Comput Geosci (2011) 15:477–488
DOI 10.1007/s10596-010-9217-z

ORIGINAL PAPER

Observation targeting with a second-order adjoint method for increased predictability

Humberto C. Godinez · Dacian N. Daescu

Received: 22 June 2010 / Accepted: 8 November 2010 / Published online: 18 December 2010

Abstract The efficiency of current adjoint-based observations targeting strategies in variational data assimilation is closely determined by the underlying assumption of a linear propagation of initial condition errors into the model forecasts. A novel targeting strategy is proposed in the context of four-dimensional variational data assimilation (4D-Var) to account for nonlinear error growth as the forecast lead time increases. A quadratic error growth model is shown to maintain the accuracy in tracking the nonlinear evolution of initial condition perturbations, as compared to the first-order approximation. A second-order adjoint model is used to provide the derivative information that is necessary in the higher-order Taylor series approximation. The observation targeting approach relies on the dominant eigenvectors of the Hessian matrix associated with a specific forecast error aspect as an indicator of the directions of largest quadratic error growth. A comparative qualitative analysis between observation targeting based on first- and second-order adjoint information is presented in idealized 4D-Var experiments with a two-dimensional global shallow-water model. The results indicate that accounting for the quadratic error growth in the targeting strategy is of particular benefit as the forecast lead time increases.

Keywords Second-order adjoint · Targeting strategies · Data assimilation · 4D-Var

Mathematics Subject Classifications (2010) 35B20 · 34E10 · 65M99 · 65M32 · 46N40

1 Introduction

Observation targeting strategies aim to identify optimal regions where collection of supplemental data can improve the state forecast, of a data assimilation and forecast system, at a specified lead time and over a selected geographical region. A statistical formulation and the mathematical framework for the adaptive observations problem in numerical weather prediction is provided in the work of Berliner et al. [4]. Adjoint modeling has been an essential tool for the development of targeting strategies in the context of variational data assimilation methods. The adjoint of the tangent linear model associated with an atmospheric model is a key ingredient to implementing various targeting strategies, such as gradient sensitivity, dominant singular vectors, and sensitivity to observations [2, 3, 8, 9, 22, 23, 35]. Analysis schemes based on ensemble of forecasts provide a feasible alternative to variational methods and observation targeting strategies in an ensemble Kalman

This research was supported by the Department of Energy at Los Alamos National Laboratory under contracts DE-AC52-06NA25396 and the DOE Office of Science Advanced Computing Research (ASCR) program in Applied Mathematical Sciences, and by the National Science Foundation under grant DMS-0914937.

H. C. Godinez (✉)
Applied Mathematics and Plasma Physics, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
e-mail: hgodinez@lanl.gov

D. N. Daescu
Department of Mathematics & Statistics, Portland State University, Portland, OR 97207, USA
e-mail: daescu@pdx.edu

filter data assimilation system are discussed in references [1, 5, 6, 18].

The first-order adjoint (FOA) model provides an efficient approach to evaluate the gradient of a scalar-valued forecast aspect, typically a forecast error measure, with respect to the model initial state. As such, the FOA gives a first-order approximation to the forecast impact as a result of perturbations in the initial state of the atmospheric model. It may be used to identify the regions where initial condition errors have a potentially large impact on determining the forecast uncertainty [36]. The accuracy of this linear approximation is limited by the magnitude of the perturbation and by the time length of the forecast.

Properly accounting for the nonlinear error growth is also an unresolved issue in the design of observation targeting strategies using singular vectors. The dominant singular vectors of the tangent linear model associated with a nonlinear dynamical system provide information on the initial condition perturbations of maximum linear growth at a specified lead time. The computational cost of evaluating the singular vectors is significant however, the experience gained from observation targeting experiments is that adaptive sampling of the sensitivity fields determined by a small number of leading singular vectors is effective in reducing the short-range forecast errors [7, 14, 33]. The singular vectors are closely determined by the metric used to measure the error growth and may be used to accurately characterize the propagation of initial condition errors as long as the underlying assumption of linear error evolution remains valid [24, 35].

As the forecast lead time increases, nonlinearities in the dynamical model may become increasingly dominant and the accuracy of the FOA and singular vector methods to track the initial-condition error propagation is impaired. This poses a limitation on the time window for which targeting strategies based on first-order derivative information are reliable. To overcome this practical difficulty and to increase the effectiveness of adjoint targeting strategies, a second-order adjoint (SOA) model may be considered to capture the quadratic terms in the error growth approximation. An overview of the SOA model implementation and applications to variational data assimilation is provided in [25]. In the present work a novel targeting strategy based on SOA information is considered and numerical experiments are used to provide a comparative analysis between the first-order and the second-order adjoint-based observation targeting guidance. Limitations of the linear approximation and the necessity to incorporate high-order derivative information are investigated by using first and second-order Taylor approximations to model the nonlinear impact of initial-condition perturbations on a forecast error functional. The SOA model provides the product between the Hessian matrix associated to the model forecast functional and a predetermined vector direction. It allows to evaluate the second derivative term in the Taylor approximation to the evolution of initial condition perturbations. As such SOA provides the information necessary to model a quadratic error growth. The dominant eigenvectors and eigenvalues of the Hessian matrix provide information about the maximum quadratic error growth due to their mathematical properties. Based on this concept, a new targeting strategy is proposed that uses the dominant eigenvectors and eigenvalues of the Hessian matrix to identify a sensitivity field that best captures the quadratic error growth.

Section 2 briefly revisits the four dimensional variational (4D-Var) data assimilation and the FOA and SOA models. In Section 3 the implementation of the FOA and SOA models to a shallow water (SW) model is presented. The accuracy of the first and second-order Taylor approximations to the nonlinear forecast error functional is analyzed in terms of the magnitude of the initial condition perturbations and the forecast lead time. A novel targeting strategy based on the eigenvalues and eigenvectors of the Hessian matrix of the forecast aspect is presented in Section 4 using the SW model. Numerical experiments are presented in Section 5 along with a comparative analysis between FOA and SOA observation targeting guidance. Conclusions and future work are presented in Section 6.

2 Data assimilation and adjoint modeling

Atmospheric data assimilation techniques combine information from a model of the atmospheric dynamics, observational data, and error statistics to provide an optimal estimate (analysis) of the model state. Theoretical foundations of modern data assimilation techniques can be found in the books of Daley [12], Kalnay [21], Lewis et al. [29] and Evensen [13].

Given an initial state $\mathbf{x}_0$, let $\mathcal{M}_{t_0 \to t_N}$ denote the discrete atmospheric model (forward model) that evolves the state from the initial time t_0 to time t_N

$$\mathbf{x}_N = \mathcal{M}_{t_0 \to t_N}(\mathbf{x}_0)\,. \tag{1}$$

The 4D-Var data assimilation [26] provides an optimal estimate $\mathbf{x}_0^a$ of the true initial condition $\mathbf{x}_0^t$ by minimizing a cost functional that measures the discrepancy be-

tween the model state, a prior (background) estimate, and time distributed observational data

$$\mathcal{J}(\mathbf{x}_0) = \left(\mathbf{x}_0 - \mathbf{x}^b\right)^T \mathbf{B}^{-1} \left(\mathbf{x}_0 - \mathbf{x}^b\right) + \sum_{i=0}^{k} \left(\mathbf{y}_i - H_i[\mathbf{x}_i]\right)^T \mathbf{R}_i^{-1} \left(\mathbf{y}_i - H_i[\mathbf{x}_i]\right) \quad (2)$$

where $\mathbf{x}^b$ is the background estimate to the initial state, $\mathbf{y}_i \in \mathbb{R}^{n_i}, i = 1, \ldots, k$ is the observation vector at t_i, $\mathbf{x}_i = \mathcal{M}_{t_0 \to t_i}(\mathbf{x}_0)$ is the model state at t_i, and H_i is the observational operator mapping the state into observations at t_i. Statistical information on the background error and observational errors is used to specify the weighting matrices $\mathbf{B}$ and $\mathbf{R}_i$ that are representations in the data assimilation system of the background error covariance $\mathbf{B}^t$ and the observation error covariance $\mathbf{R}_i^t$ at time t_i respectively.

Adaptive observations are supplementary data collected to reduce the error of a forecast aspect at verification time $t_v > t_k$ over a verification domain $\mathcal{D}_v$, expressed as

$$\mathcal{J}_v(\mathbf{x}_v) = \frac{1}{2} \left\langle \mathbf{P}\left(\mathbf{x}_v - \mathbf{x}_v^t\right), \mathbf{P}\left(\mathbf{x}_v - \mathbf{x}_v^t\right) \right\rangle_{\mathbf{E}} \quad (3)$$

where $\mathbf{x}_v^t$ denotes the true state at the verification time, and $\mathbf{x}_v$ is the model forecast state of the system at time t_v. The operator $\mathbf{P}$ is the projection operator on $\mathcal{D}_v$ that selects the state vector components inside the verification domain from the global domain and it is represented as a diagonal matrix satisfying $\mathbf{P}^*\mathbf{P} = \mathbf{P}^2 = \mathbf{P}$. The inner product $\langle , \rangle_{\mathbf{E}}$ is defined as $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{E}} = \langle \mathbf{x}, \mathbf{E}\mathbf{y} \rangle$, where $\mathbf{E}$ is a symmetric positive definite matrix. The total energy norm corresponds to a diagonal matrix $\mathbf{E}$ and it is often used to facilitate the practical implementation [19, 35]. The measure (Eq. 3) is the *forecast error functional* at the verification time t_v over the verification domain $\mathcal{D}_v$. Targeting strategies are used to identify the optimal locations for adaptive observations within a data assimilation framework. In practical applications, the forecast aspect must rely on the model alone since at the experimental design stage the actual forecast error is unknown. Therefore, the forecast error functional (Eq. 3) may be used only in a *posteriori* studies for testing the efficiency of various observation targeting methods. Ensemble forecasting [28] provides a feasible approach to assess the forecast errors and the ensemble-mean state may be used to define the forecast aspect for sensitivity calculations, as discussed in references [1, 37].

2.1 Taylor expansion of the forecast error functional

To assess the propagation of the errors in the initial condition, the forecast aspect is considered as a functional of the initial state. The functional $\mathcal{J}_v$ implicitly depends on the initial condition $\mathbf{x}_0$ of Eq. 1

$$\mathcal{J}_v(\mathbf{x}_N) = \mathcal{J}_v\left(\mathcal{M}_{t_0 \to t_N}(\mathbf{x}_0)\right), \quad (4)$$

where $\mathcal{M}_{t_0 \to t_N}$ is the nonlinear model integration from t_0 to $t_N = t_v$,

$$\mathcal{M}_{t_0 \to t_N} = \mathcal{M}_{N-1} \circ \cdots \circ \mathcal{M}_0(\mathbf{x}_0). \quad (5)$$

It is noticed that the measure (Eq. 3) is a quadratic functional of the forecast state $\mathbf{x}_v = \mathbf{x}_N$ however, its nonlinear dependence on the initial state is of more general type due to the nonlinearities in the model forecast (Eqs. 4 and 5). A perturbation $\delta\mathbf{x}_0$ in the initial state will result in a perturbation $\delta\mathcal{J}_v = \mathcal{J}_v(\mathbf{x}_N + \delta\mathbf{x}_N) - \mathcal{J}_v(\mathbf{x}_N)$, where $\mathbf{x}_N + \delta\mathbf{x}_N = \mathcal{M}_{t_0 \to t_N}(\mathbf{x}_0 + \delta\mathbf{x}_0)$, and the second order Taylor approximation is expressed in terms of $\delta\mathbf{x}_0$ as

$$\delta\mathcal{J}_v \approx \delta\mathbf{x}_0^T \nabla_{\mathbf{x}_0} \mathcal{J}_v(\mathbf{x}_N) + \frac{1}{2} \delta\mathbf{x}_0^T \nabla_{\mathbf{x}_0}^2 \mathcal{J}_v(\mathbf{x}_N)\, \delta\mathbf{x}_0 \quad (6)$$

Algorithmic differentiation and techniques for adjoint code development are discussed in references [16, 17]. Using chain rule differentiation, the tangent linear model (TLM) of the nonlinear forward integration (Eq. 5) is expressed as

$$\mathbf{M}_{0,N} = \mathbf{M}_{N-1}(\mathbf{x}_{N-1})\mathbf{M}_{N-2}(\mathbf{x}_{N-2}) \ldots \mathbf{M}_0(\mathbf{x}_0) \quad (7)$$

where $\mathbf{M}_i(\mathbf{x}_i)$ denotes the Jacobian matrix associated with the model integration from the time t_i to t_{i+1}, $\mathbf{x}_{i+1} = \mathcal{M}_{t_i \to t_{i+1}}(\mathbf{x}_i)$. The adjoint of the TLM is obtained as the transpose of Eq. 7,

$$\mathbf{M}_{0,N}^* = \mathbf{M}_0^*(\mathbf{x}_0) \ldots \mathbf{M}_{N-2}^*(\mathbf{x}_{N-2})\mathbf{M}_{N-1}^*(\mathbf{x}_{N-1}) \quad (8)$$

where $\mathbf{M}_i^*(\mathbf{x}_i)$ denotes the adjoint (transpose) of $\mathbf{M}_i(\mathbf{x}_i)$, and the gradient $\nabla_{\mathbf{x}_0} \mathcal{J}_v(\mathbf{x}_N)$ is evaluated through the backward FOA model integration

$$\lambda_N = \nabla_{\mathbf{x}_N} \mathcal{J}_v(\mathbf{x}_N) \quad (9)$$

$$\lambda_i = \mathbf{M}_i^*(\mathbf{x}_i)\, \lambda_{i+1}, \quad i = N-1, \ldots, 0 \quad (10)$$

The solution of the adjoint Eqs. 9–10 provides the gradient of Eq. 3 with respect to the initial conditions as $\lambda_0 = \nabla_{\mathbf{x}_0} \mathcal{J}_v(\mathbf{x}_N)$.

Second-order derivative information, as the product of the Hessian $\nabla_{\mathbf{x}_0}^2 \mathcal{J}_v(\mathbf{x}_N)$ times a user-defined vector, may be obtained by developing a SOA model.

2.2 SOA model equations

The equations of the discrete SOA model associated with Eqs. 1 and 4 are obtained from a linearization of both forward and adjoint model equations

$$\nu_N = \nabla^2_{\mathbf{x}_N} \mathcal{J}_v(\mathbf{x}_N)\,\mu_N = \mathbf{P}^T\mathbf{E}\mathbf{P}\mu_N \tag{11}$$

$$\nu_i = \mathbf{M}_i^*(\mathbf{x}_i)\,\nu_{i+1} + \frac{\partial}{\partial \mathbf{x}_i}\left[\mathbf{M}_i^*(\mathbf{x}_i)\,\bar{\lambda}_{i+1}\right]\mu_i, \quad i = N-1,\ldots,0 \tag{12}$$

where μ_i is the solution to the TLM integration

$$\mu_0 = \mathbf{w} \tag{13}$$

$$\mu_{i+1} = \mathbf{M}_i(\mathbf{x}_i)\,\mu_i, \quad i = 0,\ldots,N-1, \tag{14}$$

$\mathbf{w}$ is an user-defined vector, and the notation $\bar{\lambda}_{i+1}$ in the last term of Eq. 13 indicates that the state derivative applies to the $\mathbf{M}_i^*(\mathbf{x}_i)$ operator only while treating the adjoint variables λ_{i+1} as constants [11, 25].

The solution of the SOA model (Eqs. 11–13) provides the Hessian vector product $\nabla^2_{\mathbf{x}_0}\mathcal{J}(\mathbf{x}_0)\,\mu_0 = \nu(t_0)$ that is required to evaluate the second-order term in the Taylor approximation (Eq. 6), thus providing the quadratic term for the evolution of perturbations in the initial state of the forecast model.

3 The SW model, FOA and SOA Taylor approximations

Testing in an idealized framework and presentation of a proof-of-concept is a necessary stage in the development of new targeting methodologies, and needs to be performed prior to the implementation and validation in the much more complex environment of an operational system. A global two dimensional shallow water (SW) model on a sphere is used for the numerical experiments. Shallow water models have been shown to capture important features of atmospheric and oceanic dynamics and have been used as prototypes in data assimilation studies [31, 34, 38]. The equations of the SW model are

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -f\mathbf{k}\times\mathbf{v} - \nabla\phi, \tag{15}$$

$$\frac{\partial\phi}{\partial t} = -\nabla\cdot\left[(\phi-\phi_s)\,\mathbf{v}\right], \tag{16}$$

where $\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla$, $\mathbf{v} = u\mathbf{i} + v\mathbf{j}$ with $\mathbf{i}, \mathbf{j}, \mathbf{k}$ being the unit vectors in the three orthonormal directions on the sphere, u and v are the zonal and meridional velocity components, respectively, h is the fluid depth, h_s is the bottom topography, g is the gravitational constant, $\phi = gh$, $\phi_s = gh_s$, and f is the Coriolis parameter. The grid-point norm on the state space $\mathbf{x} = (u, v, h)$ is the total energy norm, induced by the inner product

$$\langle \mathbf{x}, \mathbf{x}\rangle_{\mathbf{E}} = \frac{1}{2}\left(u^2 + v^2 + \frac{g}{h_0}h^2\right) \tag{17}$$

such that $\mathbf{E}$ is a diagonal matrix. The units in Eq. 17 are of $\mathrm{m}^2\mathrm{s}^{-2}$. A Godunov type finite volume discretization method with the Van-Leer transport scheme, as described in [30], is used to provide the numerical solution to the SW equations. Computations are done on a 2.5° × 2.5° grid with a time step $\Delta t = 450$s and the verification time is set at $t_v = t_0 + 24$ h. The reference state ('truth') $\mathbf{x}_0^t$ is taken from the trajectory produced by a numerical integration of the SW model using as initial condition the 500 hPa ERA-40 data set from the European Centre for Medium-Range Weather Forecasts (ECMWF), valid for March 15 2002 at 06:00 h. The background state $\mathbf{x}^b$ is taken from a 6-hour model simulation initialized at $t_0 - 6$ h with the ERA-40 data set valid for March 15 2002 at 00:00 h. The difference between the 24 h forecasts initiated from $\mathbf{x}_0^t$ and $\mathbf{x}^b$, respectively, exhibits a high discrepancy in the region $[55°W, 35°W] \times [52°N, 65°N]$ which is taken as the verification domain $\mathcal{D}_v$ at t_v. The discrete TLM, FOA, and SOA models were implemented using the Automatic Differentiation package TAMC [15]. Figure 1 illustrates the computational flow chart of the tangent linear model and the FOA model. The FOA model is integrated backward in time to obtain the sensitivity of the final state with respect to the initial state. This backward integration requires storage of the full trajectory of the forward model to obtain the necessary derivative information at each time step. The discrete SOA can be seen as the action of the Hessian matrix of the scalar forecast aspect of interest on a vector. The code for the second-order adjoint can be implemented using the forward over reverse mode by taking the tangent of a forward-backward integration. Figure 2 shows the flow chart for the computations involved in the second-order adjoint model. A single backward integration of the

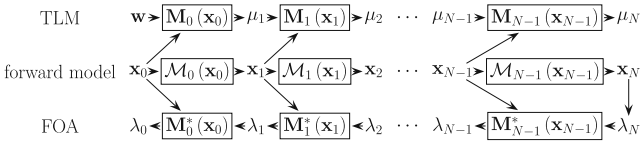

Fig. 1 Flow chart for the computation of the tangent linear model and the adjoint model

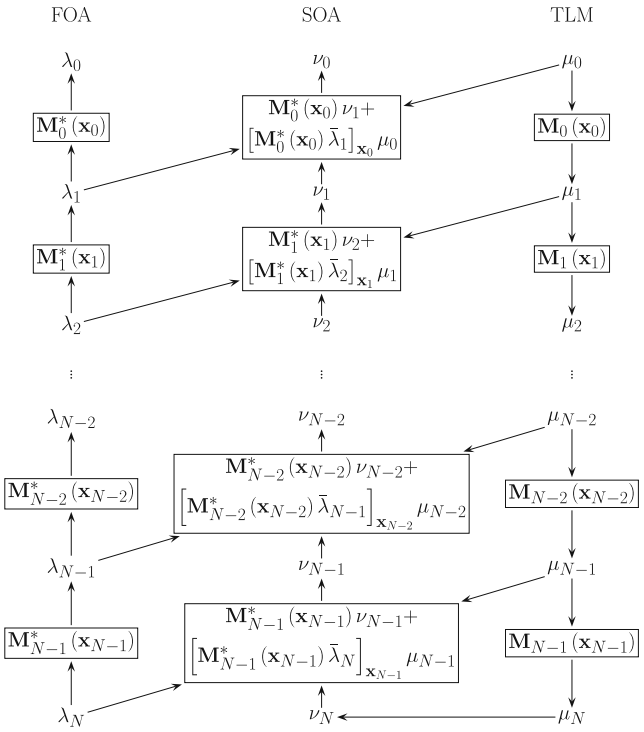

Fig. 2 Flow chart for the computation of the SOA model

SOA model requires storage of the full trajectory of the forward model (Eq. 1) and of the tangent linear model (Eqs. 13, 14), as well as the FOA model integration (Eqs. 9, 10). Although the potential benefits of incorporating second-order adjoint information in atmospheric data assimilation are unanimously recognized [25], to date, the increased storage and computational demands as well as the complexity of the code development have prevented the implementation of a second-order adjoint model in operational systems.

3.1 Taylor approximation with adjoint models

Given the nonlinearity of the forward model, the validity of the Taylor approximation (Eq. 6) is closely determined by the magnitude of perturbations in the initial condition $\mathbf{x}_0$ and the forecast lead time. To investigate this aspect, forecast perturbations are computed with the SW model together with their first and second-order Taylor approximations using the FOA and SOA models. To corroborate the accuracy of the approximations to $\delta\mathcal{J}_v$, the initial condition, taken from the background state, is perturbed according to

$$\mathbf{x}_0(\varepsilon) = \mathbf{x}_0 + \delta\mathbf{x}_0(\varepsilon), \quad \delta\mathbf{x}_0(\varepsilon) = \varepsilon\left(\mathbf{x}_0 - \mathbf{x}^t(t_0)\right),$$

where ε is a coefficient that controls the magnitude of the perturbation in the initial condition of the forward model. The perturbation $\delta\mathcal{J}_v$, as well as the adjoint-based approximations, are computed for values of the perturbation coefficient ε ranging from -0.4 to 0.4 with increments of 0.01 and fixed $t_v = 24$ h, then for a time-varying forecast lead time $t_v - t_0$ ranging from 6-hour to 72-hour with 1 h increments.

Figure 3 (left) shows the perturbation $\delta\mathcal{J}_v$ of the forecast error, and its first and second order Taylor approximations. It is noticed that the second-order approximation remains accurate over a wider range of perturbations as compared to the first-order approximation. Figure 3 (right) shows the time evolution of the Taylor approximation using the FOA and SOA models, as a function of t_v, while keeping $\varepsilon = 0.1$ fixed. As the verification time increases the second-order approximation remains significantly more accurate than the first-order approximation. Similar results were obtained with various values of $\varepsilon \neq 0$.

It must be noticed that the perturbation growth is time dependent, this is, the solutions of the forward and adjoint models depend on the verification time t_v. An important question to address is the accuracy of the approximation as the verification time is increased. The linear approximation can lose accuracy if there is a strong nonlinear time dependence of the forward model which is not accurately captured in the FOA model. Traditional targeting strategies account only for

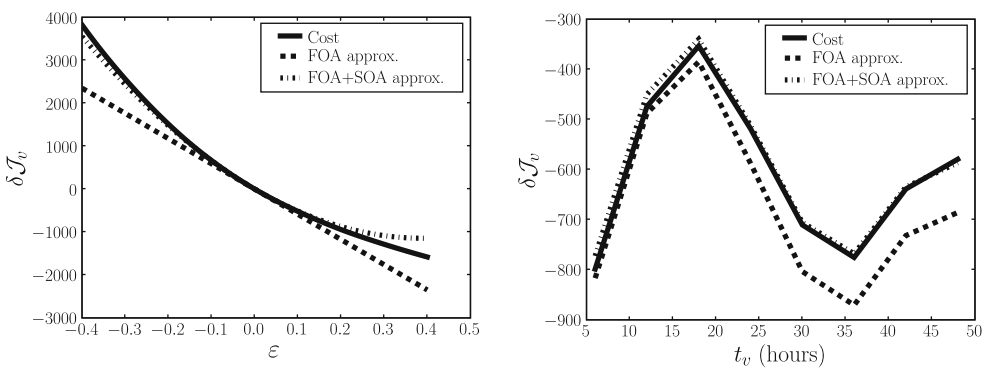

Fig. 3 *Left figure*: Taylor approximation (Eq. 6) of the perturbation of the forecast error functional as a function of initial condition perturbations coefficient ε. *Right figure*: Time evolution of Eq. 6 as a function of the verification time t_v, with $\varepsilon = 0.1$

linear error propagation and differ on the selection of the norm used to measure the error growth propagation (e.g. total energy vs. error covariance metric [35]). The limited accuracy of the first-order approximation is a major difficulty in extending the forecast time interval and the use of second-order derivative information may prove to be of relevance to practical applications.

4 Targeting using SOA information

Applications of the FOA sensitivity analysis during field experiments to collect targeted observations are presented in [23], which gives the fundamentals for targeting strategies based on adjoint modeling.

The FOA model provides the gradient of the forecast error functional $\mathcal{J}_v$ with respect to the initial condition $\mathbf{x}_0$ of the forward model. The gradient is used to define a space-distributed sensitivity field

$$F_v(x, y) = \left\| \nabla_{\mathbf{x}_0} \mathcal{J}_v(x, y) \right\|_2, \tag{18}$$

where (x, y) are grid-point coordinates. The function F_v is evaluated at each grid node were $\left\| \nabla_{\mathbf{x}_0} \mathcal{J}_v(x, y) \right\|_2$ denotes the Euclidean norm of the elements in the gradient that correspond to the (x, y) grid-point. In the FOA targeting approach supplementary observations are taken at locations where F_v exhibits the largest magnitude.

To accurately track the propagation of perturbations, we propose a new targeting method to incorporate SOA information. Consider the second term in the Taylor approximation (Eq. 6)

$$\frac{1}{2} \delta \mathbf{x}_0^T \mathbf{H} \delta \mathbf{x}_0, \tag{19}$$

where $\mathbf{H} = \nabla^2_{\mathbf{x}_0} \mathcal{J}_v(\mathbf{x}_N)$ denotes the Hessian matrix associated with $\mathcal{J}_v$. Without loss of generality, consider initial perturbations with unit two-norm, that is $\|\delta \mathbf{x}_0\|_2 = 1$, then Eq. 19 is a *Rayleigh–Ritz* ratio

$$\frac{\delta \mathbf{x}_0^T \mathbf{H} \delta \mathbf{x}_0}{\delta \mathbf{x}_0^T \delta \mathbf{x}_0}. \tag{20}$$

The vector for which the Rayleigh–Ritz ratio (Eq. 20) is maximized provides the direction of maximal quadratic error propagation in the second-order term of the Taylor approximation. Since $\mathbf{H}$ is a symmetric matrix, all its eigenvalues are real and the largest magnitude of Eq. 20 is provided by the eigenvector associated with the eigenvalue of largest magnitude [20].

In [8, 35] the singular vectors of the tangent linear model are used to define a sensitivity field for targeted observations. Following a similar approach, we define the sensitivity function based on the leading eigenvectors of the Hessian.

Let σ_i be the ith eigenvalue of $\mathbf{H}$, ordered so that $|\sigma_1| \geq |\sigma_2| \geq \ldots \geq |\sigma_n|$, and let $\mathbf{v}_i$ be its corresponding eigenvector. Consider the first m leading eigenvectors $\mathbf{v}_i$, $i = 1, \ldots, m$, where $m \ll n$, n being the dimension of the matrix. The SOA-based sensitivity field is defined as

$$F_m(x, y) = \sum_{i=1}^{m} \frac{|\sigma_i|}{|\sigma_1|} \|\mathbf{v}_i(x, y)\|_2^2, \tag{21}$$

where (x, y) are the grid-point coordinates. The function F_m is evaluated at each grid-point where $\|\mathbf{v}_i(x, y)\|_2$ denotes the Euclidean norm of the eigenvector components at the (x, y) grid-point.

5 Numerical experiments

The numerical experiments and analysis presented in this section were performed in the twin experiment framework using a global two-dimensional shallow water model on a sphere. The setup of the model is the same as in Section 3. A 4D-Var data assimilation scheme is implemented using the SW model and the

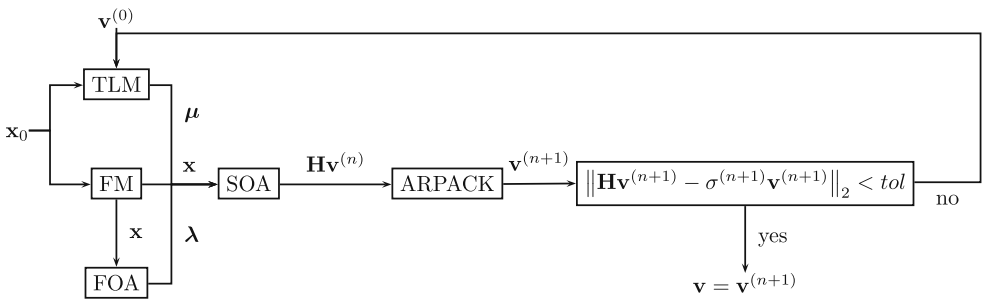

Fig. 4 Flow chart for the computation of the eigenvectors of the Hessian matrix using ARPACK

Table 1 CPU time for the computation of the leading 10 eigenvalues and eigenvectors for $t_i = 0$ to 6 h, at 1-hour increments

t_i (h)	0	1	2	3
CPU (s)	1698.52	1651.28	1589.41	1542.74
t_i (h)	4	5	6	
CPU (s)	1477.11	1429.97	1368.50	

Experiments performed with a 1.86 GHz Xeon Quad Core 5320 Processor

assimilation window of [0 h, 6 h]. To facilitate the comparison, the data assimilation setup differs only in the insertion of adaptive observations in the cost functional (Eq. 2). For each experiment, the minimization is performed using a limited memory quasi-Newton algorithm [32] and a discrete FOA model to provide the necessary gradient information. In all experiments, the verification state used to quantify the forecast errors is obtained from a nonlinear model forecast initialized from the true state, $\mathbf{x}_v^t = \mathcal{M}_{t_0 \to t_v}(\mathbf{x}_0^t)$. Numerical results and analysis are first presented for the verification time $t_v = 24$ h and in a second set of experiments for $t_v = 48$ h.

5.1 Eigenvectors of the Hessian of the SW forecast

At the 2.5-degree grid resolution the dimension of the discrete state vector $\mathbf{x}$ is $n \sim 3 \times 10^4$, such that the number of entries in the Hessian matrix $\mathbf{H}$ is of order of 10^9. Storage of this matrix is not feasible, however, iterative methods that only require the action of the matrix on a vector can be implemented to obtain the Hessian eigenvalues and eigenvectors. The Arnoldi Package (ARPACK) [27] is used to compute the leading eigenpairs of $\mathbf{H}$. Figure 4 illustrates the flow chart to compute the eigenvectors of the Hessian matrix with ARPACK using the SOA model. Each iteration involves a forward integration of the forecast and TLM models as well as a backward integration of the FOA and SOA models. This process requires storage of the trajectories $\mathbf{x}$, $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$. The output $\boldsymbol{v}_0$ of the SOA model is the Hessian vector product necessary in the Arnoldi iteration. Table 1 shows the CPU time (in seconds) necessary for the computation of 10 leading eigenpairs of the Hessian matrix associated with a 24-hour forecast error measure and a time-varying initial state within the 6-hour assimilation window, at 1-hour increments. The computational overhead of a SOA integration is about 3 times that of a FOA integration. The sensitivity field for the SOA targeting strategy is defined by Eq. 21 and was constructed using 10 eigenpairs. Additional experiments performed with 20 and 30 eigenpairs of the Hessian provided no significant changes in the configuration of the sensitivity field and did not modified the location of the adaptive observations.

The computational time increases as the forecast lead time increases, since the trajectory length involved in the forward-backward integration of the forward, TLM, FOA, and SOA models increases for eigenvectors with targeting time closer to $t_i = 0$. In our experiments, a 6-hour increase in the forecast lead time resulted in an increase of ∼24% in the computational time.

5.2 Targeted observations and data assimilation experiments

5.2.1 24-hour forecast lead time assimilation

A 24-hour forecast lead time experiment for targeted observations using FOA and SOA information is presented. Table 2 shows the leading eigenvalue of $\mathbf{H}$ for a targeting time t_i from 0 h to 6 h, at 1-hour increments with a target time $t_v = t_0 + 24$ h. The leading eigenvalue decreases in magnitude as t_i increases, which may indicate that as the forecast lead time is shortened the impact of the second-order derivative information decreases. This is consistent with the analysis in Section 3.1, where the contribution of the SOA was shown to diminish for forecasts closer to the initial time.

The sensitivity fields obtained with the FOA sensitivity function (Eq. 18) are shown in Fig. 5 for $t_i = t_0$ (top-left) and $t_i = t_0 + 6$ h (top-right). The locations of 20 adaptive observations (marked with •) correspond to the grid-points where the sensitivity fields have the largest magnitude. For comparison, Fig. 5 also displays the sensitivity fields obtained with the SOA sensitivity function (Eq. 21), using $m = 10$ leading eigenvectors of the Hessian matrix, at $t_i = t_0$ (bottom-left) and $t_i = t_0 + 6$ h (bottom-right). The difference in the sensitivity fields between the FOA and SOA methods illustrates the different type of perturbation growth being measured. Both FOA and SOA fields are time-varying, thus the location of targeted observations depends on the targeting instant.

Table 2 Leading eigenvalues of the Hessian matrix for $t_i = 0$ h to 6 h, at 1-hour increments

t_i (h)	0	1	2	3
σ_1	8.455	7.272	6.721	5.932
t_i (h)	4	5	6	
σ_1	5.270	4.355	3.541	

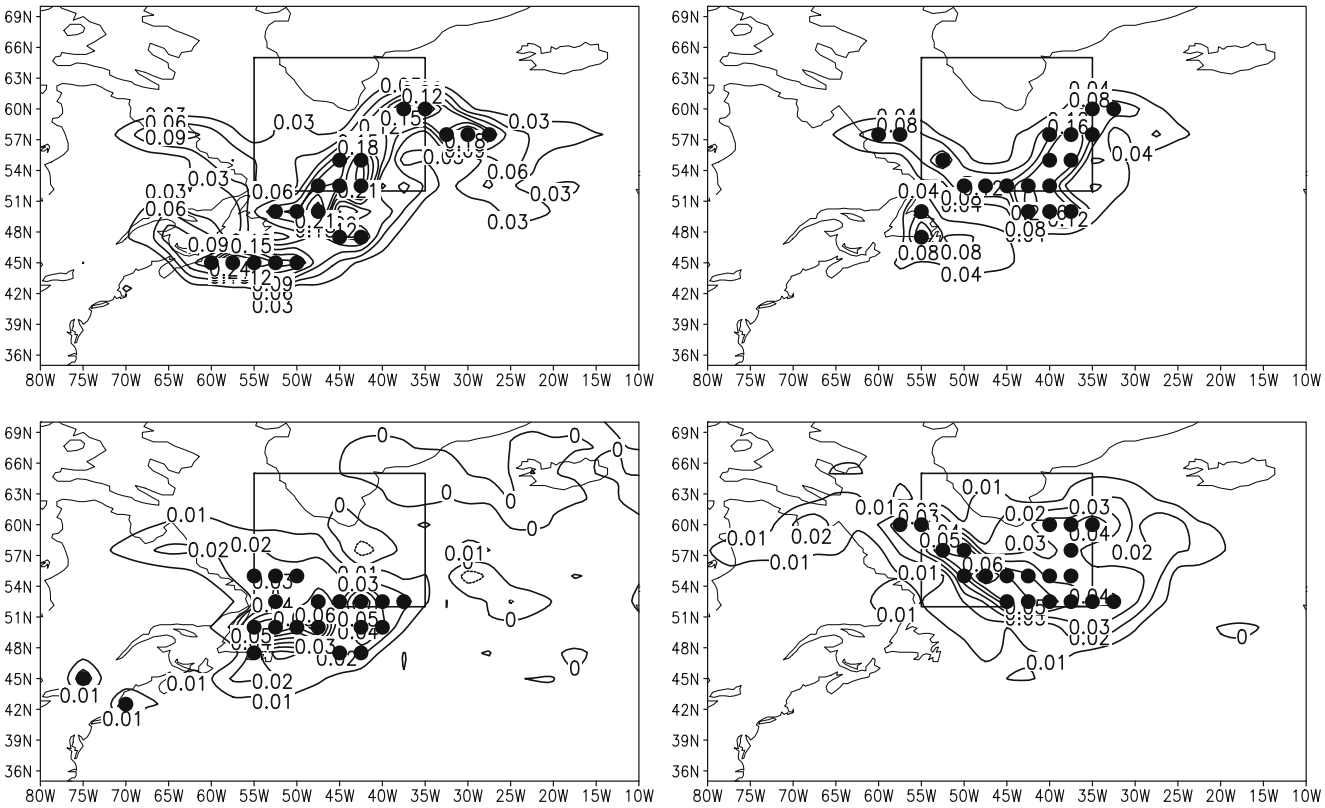

Fig. 5 Sensitivity fields with forecast lead time of 24 h. *Top figures*: sensitivity field (Eq. 18) at $t_i = t_0$ (left) and at $t_i = t_0 + 6$ h (*right*). *Bottom figures*: sensitivity field (Eq. 21) with $m = 10$ eigenvectors at $t_i = t_0$ (*left*) and at $t_i = t_0 + 6$ h (*right*). The verification domain $\mathcal{D}_v$ is the region within the *rectangle*

A 4D-Var data assimilation is performed using 20 adaptive observations at $t = 0$ and 6 h where the sensitivity field (Eq. 21) has the highest values, as marked in Fig. 5 (bottom). The performance of adaptive observations obtained from the SOA sensitivity function (Eq. 21) is compared with that obtained from the FOA sensitivity function (Eq. 18), as marked in Fig. 5 (top).

Grid-point meridional- and zonal-averaged forecast errors over the verification domain are shown in Fig. 6 for the forecast initialized from the background state estimate and from the analysis provided by each 4D-Var observation targeting experiment. It is noticed that assimilation of the adaptive observations resulted in a substantial forecast error reduction and that, in average, the SOA targeting guidance provided improved results as compared to the FOA methodology. Isopleths of the 24-hour forecast error of the assimilation of adaptive observations defined by the sensitivity fields

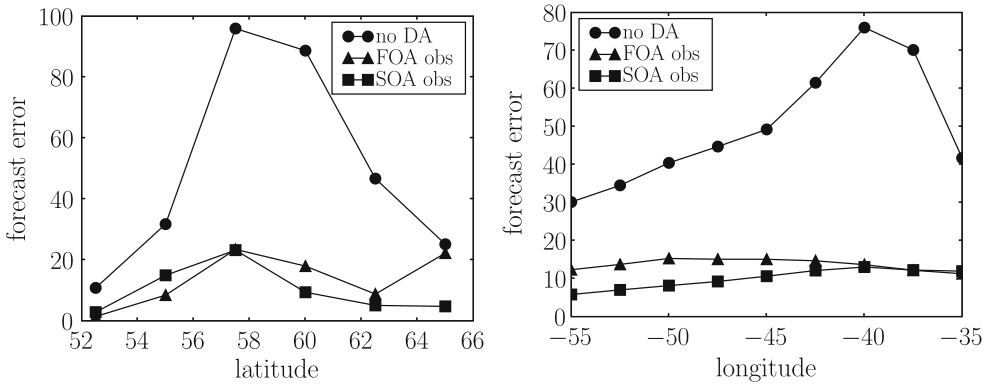

Fig. 6 Meridional- and zonal-averaged total energy forecast error ($m^2 s^{-2}$) at the verification time $t_v = 24$ h over the verification domain for the background forecast (*circles*), and the analysis forecast provided by data assimilation with adaptive observations at both $t = 0$ h and $t = 6$ h based on the FOA approach (*triangles*) and based on the eigenvectors of the Hessian (*squares*)

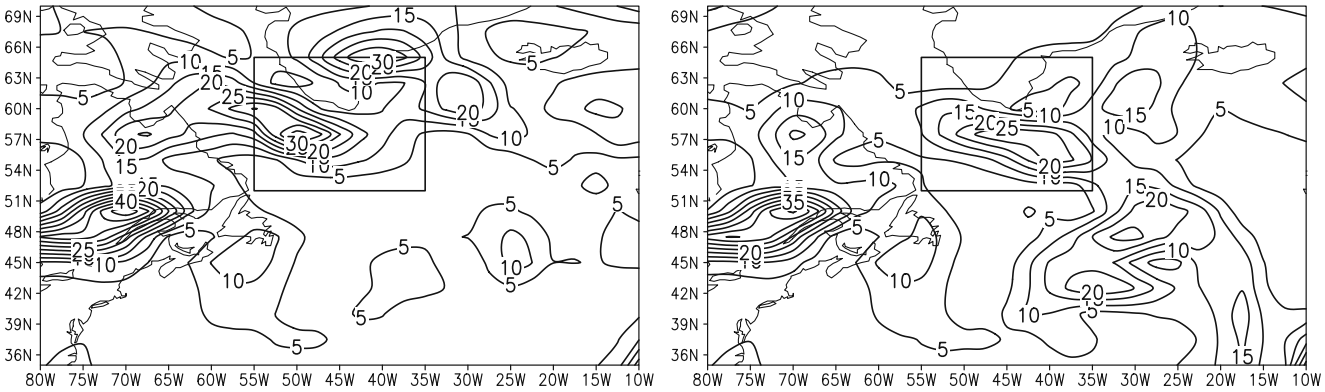

Fig. 7 Isopleths of the analysis forecast error at $t_v = 24$ h with adaptive observations selected at both $t = 0$ and $t = 6$ h from FOA (*left*), and from the Hessian eigenvectors (*right*). The contour interval is 5 $m^2 s^{-2}$

(Eq. 21) and (Eq. 18) are shown in Fig. 7. The spatial distribution of the forecast errors reveals that each set of targeted observations is particularly effective in reducing the errors over specific subregions of the verification domain: the FOA targeted observations were most effective in reducing the errors over the south-eastern region whereas the SOA targeted observations were most effective in reducing the errors over the north-western region.

The authors noticed that the insertion of targeted observations at $t = 6$ h, as opposed to only $t = 0$ h, is of significant benefit to the forecast. Results of data assimilation performed with targeted observations at the initial time only are shown in Fig. 8 and, by comparison with the results in Fig. 6, a large forecast impact is noticed from observations at $t = 6$ h. This is consistent to the observation sensitivity study in [9] where is was found that the forecast sensitivity to observations increases for observations near the end of the assimilation window, and thus closer to the verification time. In addition, accounting for data interaction is essential when multiple targeting instants are considered [10], and this is an area where further research is much needed. Nevertheless, the results show the relevance of the SOA in targeting strategies when dealing with large time intervals and/or highly nonlinear forecast models.

5.2.2 48-hour forecast lead time assimilation

In a second set of experiments the forecast lead time is increased to 48 h, while the rest of the experiment setup remains the same. Figure 9 shows the sensitivity fields obtained with FOA, for $t_i = t_0$ (top-left) and $t_i = t_0 + 6$ h (top-right), and SOA, for $t_i = t_0$ (bottom-left) and $t_i = t_0 + 6$ h (bottom-right). Even though the sensitivity fields are computed at the same targeting instants as in the previous experiment, the increased forecast lead time has altered the structure of the sensitivity fields and consequently, the location of the targeted observations. The modified structure of sensitivity fields reflects the dynamical dependence of the forecast on the initial condition error propagation through the model.

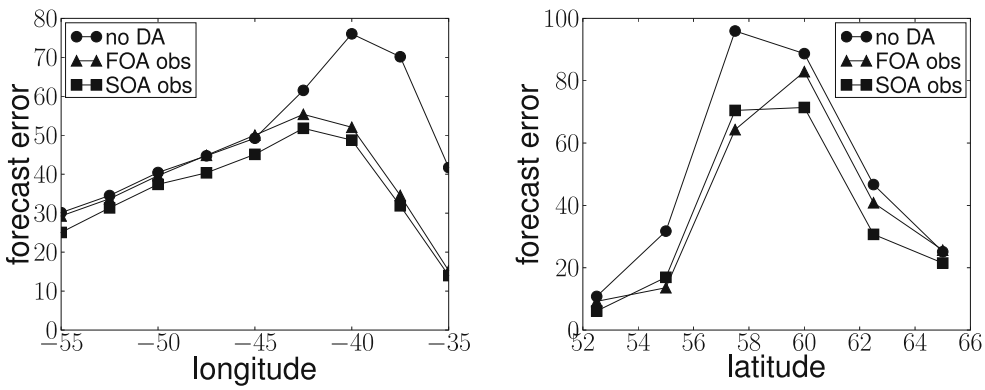

Fig. 8 Same as Fig. 6 for adaptive observations at $t = 0$ h only

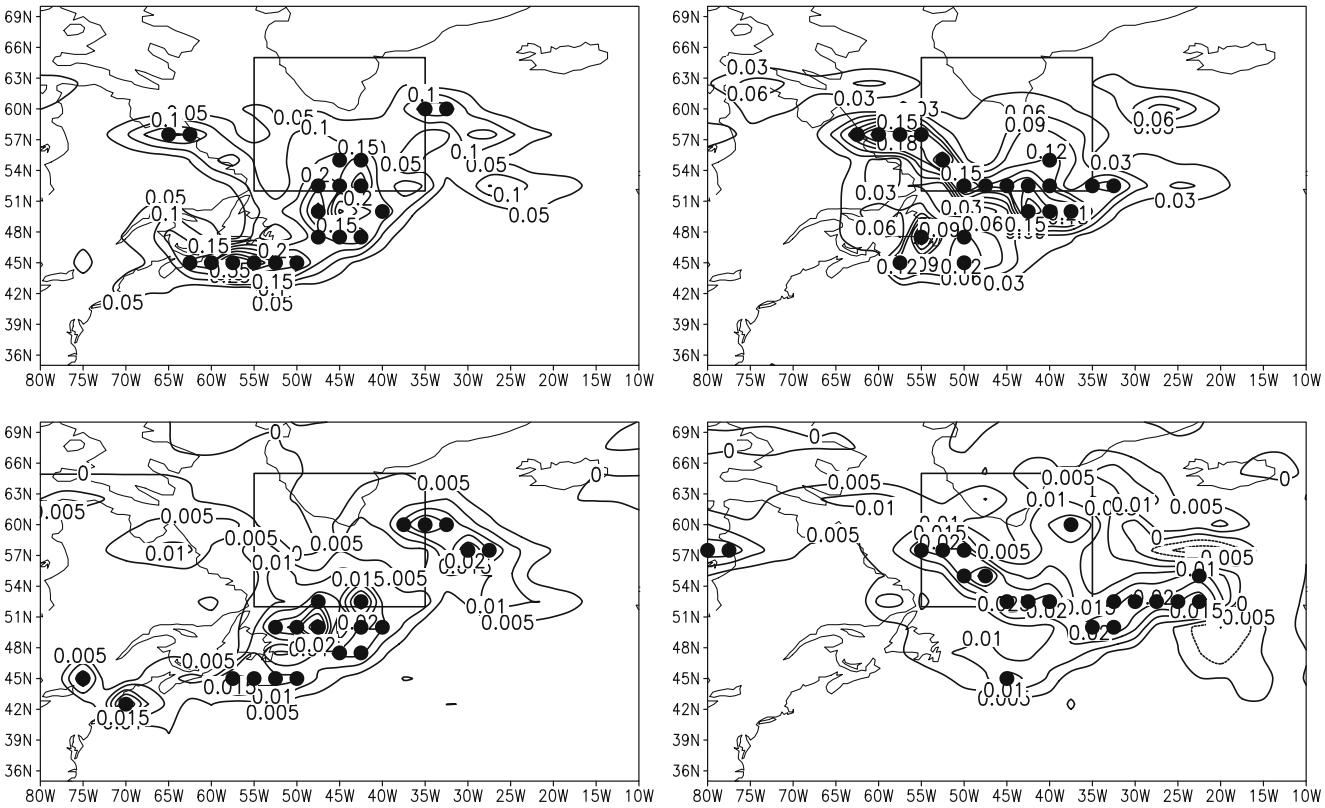

Fig. 9 Sensitivity fields with forecast lead time of 48 h. *Top figures*: sensitivity field (Eq. 18) at $t_i = t_0$ (*left*) and at $t_i = t_0 + 6$ h (*right*). *Bottom figures*: sensitivity field (Eq. 21) with $m = 10$ eigenvectors at $t_i = t_0$ (*left*) and at $t_i = t_0 + 6$ h (*right*). The verification domain $\mathcal{D}_v$ is the region within the *rectangle*

As in the previous experiments, 20 adaptive observations defined by the FOA and SOA sensitivity fields are assimilated at $t = 0$ h and $t = 6$ h with a verification lead time of 48 h. Figure 10 displays the meridional- and zonal-averaged forecast error at $t_v = t_0 + 48$ h over the verification domain. It is noticed that at a 48-hour forecast lead time the impact of the nonlinearities in the model is significant enough to impair the ability of the FOA targeting strategy to accurately track the evolution of initial condition errors. In this case the information provided by the SOA model, in the form of the sensitivity field defined by the eigenvalues and eigenvectors of the Hessian matrix, becomes particularly useful to track the evolution of errors in the initial model state and a significant reduction in the forecast error is noticed when the SOA targeted observations are assimilated as compared to the FOA targeted observations.

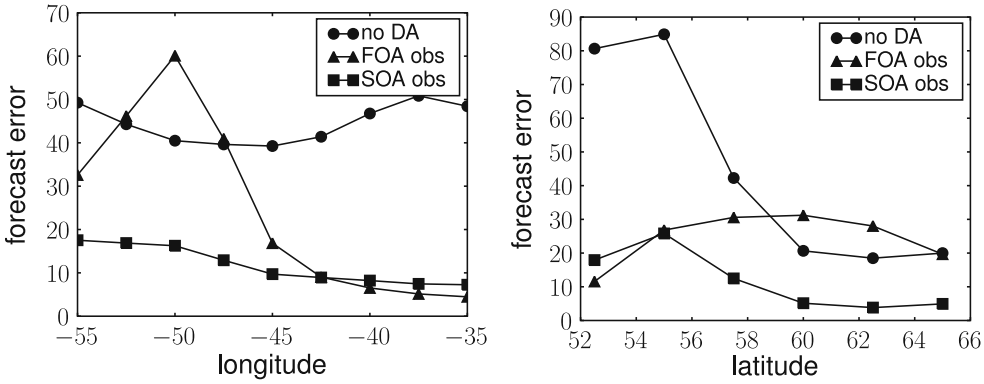

Fig. 10 Same as Fig. 6 with forecast lead time at 48 h

6 Conclusions

To date, properly accounting for the nonlinear growth of initial condition errors into the forecasts issued by a data assimilation system is an unresolved issue in the design of effective observation targeting strategies for numerical weather prediction. The practical applicability of adaptive observation methods based on first-order sensitivity information is hampered by a bottleneck in the experimental design: on the one hand the forecast lead time must be long enough to allow for preparation, observational resources deployment, data assimilation, and forecast updates; on the another hand the underlying assumption of linear error growth is valid only at short range, typically for a 24–48 h forecast lead time. In this study a novel approach based on second-order adjoint modeling is proposed to account for the quadratic initial-condition error growth in the model forecast. Using second-order adjoint information it is possible to account for quadratic error growth and maintain a more accurate approximation to the evolution of initial condition errors over longer time intervals. The dominant eigenvectors of the Hessian matrix associated with a specified forecast error aspect served as an indicator of the directions of maximal quadratic error growth. These vectors are a natural choice to be included in the targeting strategies given their mathematical properties. A discrete SOA model was implemented using automatic differentiation and coupled with the ARPACK software package to compute the Hessian eigenpairs. The eigenvectors associated with the largest eigenvalues were used to define a space-distributed sensitivity field function in order to identify the regions of maximal quadratic initial condition error growth. A comparative qualitative analysis between gradient sensitivity targeting strategy and our novel SOA targeting strategy, in the framework of 4D-Var, is performed using a two-dimensional global shallow-water model. The numerical experiments reveal that for an extended lead time the forecast error over the verification domain is significantly smaller when assimilating observations taken from the SOA targeting strategy as compared to a gradient sensitivity method. Preliminary results indicate that accounting for the quadratic error growth in the targeting strategy is of particular benefit as the nonlinear terms in the Taylor series approximation become dominant in the evolution of initial condition perturbations in the model.

The efficiency of the SOA targeting method remains to be further tested and validated in a priori observation targeting experiments. In this context, the use of an ensemble of model forecasts provides a feasible approach to define the forecast error functional and to initialize the adjoint fields for first- and second-order sensitivity computations. To fully exploit the benefits of the SOA model, novel targeting strategies must combine and properly balance the information from both FOA and SOA models to form a more cohesive and accurate strategy.

References

1. Ancell, B., Hakim, G.J.: Comparing adjoint-and ensemble-sensitivity analysis with applications to observation targeting. Mon. Weather Rev. **135**, 4117–4134 (2007)
2. Baker, N.L., Daley, R.: Observation and background adjoint sensitivity in the adaptive observation-targeting problem. Q. J. R. Meteorol. Soc. **126**, 1431–1454 (2000)
3. Bergot, T., Doerenbecher, A.: A study on the optimization of the deployment of targeted observations using adjoint-based methods. Q. J. R. Meteorol. Soc. **128**, 1689–1712 (2002)
4. Berliner, L.M., Lu, Z.Q., Snyder, C.: Statistical design for adaptive weather observations. J. Atmos. Sci. **56**, 2536–2552 (1999)
5. Bishop, C.H., Etherton, B.J., Majumdar, S.J.: Adaptive sampling with the ensemble transform Kalman filter. Part I: theoretical aspects. Mon. Weather Rev. **129**(3), 420–436 (2001)
6. Bishop, C.H., Toth, Z.: Ensemble transformation and adaptive observations. J. Atmos. Sci. **56**, 1748–1765 (1999)
7. Buizza, R., Cardinali, C., Kelly, G., Thépaut, J.N.: The value of observations. II: The value of observations located in singular-vector-based target areas. Q. J. R. Meteorol. Soc. **133**, 1817–1832 (2007)
8. Buizza, R., Montani, A.: Targeted observations using singular vectors. J. Atmos. Sci. **56**, 2965–2985 (1999)
9. Daescu, D.N.: On the Sensitivity equations of four-dimensional variational (4D-Var) data assimilation. Mon. Weather Rev. **136**, 3050–3065 (2008)
10. Daescu, D.N., Navon, I.M.: Adaptive observations in the context of 4D-Var data assimilation. Meteorol. Atmos. Phys. **85**, 205–226 (2004)
11. Daescu, D.N., Navon, I.M.: Efficiency of a POD-based reduced second-order adjoint model in 4D-Var data assimilation. Int. J. Numer. Methods Fluids **53**, 985–1004 (2007)
12. Daley, R.: Atmospheric Data Analysis. Cambridge University Press (1991)
13. Evensen, G.: Data Assimilation: The Ensemble Kalman Filter, 2nd edn. Springer (2009)
14. Gelaro, R., Langland, R.H., Rohaly, G.D., Rosmond, T.E.: As assessment of the singular-vector approach to targeted observing using the FASTEX dataset. Q. J. R. Meteorol. Soc. **125**, 3299–3327 (1999)
15. Giering, R.: Tangent linear and adjoint model compiler, users manual. Center for Global Change Sciences, Department of Earth, Atmospheric, and Planetary Science. MIT, Cambridge (1997)
16. Giering, R., Kaminski, T.: Recipes for adjoint code construction. Trans. Math. Software **24**, 437–474 (1998)
17. Griewank, A.: Evaluating derivatives: principles and techniques of algorithmic differentiation. Society for Industrial and Applied Mathematics, Philadelphia, PA (2000)
18. Hamill, T.M., Snyder, C.: Using improved background-error covariances from an ensemble Kalman filter for

adaptive observations. Mon. Weather Rev. **130**, 1552–1572 (2002)
19. Hamill, T.M., Snyder, C., Whitaker, J.S.: Ensemble forecasts and the properties of flow-dependent analysis-error covariance singular vectors. Mon. Weather Rev. **131**, 1741–1758 (2003)
20. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press (1985)
21. Kalnay, E.: Atmospheric Modeling, Data Assimilation, and Predictability. Cambridge University Press (2003)
22. Langland, R.H.: Issues in targeted observing. Q. J. R. Meteorol. Soc. **131**, 3409–3425 (2005)
23. Langland, R.H., Gelaro, R., Rohaly, G.D., Shapiro, M.A.: Targeted observations in FASTEX: Adjoint-based targeting procedures and data impact experiments in IOP17 and IOP18. Q. J. R. Meteorol. Soc. **125**, 3241–3270 (1999)
24. Lawrence, A.R., Leutbecher, M., Palmer, T.N.: The characteristics of Hessian singular vectors using an advanced data assimilation scheme. Q. J. R. Meteorol. Soc. **135**, 1117–1132 (2009)
25. Le Dimet, F.X., Navon, I.M., Daescu, D.N.: Second-order information in data assimilation. Mon. Weather Rev. **130**, 629–648 (2002)
26. Le Dimet, F.X., Talagrand, O.: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. Tellus **38A**(2), 97–110 (1986)
27. Lehoucq, R.B., Sorensen, D.C., Yang, C.: ARPACK Users' Guide: Solution of Large-scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods. Society for Industrial and Applied Mathematics Philadelphia, PA (1998)
28. Leutbecher, M., Palmer, T.N.: Ensemble forecasting. J. Comput. Phys. **227**, 3515–3539 (2008)
29. Lewis, J.M., Lakshmivarahan, S., Dhall, S.K.: Dynamic Data Assimilation: A Least Squares Approach. Cambridge University Press (2006)
30. Lin, S.J., Chao, W.C., Sud, Y.C., Walker, G.K.: A class of the van Leer-type transport schemes and its application to the moisture transport in a general circulation model. Mon. Weather Rev. **122**, 1575–1593 (1994)
31. Lin, S.J., Rood, R.B.: An explicit flux-form semi-Lagrangian shallow-water model on the sphere. Q. J. R. Meteorol. Soc. **123**, 2477–2498 (1997)
32. Liu, D.C., Nocedal, J.: On the limited memory BFGS method for large scale optimization. Math. Program **45**, 503–528 (1989)
33. Majumdar, S.J., Aberson, S.D., Bishop, C.H., Buizza, R., Peng, M.S., Reynolds, C.A.: A comparison of adaptive observing guidance for Atlantic tropical cyclones. Mon. Weather Rev. **134**, 2354–2372 (2006)
34. Neta, B., Giraldo, F.X., Navon, I.M.: Analysis of the Turkel-Zwas scheme for the two-dimensional shallow water equations in spherical coordinates. J. Comput. Physics **133**, 102–112 (1997)
35. Palmer, T.N., Gelaro, R., Barkmeijer, J., Buizza, R.: Singular Vectors, Metrics, and Adaptive Observations. J. Atmos. Sci. **55**, 633–653 (1998)
36. Rabier, F., Klinker, E., Courtier, P., Hollingsworth, A.: Sensitivity of forecast errors to initial conditions. Q. J. R. Meteorol. Soc. **122**, 121–150 (1996)
37. Torn, R.D., Hakim, G.J.: Ensemble-based sensitivity analysis. Mon. Weather Rev. **136**, 663–677 (2008)
38. Xu, L., Daley, R.: Data Assimilation with a barotropically unstable shallow water system using representer algorithms. Tellus **54A**, 125–137 (2002)