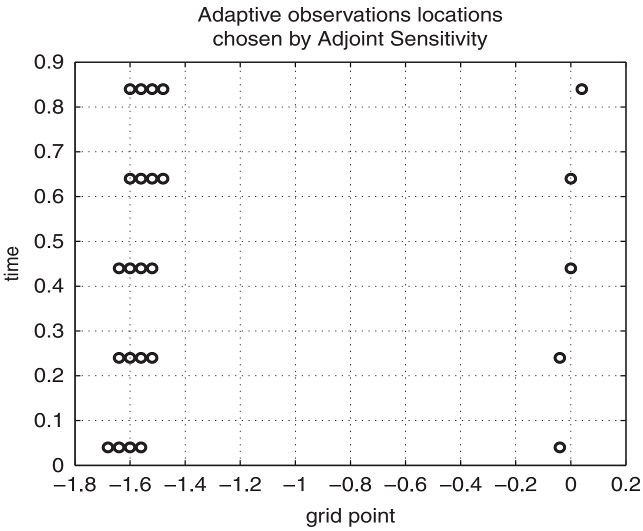


Figure 5. Locations of adaptive observations based on the adjoint sensitivity vector at each targeting instant τ_k.

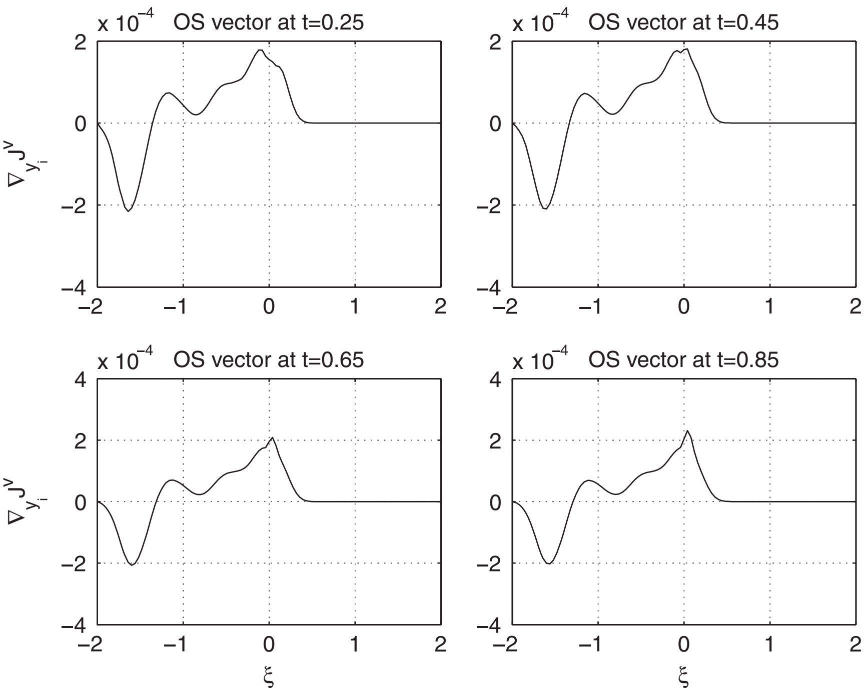


Figure 6. Observation sensitivity vector at different time steps for $R=50$.

Numerical experiments are performed to assimilate both routine and adaptive observations by using the 4D-Var method. First, we minimize the cost functional with routine observations. The initial guess provided to the minimization routine is produced by perturbing the given initial condition x_0^{t} with the Gaussian random errors of 3%. At each iteration of the minimization process, we monitor the current analysis estimate x_0^{a}, the corresponding value of the cost function $\mathscr{J}(x_0^{\mathrm{a}})$, as well as the gradient $\nabla\mathscr{J}(x_0^{\mathrm{a}})$. The evolution of the forecast error functional $\mathscr{J}^{\mathrm{v}}$ at t_{v} is also monitored through additional model forecasts at t_{v}. We then use adaptive observations along with routine observations to perform the 4D-Var data assimilation and monitor the same DAS output aspects as described above. The only difference from the assimilation of

DOI: 10.1002/fld

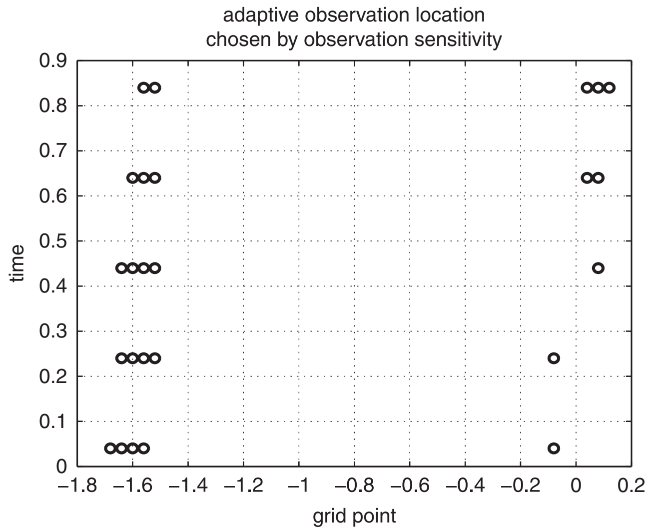
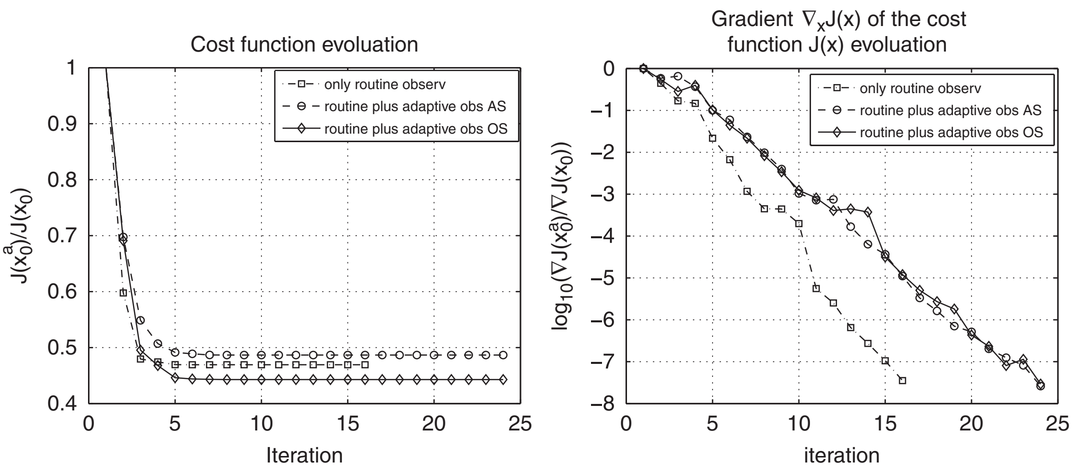


Figure 7. Locations of the adaptive observations based on the observation sensitivity vector at each targeting instant τ_k.



(a) Evolution of cost function for $R = 50$

(b) Evolution of the gradient for $R = 50$

Figure 8. The normalized value of the cost function $\mathscr{J}$ in (a) and gradient of cost function (b) when both routine and adaptive observations are assimilated. The gradients are shown on a semi-logarithmic scale.

routine observations only is that we incorporate one additional term for adaptive observations into the cost function in (1), $\mathscr{J}_{\mathrm{o}} = \mathscr{J}_{\mathrm{o}}^{\mathrm{r}} + \mathscr{J}_{\mathrm{o}}^{\mathrm{a}}$, where the cost component $\mathscr{J}_{\mathrm{o}}^{\mathrm{r}}$ incorporates the routine observations and the cost component $\mathscr{J}_{\mathrm{o}}^{\mathrm{a}}$ incorporates the additional adaptive observations. For each assimilation experiment, the evolution of the 4D-Var cost function and its gradient are displayed in Figures 8(a) and (b), respectively. The evolution of the forecast error $\mathscr{J}^{\mathrm{v}}(x_0^{\mathrm{a}})$ during the assimilation process is displayed in Figure 9 and it is noticed that although only a few additional adaptive observations were assimilated, their impact on the forecast error reduction is significant.

5. EFFECT OF RANDOM PERTURBATIONS ON THE FORECAST

An improved forecast can be obtained by using a few adaptive observations at locations that are dynamically identified based on the information provided by the AS or the OS fields. In our

DOI: 10.1002/fld

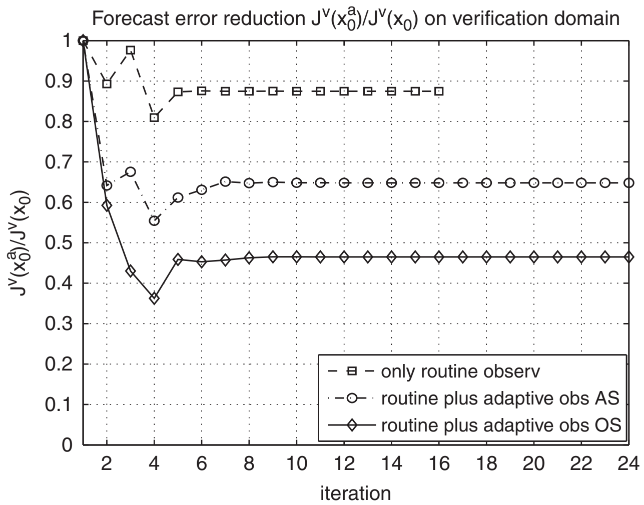


Figure 9. The evolution of the normalized forecast error at t_v over verification domain during the iterative data assimilation process.

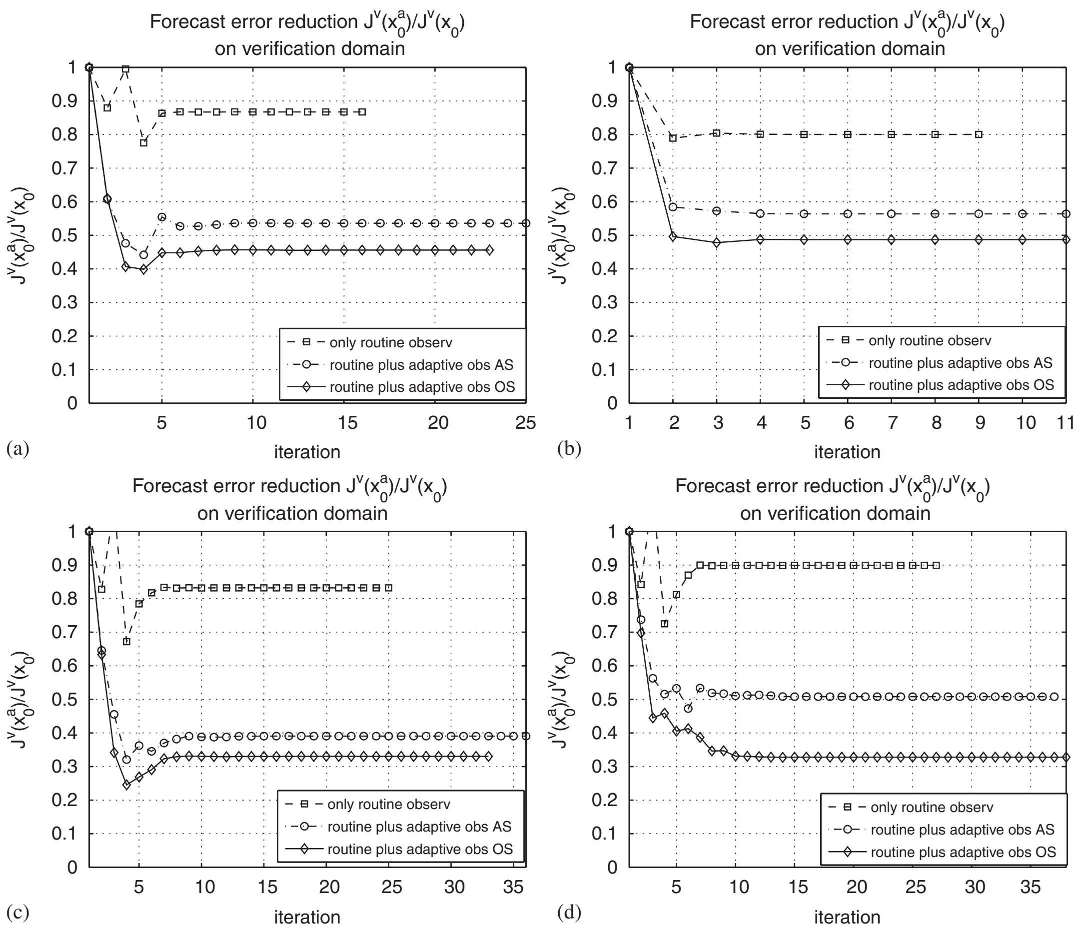


Figure 10. The normalized forecast errors at t_v over the verification domain for various sets of random perturbations in the data assimilation system input: (a) $\sigma_b = 0.08$, $\sigma_o = 0.06$ and $\sigma_i = 0.05$; (b) $\sigma_b = 0.05$, $\sigma_o = 0.09$ and $\sigma_i = 0.03$; (c) $\sigma_b = 0.08$, $\sigma_o = 0.04$ and $\sigma_i = 0.05$; and (d) $\sigma_b = 0.09$, $\sigma_o = 0.04$ and $\sigma_i = 0.05$.

DOI: 10.1002/fld

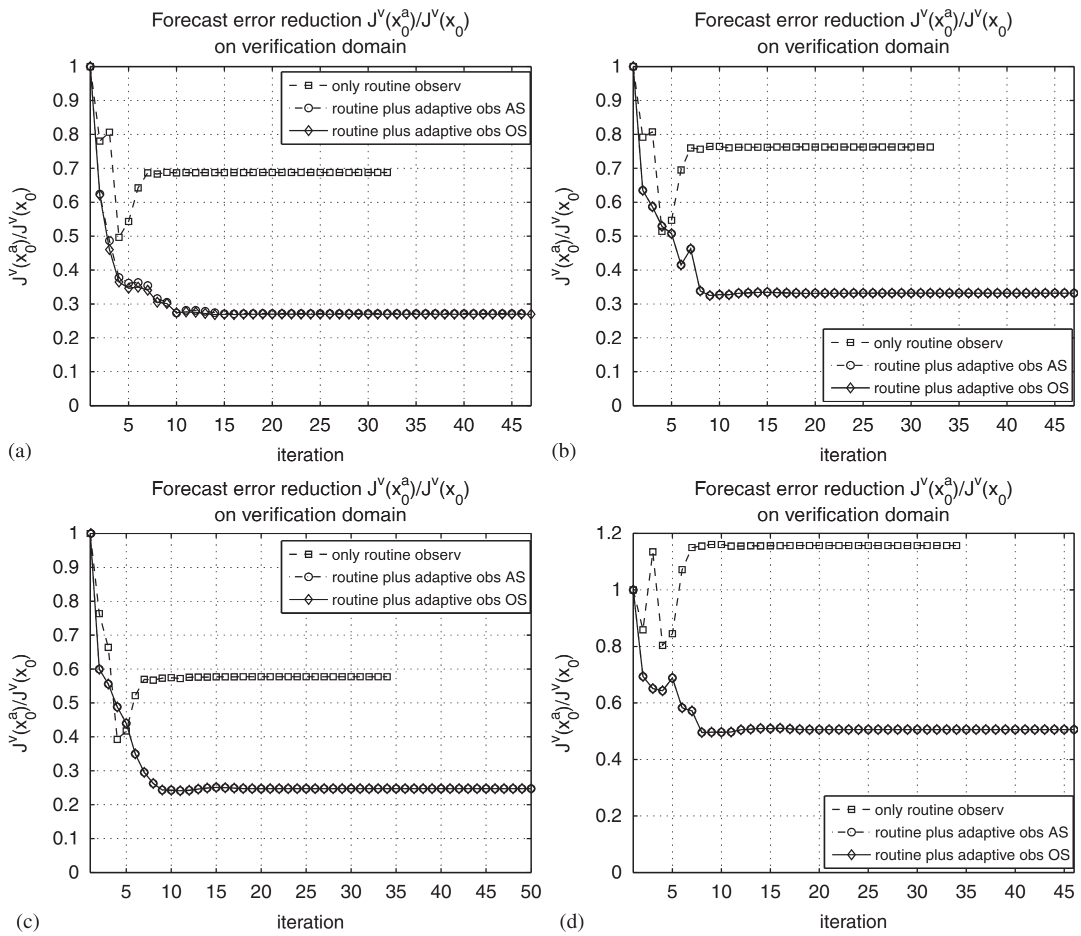


Figure 11. The normalized forecast errors at t_v over the verification domain with various sets of random perturbations in the data assimilation system input: (a) $\sigma_b = 0.08$, $\sigma_o = 0.03$ and $\sigma_i = 0.05$; (b) $\sigma_b = 0.09$, $\sigma_o = 0.03$ and $\sigma_i = 0.05$; (c) $\sigma_b = 0.15$, $\sigma_o = 0.05$ and $\sigma_i = 0.10$; and (d) $\sigma_b = 0.12$, $\sigma_o = 0.04$ and $\sigma_i = 0.05$.

experiments, random perturbations taken from a normal distribution $N(0, \sigma^2)$ are used to generate the background vector, observations, and the initial condition for data assimilation. Let σ_b, σ_o, and σ_i denote the standard deviation of the errors in background, observations, and the initial condition, respectively. In order to assess the effect of random perturbations on the performance of the AS and of the OS methods, we use several sets of standard deviations. We monitor the distribution of the forecast error defined in (3) over the verification domain at the verification time after the data assimilation has taken place with routine plus adaptive observations. The results for forecast errors are presented in Figures 10 and 11. These figures show that the forecast errors improve if adaptive observations are used together with routine observations. In our experiments, we have found that the impact of targeted observations on the forecast error reduction is closely determined by the ratios $r_o = \sigma_b/\sigma_o$ and $r_i = \sigma_b/\sigma_i$. Both AS and OS targeting methods work well if $r_i \leqslant 2$, but their performance still depends on the specification of r_o. Experiments performed with different sets of random perturbations reveal that the OS targeting outperforms the AS targeting if $r_o \leqslant 2.5$, as shown in Figures 10(a)–(d). The performance of the OS is almost the same as the performance of AS if $r_o > 2.5$, as shown in Figures 11(a)–(c). Furthermore, if $r_i > 2$ then the 4D-Var method fails to provide an improved forecast when only routine observations are used, see Figure 11(d), whereas the AS and OS observation targeting techniques will still provide improved forecasts and their performance is similar.

DOI: 10.1002/fld

6. SUMMARY AND CONCLUSIONS

In this work, the implementation and a comparative performance analysis are presented for two observation-targeting methods—the AS and the OS to identify optimal sites of the additional observational data of large impact on the forecast error reduction. We have found that the forecast error may be significantly reduced by adding only a few properly selected adaptive observations to the routine observations. The performance of the adaptive observations techniques was tested in a twin experiments' framework and it was found to closely depend on the relative magnitude of the random perturbations used to generate the background vector and the observations. In our experiments, we noticed that if the ratio r_o between the standard deviation of the errors in the background vector and the standard deviation of the errors in observations is less than 2.5, then the OS method outperforms the AS method, whereas if the above ratio is greater than 2.5 then both targeting methods perform almost the same. The AS method can be used for a large perturbed background vector and small perturbed observational data to reduce the forecast errors. The increased computational cost of implementing the OS method is justified in the presence of increased observational errors where our preliminary numerical experiments indicate that the OS approach may significantly outperform the AS approach to observation-targeting. The availability of the Hessian matrix information is a key ingredient in the implementation of the OS approach and providing this information for large-scale modeling is a particularly difficult task due to the present day limitations on computer memory and the complexity of the necessary SOA code development. Given the simplicity of our experimental setup, the numerical results in this work should be merely regarded as a presentation of a proof-of-concept. A necessary next step in this research is to implement and further test the OS method using a more realistic model and DAS. In this context, additional simplifications are necessary to facilitate the implementation and techniques of reduced-order modeling may be used to alleviate the computational cost of obtaining second-order derivative information.

Issues of threshold processes and nondifferentiability while outside the scope of the present paper could be addressed in further research.

ACKNOWLEDGEMENTS

The work of Prof. D. N. Daescu was supported by the National Science Foundation under grant DMS-0914937.

Prof. Navon's research was supported by the National Science Foundation under grant CMG-0931198.

REFERENCES

1. LeDimet FX, Ngodock HE, Luong B, Verron J. Sensitivity analysis in data assimilation. *Journal of the Meteorological Society of Japan* 1997; **75**:245–255.
2. Langland RH. Issues in targeted observing. *Quarterly Journal of the Royal Meteorological Society* 2005; **131**:3409–3425.
3. Langland RH, Gelaro R, Rohaly GD, Shapiro MA. Targeted observations in FASTEX: adjoint-based targeting procedures and data impact experiments in IOP17 and IOP18. *Quarterly Journal of the Royal Meteorological Society* 1999; **125**:3241–3270.
4. Palmer TN, Gelaro R, Barkmeijer J, Buizza R. Singular vectors, metrics, and adaptive observations. *Journal of the Atmospheric Sciences* 1998; **55**:633–653.
5. Gelaro R, Langland RH, Rohaly GD, Rosmond TE. An assessment of the singular vector approach to targeted observing using the FASTEX data-set. *Quarterly Journal of the Royal Meteorological Society* 1999; **125**: 3299–3328.
6. Daescu DN, Navon IM. Adaptive observations in the context of 4d-var data assimilation. *Meteorology and Atmospheric Physics* 2004; **85**:205–226.
7. Baker NL, Daley R. Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Quarterly Journal of the Royal Meteorological Society* 2000; **126**:1431–1454.
8. Daescu DN. On the sensitivity equations of four-dimensional variational (4D-Var) data assimilation. *Monthly Weather Review* 2008; **136**:3050–3065.
9. Doerenbecher A, Bergot T. Sensitivity to observations applied to FASTEX cases. *Nonlinear Processes in Geophysics* 2001; **8**:467–481.

10. Fourrié N, Doerenbecher A, Bergot T, Joly A. Adjoint sensitivity of the forecast to TOVS observations. *Quarterly Journal of the Royal Meteorological Society* 2002; **128**:2759–2777.
11. Langland RH, Baker NL. Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus* 2004; **56A**:459–476.
12. Zhu Y, Gelaro R. Observation sensitivity calculations using the adjoint of the Gridpoint Statistical Interpolation (GSI) analysis system. *Monthly Weather Review* 2008; **136**:335–351.
13. Trémolet Y. Computation of observation sensitivity and observation impact in incremental variational data assimilation. *Tellus* 2008; **60A**:964–978.
14. LeDimet FX, Navon IM, Daescu DN. Second order information in data assimilation. *Monthly Weather Review* 2002; **130**(3):629–648.
15. Giering R, Kaminski T. Recipes for adjoint code construction. *ACM Transactions on Mathematical Software* 1998; **24**:437–474.
16. Rabier F, Courtier P. Four-dimensional assimilation in the presence of baroclinic instability. *Quarterly Journal of the Royal Meteorological Society* 1992; **118**:649–672.
17. Xu L, Rosmond T, Goerrs J, Chua B. Toward a weak constraint operational 4D-Var system: application to the Burgers' equation. *Meteorologische Zeitschrift* 2007; **16**:741–753.
18. Liu DC, Nocedal J. On the limited memory BFGS method for large-scale minimization. *Mathematical Programming* 1989; **45**:503–528.
19. Navon IM, Zou X, Derber J, Sela J. Variational data assimilation with an adiabatic version of the NMC spectral model. *Monthly Weather Review* 1992; **120**:1433–1446.
20. Wang Z, Navon IM, LeDimet FX, Zou X. The second order adjoint analysis: theory and applications. *Meteorology and Atmospheric Physics* 1992; **50**:3–20.

DOI: 10.1002/fld