NOTES AND CORRESPONDENCE

On the Deterministic Observation Impact Guidance: A Geometrical Perspective

DACIAN N. DAESCU

Portland State University, Portland, Oregon

(Manuscript received 20 January 2009, in final form 26 April 2009)

ABSTRACT

An optimal use of the atmospheric data in numerical weather prediction requires an objective assessment of the value added by observations to improve the analyses and forecasts of a specific data assimilation system (DAS). This research brings forward the issue of uncertainties in the assessment of observation values based on deterministic observation impact (OBSI) estimations using observing system experiments (OSEs) and the adjoint-DAS framework. The state-to-observation space uncertainty propagation as a result of the errors in the verification state is investigated. For a quadratic forecast error measure, a geometrical perspective is used to provide insight and to convey some of the key aspects of this research. The study is specialized to a DAS implementing a linear analysis scheme and numerical experiments are presented using the Lorenz 40-variable model.

1. Introduction

Development of efficient methodologies to quantify the observation impact (OBSI) on the analyses and forecasts of a specific data assimilation system (DAS) is the focus of research at numerical weather prediction centers worldwide. An objective assessment of the observation value is required to achieve an optimal use of the data from in situ platforms and satellites and to design intelligent data-thinning procedures, cost-effective field experiments for targeted observations, and future observing system networks.

Traditionally (Atlas 1997), the OBSI estimation is performed through observing system experiments (OSEs) where selected datasets are systematically added to or removed from the control DAS to obtain the experiment state. The OBSI on a forecast aspect of interest is then quantified by comparison to the forecast issued from the control DAS. The OSEs' methodology requires a modest software development, once the DAS is in place, and has been used to assess the impact of various observing system components and the value of targeted observations (Andersson et al. 1991; Bouttier and Kelly 2001; Kelly et al. 2007). Since an additional assimilation experiment is required for each new dataset being evaluated, in practice only a few observing system components may be evaluated through OSEs.

Observation sensitivity techniques were initially introduced in numerical weather prediction for applications to targeted observations (Baker and Daley 2000; Doerenbecher and Bergot 2001) and adjoint-DAS techniques are currently implemented as an effective approach (all at once) to estimate the impact of *any* data subset in the DAS on a specified forecast aspect and to monitor the observation performance on short-range forecasts (Fourrié et al. 2002; Cardinali et al. 2004; Langland and Baker 2004; Langland 2005; Errico 2007; Gelaro et al. 2007; Zhu and Gelaro 2008; Trémolet 2007, 2008; Daescu and Todling 2009). These studies revealed several practical issues in the adjoint-based OBSI estimates such as the difficulty to obtain an adjoint DAS fully consistent to the analysis scheme and the need for high-order approximation measures. A particular source of uncertainty in the OBSI estimates is due to inconsistencies between the adjoint model implementation and the nonlinear model forecast (e.g., moist physical processes are not properly incorporated in the adjoint model; Ancell and Mass 2008). Nonlinearities in the DAS analysis scheme and/or in the forecast aspect

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

DOI: 10.1175/2009MWR2954.1

measure and approximation errors may contribute to further ambiguities in the OBSI interpretations. A significant software development effort is required to implement the adjoint DAS; however, this approach provides detailed OBSI information that would be difficult to be obtained by other means in a variational DAS. Comparative studies of observation impacts derived from OSEs and adjoint-DAS techniques are provided in the work of Gelaro and Zhu (2009) and Cardinali (2009). The consensus is that adjoint-based techniques are effective in measuring the response in a functional aspect of a short-range forecast due to any perturbation in the observing system, whereas OSEs are effective in measuring the impact of a single data component on any forecast aspect. Forecasts sensitivity to observations and OBSI estimation in an ensemble Kalman filter is discussed by Liu and Kalnay (2008) and Torn and Hakim (2008).

This research note brings forward an issue that has not been properly addressed in previous studies; namely, the uncertainty in the OBSI guidance derived from deterministic forecast error measures using OSEs and adjoint-DAS techniques. The OSEs' framework and the adjoint-DAS approach to observation impact estimation in a DAS implementing a linear analysis scheme are briefly reviewed in section 2. The state-to-observation space uncertainty propagation in the OBSI assessment as a result of the errors in the verification state is discussed in section 3. For a quadratic forecast error measure, a geometrical perspective is used to provide insight and to convey some of the key aspects of this research. Illustrative numerical experiments are presented in section 4 using the Lorenz 40-variable model (Lorenz and Emanuel 1998). Concluding remarks are in section 5.

2. Deterministic OBSI estimation

The deterministic approach to observation impact estimation using OSEs and the adjoint-DAS methodology is briefly reviewed in this section. For simplicity, consider a linear analysis scheme such as a Kalman filter and variational methods implementing a single outer loop iteration (Daley 1991; Kalnay 2002):

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}[\mathbf{y} - \mathbf{h}(\mathbf{x}_b)], \tag{1}$$

where $\mathbf{x}_b$ is a background estimate to the initial conditions, $\mathbf{y}$ is the vector of observational data, $\mathbf{h}$ is the observation operator, and

$$\mathbf{K} = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1} \tag{2}$$

is the optimal gain matrix expressed in terms of the background error covariance matrix $\mathbf{B}$, the observation error covariance matrix $\mathbf{R}$, and the linearized observation operator $\mathbf{H}$.

In the deterministic framework the OBSI calculations are performed for a specific forecast aspect and a typical scalar measure of the error in a forecast initiated from $\mathbf{x}$ is defined as

$$e_v(\mathbf{x}) = (\mathbf{x}^f - \mathbf{x}^v)^{\mathrm{T}}\mathbf{C}(\mathbf{x}^f - \mathbf{x}^v) = \|\mathbf{x}^f - \mathbf{x}^v\|_{\mathbf{C}}^2, \tag{3}$$

where $\mathbf{x}^f = \mathcal{M}(\mathbf{x})$ is the nonlinear model forecast at time t initiated at $t_0 < t$ from $\mathbf{x}$, $\mathbf{x}^v$ is the verifying analysis at time t, the superscript T denotes the transpose operator, and $\mathbf{C}$ is a symmetric and positive definite matrix that defines the metric on the state space (e.g., an appropriate energy norm). The subscript e_v is used to emphasize the dependence of the measure in (3) on the verification state $\mathbf{x}^v$. It is noticed that the right side of (3) is quadratic in terms of $\mathbf{x}^f$, however, the nonlinearity of $e_v(\mathbf{x})$ as a function of the initial conditions is of a more general type because of the nonlinearities in the model forecast.

a. OSE's framework

Let $\mathbf{y}$ denote the set of all observations assimilated in the control DAS in (1)–(2), $\mathbf{y}_i \subset \mathbf{y}$ denote the data subset whose impact is being evaluated, and $\overline{\mathbf{y}}_i = \mathbf{y}\backslash\mathbf{y}_i$ denote the complement set of $\mathbf{y}_i$ with respect to $\mathbf{y}$ (i.e., the set of observations in the DAS after removing the data component $\mathbf{y}_i$). In the OSEs' data denial framework an *additional* assimilation is performed using only data $\overline{\mathbf{y}}_i$ to obtain the *experiment* analysis state $\overline{\mathbf{x}}_{a,i}$:

$$\overline{\mathbf{x}}_{a,i} = \overline{\mathbf{x}}_{b,i} + \overline{\mathbf{K}}_i[\overline{\mathbf{y}}_i - \overline{\mathbf{h}}_i(\overline{\mathbf{x}}_{b,i})]. \tag{4}$$

Equation (4) accounts for the change in the observation-related input from ($\mathbf{y}$, $\mathbf{R}$, $\mathbf{h}$) in the control DAS to ($\overline{\mathbf{y}}_i$, $\overline{\mathbf{R}}_i$, $\overline{\mathbf{h}}_i$) in the OSEs, as well as for the modified background-related input from ($\mathbf{x}_b$, $\mathbf{B}$) in the control DAS to ($\overline{\mathbf{x}}_{b,i}$, $\overline{\mathbf{B}}_i$) in the OSEs. The latter must be considered when several data assimilation cycles are performed to assess the impact of data removal. Once $\overline{\mathbf{x}}_{a,i}$ is available, the impact of the selected data subset $\mathbf{y}_i$ within the control DAS on any forecast aspect of interest e_v may be easily evaluated by

$$I_v^{\mathrm{ose}}(\mathbf{y}_i) \overset{\mathrm{def}}{=} e_v(\mathbf{x}_a) - e_v(\overline{\mathbf{x}}_{a,i}), \tag{5}$$

such that $I_v^{\mathrm{ose}}(\mathbf{y}_i)$ measures the value added by the data $\mathbf{y}_i$ to the observing system $\overline{\mathbf{y}}_i$.

b. Adjoint-DAS framework

Current adjoint-based OBSI techniques are based on an observation-space estimation of the forecast impact due to assimilation of all data in the DAS:

$$\delta e_v = e_v(\mathbf{x}_a) - e_v(\mathbf{x}_b) \approx (\delta \mathbf{x}_a)^{\mathrm{T}} \mathbf{g}_v = (\delta \mathbf{y})^{\mathrm{T}} \mathbf{K}^{\mathrm{T}} \mathbf{g}_v, \quad (6)$$

where $\delta \mathbf{y} = \mathbf{y} - \mathbf{h}(\mathbf{x}_b)$ is the innovation vector, $\delta \mathbf{x}_a = \mathbf{K}\delta \mathbf{y}$ is the analysis increment in (1), and $\mathbf{g}_v$ is a properly defined vector measuring the forecast sensitivity to initial conditions (Gelaro et al. 2007). Associated to the measure in (3), the gradient to initial conditions is expressed as

$$\nabla_{\mathbf{x}} e_v(\mathbf{x}) = 2\mathbf{M}^{\mathrm{T}}(\mathbf{x})\mathbf{C}(\mathbf{x}^f - \mathbf{x}^v), \quad (7)$$

where $\mathbf{M}^{\mathrm{T}}(\mathbf{x})$ is the adjoint of the tangent linear model from t_0 to t evaluated along the model trajectory initialized from $\mathbf{x}$. For example, the observation impact methodology introduced by Langland and Baker (2004) estimates the variation δe_v using sensitivity gradients along both background and analysis trajectories:

$$\mathbf{g}_v = \frac{1}{2}\nabla_{\mathbf{x}} e_v(\mathbf{x}_a) + \frac{1}{2}\nabla_{\mathbf{x}} e_v(\mathbf{x}_b), \quad (8)$$

which provides a second-order accurate δe_v approximation (Errico 2007; Daescu and Todling 2009). In practical applications the approximation in (6)–(8) was found to provide improved results as compared with first-order approximations (Gelaro et al. 2007; Gelaro and Zhu 2009; Cardinali 2009). A measure of the contribution of *any* data subset $\mathbf{y}_i$ in the DAS to the forecast error reduction is obtained by taking the inner product between the innovation vector component $\delta \mathbf{y}_i$ and the corresponding amplification factor in (6):

$$I_v^{\mathrm{adj}}(\mathbf{y}_i) \stackrel{\mathrm{def}}{=} (\delta \mathbf{y}_i)^{\mathrm{T}} (\mathbf{K}^{\mathrm{T}} \mathbf{g}_v)_i. \quad (9)$$

Data components for which $I_v^{\mathrm{adj}}(\mathbf{y}_i) < 0$ contribute to the forecast error reduction (improve the forecast), whereas data components with $I_v^{\mathrm{adj}}(\mathbf{y}_i) > 0$ will increase the forecast error (degrade the forecast).

3. Uncertainty analysis

In practice, the verification state $\mathbf{x}^v$ used to define the forecast error measure in (3) is obtained after several assimilation cycles to the verification time are performed and represents only *an estimation* of the true atmospheric state $\mathbf{x}^t$ at t. As such, the assessment of the impact of various data types in the DAS to reduce the actual forecast error must account for the state-to-observation space uncertainty propagation based on the forecast error measure that involves the unknown true atmospheric state $\mathbf{x}^t$:

$$e_t(\mathbf{x}) = (\mathbf{x}^f - \mathbf{x}^t)^{\mathrm{T}} \mathbf{C} (\mathbf{x}^f - \mathbf{x}^t) = \left\| \mathbf{x}^f - \mathbf{x}^t \right\|_{\mathbf{C}}^2. \quad (10)$$

To account for the errors in the verification state, let n denote the dimension of the state vector and

$$\epsilon = \| \mathbf{x}^v - \mathbf{x}^t \| \quad (11)$$

denote a measure of the precision (accuracy) of $\mathbf{x}^v$. The $(n-1)$ sphere of center $\mathbf{x}^t$ and radius ϵ

$$\mathcal{S}(\mathbf{x}^t, \epsilon) = \{ \mathbf{x}^v \in \mathbb{R}^n : \| \mathbf{x}^v - \mathbf{x}^t \| = \epsilon \} \quad (12)$$

is the set of verification states of ϵ precision.

a. OBSI uncertainty propagation in OSEs

In the deterministic OSEs' framework, the uncertainty in the OBSI estimates is expressed as

$$\delta I_{t,v}^{\mathrm{ose}}(\mathbf{y}_i) = [e_t(\mathbf{x}_a) - e_t(\overline{\mathbf{x}}_{a,i})] - [e_v(\mathbf{x}_a) - e_v(\overline{\mathbf{x}}_{a,i})]. \quad (13)$$

From (3) and (10) it follows that

$$\begin{aligned} e_v(\mathbf{x}_a) - e_v(\overline{\mathbf{x}}_{a,i}) &= \| \mathbf{x}_a^f - \mathbf{x}^v \|_{\mathbf{C}}^2 - \| \overline{\mathbf{x}}_{a,i}^f - \mathbf{x}^v \|_{\mathbf{C}}^2 \\ &= \langle \mathbf{x}_a^f - \overline{\mathbf{x}}_{a,i}^f, \mathbf{x}_a^f + \overline{\mathbf{x}}_{a,i}^f - 2\mathbf{x}^v \rangle_{\mathbf{C}} \end{aligned} \quad (14)$$

and

$$\begin{aligned} e_t(\mathbf{x}_a) - e_t(\overline{\mathbf{x}}_{a,i}) &= \| \mathbf{x}_a^f - \mathbf{x}^t \|_{\mathbf{C}}^2 - \| \overline{\mathbf{x}}_{a,i}^f - \mathbf{x}^t \|_{\mathbf{C}}^2 \\ &= \langle \mathbf{x}_a^f - \overline{\mathbf{x}}_{a,i}^f, \mathbf{x}_a^f + \overline{\mathbf{x}}_{a,i}^f - 2\mathbf{x}^t \rangle_{\mathbf{C}}, \end{aligned} \quad (15)$$

respectively, where $\langle \cdot, \cdot \rangle_{\mathbf{C}}$ denotes the inner product in the state space, $\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{C}} = \mathbf{u}^{\mathrm{T}} \mathbf{C} \mathbf{v}$. After replacing (15) and (14) in (13):

$$\begin{aligned} \delta I_{t,v}^{\mathrm{ose}}(\mathbf{y}_i) &= 2(\mathbf{x}_a^f - \overline{\mathbf{x}}_{a,i}^f)^{\mathrm{T}} \mathbf{C} (\mathbf{x}^v - \mathbf{x}^t) \\ &= 2 \| \mathbf{x}_a^f - \overline{\mathbf{x}}_{a,i}^f \|_{\mathbf{C}} \| \mathbf{x}^v - \mathbf{x}^t \|_{\mathbf{C}} \cos\theta_i, \end{aligned} \quad (16)$$

where θ_i is the angle between the vectors $\mathbf{x}_a^f - \overline{\mathbf{x}}_{a,i}^f$ and $\mathbf{x}^v - \mathbf{x}^t$. Equation (16) shows that in the OSEs' framework the uncertainty in the OBSI estimation is determined by the magnitude of the difference between the control forecast and the experiment forecast and by the verification state error component along the direction $\mathbf{x}_a^f - \overline{\mathbf{x}}_{a,i}^f$. The uncertainty propagation is thus specific to the dataset being evaluated. The hyperplane of $\mathbb{R}^n$ passing through $\mathbf{x}^t$ and normal to the vector $\mathbf{x}_a^f - \overline{\mathbf{x}}_{a,i}^f$:

$$\mathcal{H}_i = \{ \mathbf{x}^v : \langle \mathbf{x}_a^f - \overline{\mathbf{x}}_{a,i}^f, \mathbf{x}^v - \mathbf{x}^t \rangle_{\mathbf{C}} = 0 \} \quad (17)$$

identifies the set of verification states that are objective in estimating the OBSI of the data component $\mathbf{y}_i$, $\delta I_{t,v}^{\mathrm{ose}}(\mathbf{y}_i) = 0$. A geometrical illustration is provided in Fig. 1. It is also noticed that in a deterministic approach verification states of equal precision may result in an ambiguous OBSI guidance; for example, Fig. 1 shows a configuration of equal-precision states $\mathbf{x}^v, \mathbf{x}^w \in \mathcal{S}(\mathbf{x}^t, \epsilon)$ such that $I_v^{\mathrm{ose}}(\mathbf{y}_i) < 0$ and $I_w^{\mathrm{ose}}(\mathbf{y}_i) > 0$.

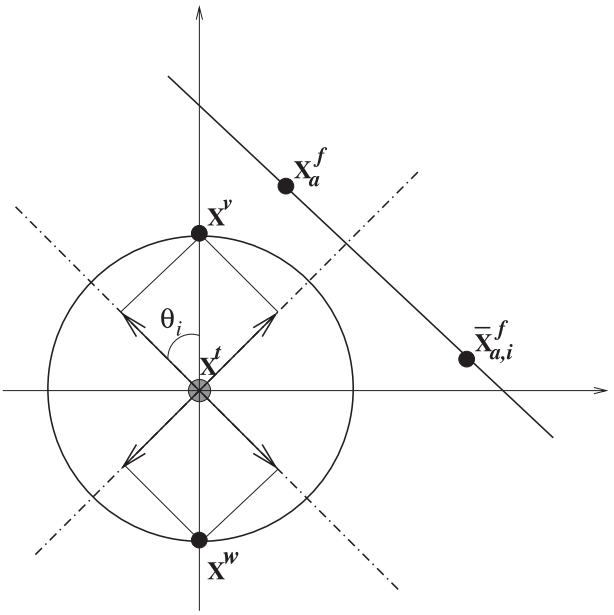

FIG. 1. Illustration of the uncertainty in the observation impact estimation using deterministic OSEs. The verification state error component along the direction $\mathbf{x}_a^f - \bar{\mathbf{x}}_{a,i}^f$ contributes to the uncertainty propagation. Verification states of equal-precision $\mathbf{x}^v$ and $\mathbf{x}^w$ may provide an ambiguous observation impact guidance.

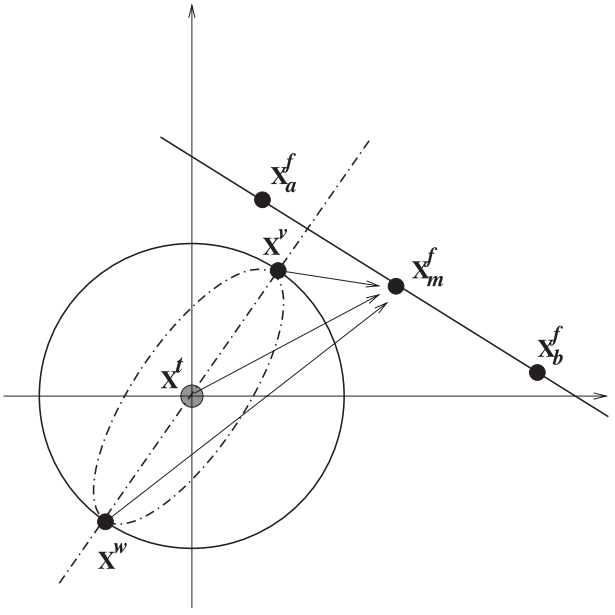

FIG. 2. Illustration of the uncertainty in the observation impact estimation in the deterministic adjoint-DAS framework. Verification states of equal-precision and equal-overall data impact may provide an ambiguous observation impact guidance through the vector $\mathbf{x}_m^f - \mathbf{x}^v$.

In a probabilistic framework, by expressing $\mathbf{x}_a^f - \bar{\mathbf{x}}_{a,i}^f = (\mathbf{x}_a^f - \mathbf{x}^t) - (\bar{\mathbf{x}}_{a,i}^f - \mathbf{x}^t)$ in (17), it is noticed that an objective OBSI assessment should be based on verification states whose errors are unbiased and are statistically independent from the errors in the forecasts produced by the DAS.

b. *OBSI uncertainty propagation in adjoint DAS*

In the deterministic adjoint-DAS framework the uncertainty in the verification state is propagated into the OBSI estimates of all the data subsets being evaluated. The uncertainty in the OBSI estimation (9) of a data subset $\mathbf{y}_i$ is expressed in a general form as

$$\delta I_{t,v}^{\mathrm{adj}}(\mathbf{y}_i) = I_t^{\mathrm{adj}}(\mathbf{y}_i) - I_v^{\mathrm{adj}}(\mathbf{y}_i) = (\delta \mathbf{y}_i)^{\mathrm{T}}[\mathbf{K}^{\mathrm{T}}(\mathbf{g}_t - \mathbf{g}_v)]_i. \quad (18)$$

For example, associated with the measure in (7)–(8):

$$\delta I_{t,v}^{\mathrm{adj}}(\mathbf{y}_i) = (\delta \mathbf{y}_i)^{\mathrm{T}}[\mathbf{K}^{\mathrm{T}}(\mathbf{M}_a^{\mathrm{T}} + \mathbf{M}_b^{\mathrm{T}})\mathbf{C}(\mathbf{x}^v - \mathbf{x}^t)]_i, \quad (19)$$

which shows that, accounting for the errors in $\mathbf{x}^v$ only, the uncertainty in the OBSI guidance is a result of the backward propagation of the errors in the verification state through the adjoint model, the adjoint-DAS operator, and amplified by the innovation vector components.

To provide a geometrical perspective, it is noticed that by analogy to the derivation in (13)–(17) in the adjoint-DAS approach:

$$\delta e_t - \delta e_v = 2(\mathbf{x}_a^f - \mathbf{x}_b^f)^{\mathrm{T}}\mathbf{C}(\mathbf{x}^v - \mathbf{x}^t) \quad (20)$$

such that the hyperplane:

$$\mathcal{H}_{\delta e_t} = \{\mathbf{x}^v : \langle \mathbf{x}_a^f - \mathbf{x}_b^f, \mathbf{x}^v - \mathbf{x}^t \rangle_{\mathbf{C}} = 0\} \quad (21)$$

identifies the set of verification states of equal-overall data impact:

$$\delta e_v = \delta e_t. \quad (22)$$

In the deterministic adjoint-DAS approach verification states of equal-precision (ϵ) and equal-overall data impact (δe_t) located on the $(n-2)$ sphere of $\mathbb{R}^n$:

$$\begin{aligned}\mathcal{S}_{\epsilon,\delta e_t} &= \mathcal{S}(\mathbf{x}^t, \epsilon) \cap \mathcal{H}_{\delta e_t} \\ &= \{\mathbf{x}^v \in \mathbb{R}^n : \|\mathbf{x}^v - \mathbf{x}^t\| = \epsilon, \langle \mathbf{x}_a^f - \mathbf{x}_b^f, \mathbf{x}^v - \mathbf{x}^t \rangle_{\mathbf{C}} = 0\}\end{aligned} \quad (23)$$

will lead to OBSI estimates that will vary in magnitude and that for certain data subsets may also vary in sign. In general, as illustrated in Fig. 2, $\mathcal{S}_{\epsilon,\delta e_t}$ consists of two points for $n = 2$, is a circle for $n = 3$, and so on. For simplicity, consider the case of linear dynamics $\mathbf{x}^f = \mathbf{M}\mathbf{x}$ where $\mathbf{M}$ is a state-independent matrix operator. In this framework, the measure in (3) is a quadratic function of the initial conditions and the approximation in (6)–(8) is thus

exact. Theoretically, for quadratic measures all second-order adjoint-based OBSI estimates are equivalent (Daescu and Todling 2009). The vector $\mathbf{g}_v$ defined by (7)–(8) is expressed as

$$\boldsymbol{g}_v = \mathbf{M}^{\mathrm{T}}\mathbf{C}(\mathbf{M}\mathbf{x}_a + \mathbf{M}\mathbf{x}_b - 2\mathbf{x}^v) = 2\mathbf{M}^{\mathrm{T}}\mathbf{C}(\mathbf{x}_m^f - \mathbf{x}^v), \quad (24)$$

where

$$\mathbf{x}_m^f = \frac{1}{2}(\mathbf{M}\mathbf{x}_a + \mathbf{M}\mathbf{x}_b) = \frac{1}{2}(\mathbf{x}_a^f + \mathbf{x}_b^f) \quad (25)$$

denotes the midpoint of the segment from $\mathbf{x}_b^f$ to $\mathbf{x}_a^f$. For each verification state $\mathbf{x}^v \in \mathcal{S}_{\epsilon,\delta e_t}$ the vector $\mathbf{x}_m^f - \mathbf{x}^v$ in (24) will vary in magnitude and orientation for example, Fig. 2 illustrates a configuration where verification states of equal-precision and equal-overall data impact may provide an ambiguous OBSI guidance in (9) since in this case

$$(x_{m,2}^f - x_2^v)(x_{m,2}^f - x_2^w) < 0.$$

4. Numerical experiments

The Lorenz 40-variable model (Lorenz and Emanuel 1998)

$$\frac{dx_j}{dt} = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F, \quad \text{where} \quad j = 1, 2, \ldots, n, \quad (26)$$

$n = 40$, $x_{-1} = x_{n-1}$, $x_0 = x_n$, and $x_{n+1} = x_1$, is used in OSEs and the adjoint-DAS approach to investigate the uncertainty in the OBSI estimation as a result of the errors in the verification state. For this model, Liu and Kalnay (2008) provided a comparative analysis of the OBSI derived from the adjoint-DAS method in (6)–(8) and an ensemble sensitivity method. The system in (26) is integrated with a fourth-order Runge–Kutta method and a constant time step $\Delta t = 0.05$ that in the data assimilation experiments is identified to a 6-h time period. The time evolution of the true state $\mathbf{x}^t$ is obtained by taking the external forcing to be $F = 8$ and a model error is considered in the forecast model by taking $F = 7.6$. An initial state $\mathbf{x}_0^t$ is obtained by a 90-day (360 time step) integration started from $x_j = 8$ for $j \neq n/2$ and $x_{n/2} = 8.008$; a background estimate $\mathbf{x}_b$ to $\mathbf{x}_0^t$ is prescribed by introducing random perturbations in $\mathbf{x}_0^t$ taken from the standard normal distribution N (0, 1). Observational data is generated for each state component (at each grid point) and at each time step, such that the observation operator is the identity. To investigate the uncertainty in the OBSI estimation when various data types are included in the DAS, it is assumed that the observing system provides four data types $\mathbf{y}^{(1)}$, $\mathbf{y}^{(2)}$, $\mathbf{y}^{(3)}$, and $\mathbf{y}^{(4)}$, each being a 10-dimensional vector, and that data type $\mathbf{y}^{(i)}$ is taken at locations $4\kappa + i$, $\kappa = 0, 1, \ldots, 9$. Componentwise, the structure of the observation vector is thus $\mathbf{y} = [y_1^{(1)} y_1^{(2)} y_1^{(3)} y_1^{(4)} y_2^{(1)} y_2^{(2)} y_2^{(3)} y_2^{(4)} \ldots y_{10}^{(1)} y_{10}^{(2)} y_{10}^{(3)} y_{10}^{(4)}]^{\mathrm{T}}$. The observation errors are normally distributed $N[0, \sigma^{(i)}]$ with the standard deviation prescribed as $\sigma^{(1)} = 0.1$, $\sigma^{(2)} = 0.2$, $\sigma^{(3)} = 0.4$, and $\sigma^{(4)} = 0.8$. The DAS implements an extended Kalman filter algorithm that provides the background estimate at time t_{i+1} and the associated background error covariance matrix according to

$$\mathbf{x}_b(t_{i+1}) = \mathcal{M}_{t_i \to t_{i+1}}[\mathbf{x}_a(t_i)] \quad \text{and} \quad (27)$$

$$\mathbf{B}(t_{i+1}) = \mathbf{M}(t_i)\mathbf{A}(t_i)\mathbf{M}^{\mathrm{T}}(t_i) + \mathbf{Q}(t_i), \quad (28)$$

where $\mathbf{x}_a(t_i)$ is the analysis (1) at time t_i,

$$\mathbf{A}(t_i) = [\mathbf{I} - \mathbf{K}(t_i)\mathbf{H}(t_i)]\mathbf{B}(t_i) \quad (29)$$

is the analysis error covariance matrix, and $\mathbf{M}(t_i)$ is the state-dependent Jacobian matrix of the numerical model from t_i to t_{i+1}. The model error covariance matrix $\mathbf{Q}$ is taken to be time invariant, diagonal and with constant entries, and by trial and error the specification of the diagonal entries $Q_{jj} = 0.01$ was found to provide improved analyses as compared to other selections. Numerical experiments are set with two data assimilation systems that differ only in the specification of the observation error covariance $\mathbf{R}$: in a first set of experiments, hereinafter referred to as DAS-I, $\mathbf{R}$ is taken to be diagonal and with entries R_{jj} that are statistically consistent with the observation errors in each data type, as described above; in a second set of experiments, hereafter referred to as DAS-II, all the diagonal entries are taken $\mathbf{R}_{jj} = 0.04$, which corresponds to an observation error standard deviation of 0.2. The DAS-I attempts to simulate an optimal DAS, whereas in the DAS-II the errors in data type $\mathbf{y}^{(1)}$ are overestimated and the errors in data types $\mathbf{y}^{(3)}$ and $\mathbf{y}^{(4)}$ are underestimated. Each DAS configuration is run for 7560 analysis cycles and time-averaged estimates of the OBSI on the 24-h forecast error are collected using OSEs and the adjoint-DAS approach over the last $N = 7200$ analysis cycles (a 5-yr time period). The time-averaged relative error in the analyses over this period was found to be of ~2.6% in the DAS-I and of ~5.1% in the DAS-II. Therefore, the relative error in the verification states $\mathbf{x}^v$ produced by the DAS is about 2 times larger in the DAS-II as compared to the DAS-I. OSEs are set by removing a single data point from the observing system, for each data location, for a total of n additional data assimilation experiments. The adjoint-DAS approach to OBSI estimation

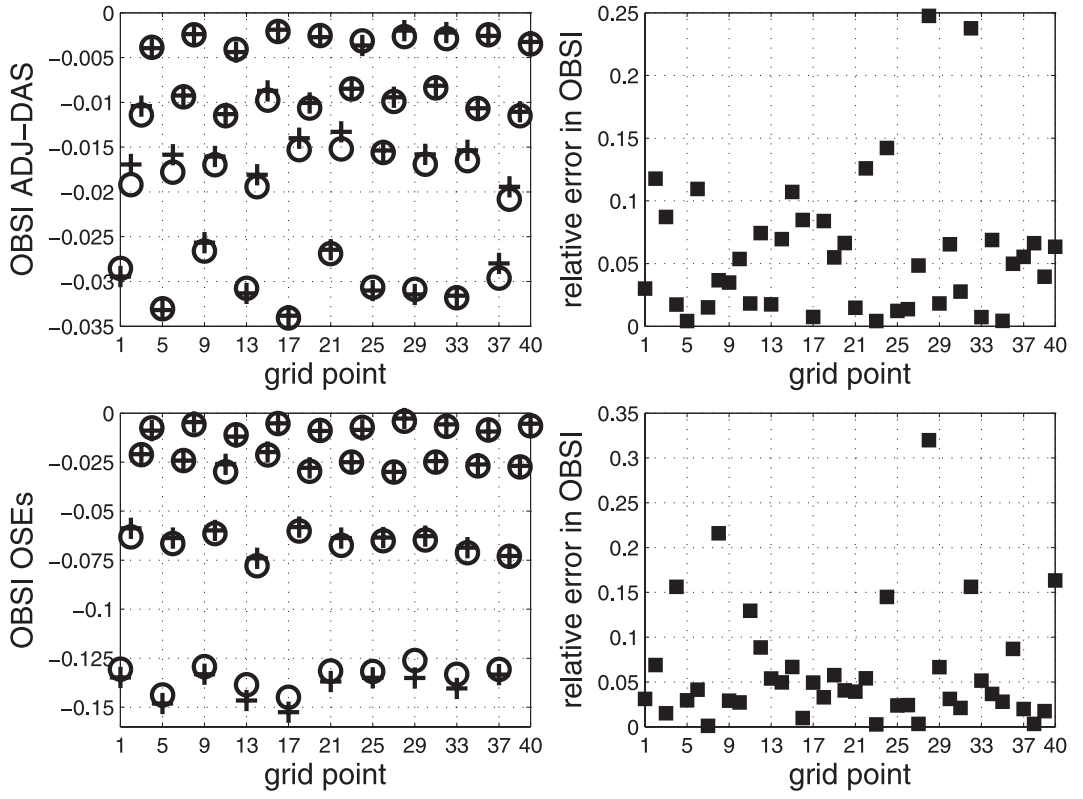

FIG. 3. Time-averaged adjoint-DAS and OSEs' observation impact estimates on the 24-h forecast error in the DAS-I setup. Results with a verification state provided by the DAS (e_v measure, open circles) and with a verification state taken to be the true state (e_t measure, crosses). For each approach, the figure on the right displays the relative error in the OBSI estimation as a result of the errors in the verification state.

implements the second-order measure in (6)–(9) and the time-averaged relative error in the δe_v approximation in (6) was found to be of ~3%. This additional source of uncertainty is characteristic to the current formulation of the adjoint-DAS approach and in operational systems was found to range from 15% to 25% as a result of the difficulties of the practical implementation (Langland 2005; Gelaro et al. 2007). For both OSEs and the adjoint-DAS approach the uncertainty in the OBSI estimation due to the errors in the verification state is monitored by comparing the results from a forecast error measure e_v based on a verification state $\mathbf{x}^v$ produced by the corresponding DAS to the OBSI estimates from a forecast error measure e_t based on the true state $\mathbf{x}^t$.

The results obtained in the DAS-I setup are displayed in Fig. 3. It is noticed that in this idealized DAS both the OSEs and the adjoint-DAS properly identify each data-type impact on the forecast error reduction, at all data locations. Negative OBSI values indicate that each data is of benefit to the forecasts, with data type $\mathbf{y}^{(1)}$ having the largest contribution (grid point locations 1, 5, ... , 37), followed in decreasing order of magnitude by data types $\mathbf{y}^{(2)}$, $\mathbf{y}^{(3)}$, and $\mathbf{y}^{(4)}$. For both OSEs and the adjoint-DAS approach the relative error in the e_v-based OBSI as compared to the e_t-based OBSI remains in general within a small factor from the relative error in the verification state; however, it is noticed that data type $\mathbf{y}^{(4)}$ (grid point locations 4, 8, ... , 40) may be subject to a larger OBSI uncertainty. The complementary nature of the OBSI information extracted from OSEs and the adjoint-DAS approach is discussed in detail in the work of Gelaro and Zhu (2009) and Cardinali (2009) and it is not further addressed here. Since the adjoint-DAS OBSI operator in (9) is linear in the observation space, the OBSI information from the adjoint DAS in Fig. 3 allows a quantification of the impact of any data subset whereas additional experiments are required in the OSEs to analyze the OBSI of other data subsets.

The results obtained in the DAS-II setup are displayed in Fig. 4. Both OSEs and the adjoint-DAS OBSI estimates reveal an increased uncertainty at all locations of data type $\mathbf{y}^{(4)}$ whose observation error is largely underestimated in the DAS-II. It is also noticed that in these experiments the adjoint-DAS OBSI based on the e_v measure attributes a benefic forecast impact from the assimilation of any subset of data type $\mathbf{y}^{(4)}$, whereas both

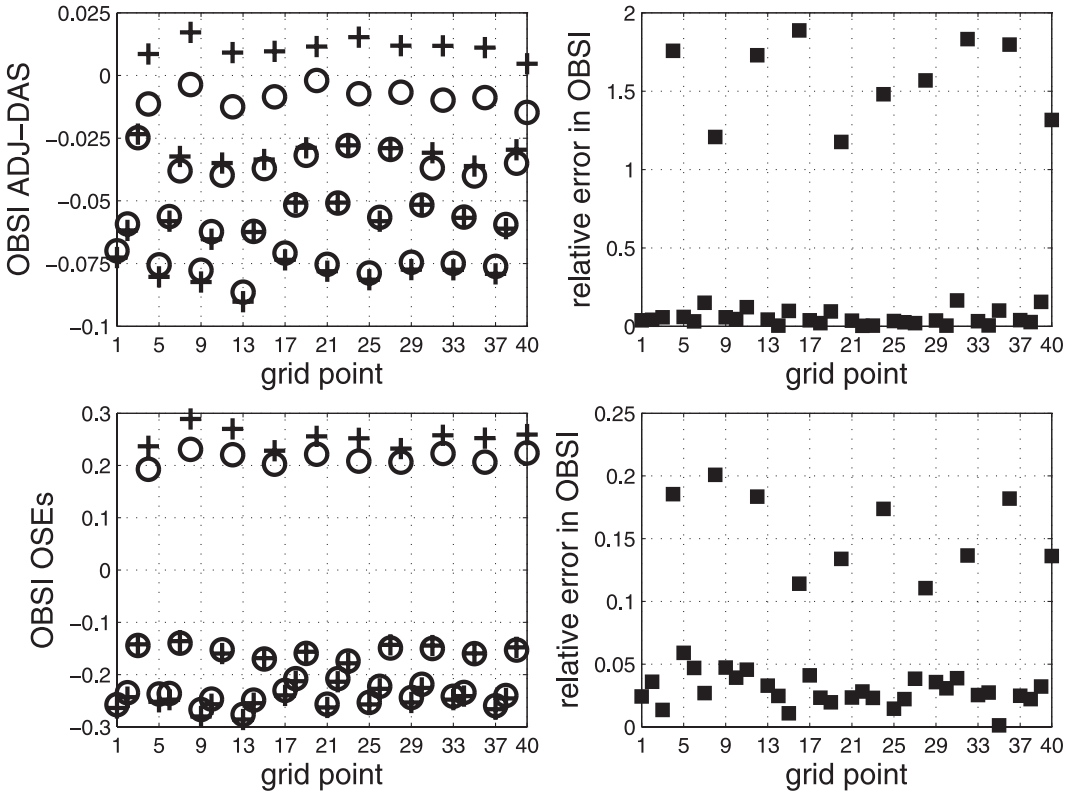

FIG. 4. As in Fig. 3, but for the OBSI estimation in the DAS-II setup.

in the adjoint-DAS approach and OSEs the observation impacts derived from the measure e_t identify the assimilation of this data type as detrimental to the forecasts. The compounded propagation of errors in the verification state and in the δe_v adjoint-based approximation measure in the OBSI estimation needs to be further investigated.

5. Conclusions

This research brings forward the issue of objective assessment of observations value based on the observation impact calculations from deterministic OSEs and the adjoint-DAS approach. A geometrical perspective is provided to the uncertainty in the OBSI estimation in the presence of errors in the verification state. Numerical experiments with a simple model are used to illustrate that the practical difficulty of providing an accurate representation of the true atmospheric state in the forecast error measure may lead to an increased uncertainty in the OBSI estimation derived from deterministic forecast error measures, in particular when suboptimal input error statistics are specified in the DAS. To increase the confidence in the OBSI guidance, a probabilistic framework may be considered using an ensemble of verification states and/or an ensemble of data assimilation systems (Tan et al. 2007; Ancell and Hakim 2007). An objective OBSI assessment relies on the ability to provide verification states whose errors are unbiased and are statistically independent from the errors in the forecasts produced by the DAS.

Acknowledgments. This work was supported by the NASA Modeling, Analysis, and Prediction Program under Award NNG06GC67G.

REFERENCES

Ancell, B. C., and G. J. Hakim, 2007: Comparing ensemble and adjoint sensitivity analysis with applications to observation targeting. *Mon. Wea. Rev.,* **135,** 4117–4134.

——, and C. F. Mass, 2008: The variability of adjoint sensitivity with respect to model physics and basic-state trajectory. *Mon. Wea. Rev.,* **136,** 4612–4628.

Andersson, E., A. Hollingsworth, G. Kelly, P. Lönnberg, J. Pailleux, and Z. Zhang, 1991: Global observing system experiments on operational statistical retrievals of satellite sounding data. *Mon. Wea. Rev.,* **119,** 1851–1865.

Atlas, R., 1997: Atmospheric observations and experiments to assess their usefulness in data assimilation. *J. Meteor. Soc. Japan,* **75,** 111–130.

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

Bouttier, F., and G. Kelly, 2001: Observing-system experiments in the ECMWF 4D-Var data assimilation system. *Quart. J. Roy. Meteor. Soc.,* **127,** 1469–1488.

Cardinali, C., 2009: Monitoring the observation impact on the short-range forecast. *Quart. J. Roy. Meteor. Soc.,* **135,** 239–250.

——, S. Pezzulli, and E. Andersson, 2004: Influence-matrix diagnostic of a data assimilation system. *Quart. J. Roy. Meteor. Soc.,* **130,** 2767–2786.

Daescu, D. N., and R. Todling, 2009: Adjoint estimation of the variation in model functional output due to the assimilation of data. *Mon. Wea. Rev.,* **137,** 1705–1716.

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

Doerenbecher, A., and T. Bergot, 2001: Sensitivity to observations applied to FASTEX cases. *Nonlinear Processes Geophys.,* **8,** 467–481.

Errico, R. M., 2007: Interpretations of an adjoint-derived observational impact measure. *Tellus,* **59A,** 273–276.

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.

Gelaro, R., and Y. Zhu, 2009: Examination of observation impacts derived from observing system experiments (OSEs) and adjoint models. *Tellus,* **61A,** 179–193.

——, ——, and R. M. Errico, 2007: Examination of various-order adjoint-based approximations of observation impact. *Meteor. Z.,* **16,** 685–692.

Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 364 pp.

Kelly, G., J. N. Thépaut, R. Buizza, and C. Cardinali, 2007: The value of observations. I: Data denial experiments for the Atlantic and the Pacific. *Quart. J. Roy. Meteor. Soc.,* **133,** 1803–1815.

Langland, R. H., 2005: 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.

Liu, J., and E. Kalnay, 2008: Estimation of observation impact without adjoint model in an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc.,* **134,** 1327–1335.

Lorenz, E. N., and K. A. Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.,* **55,** 399–414.

Tan, D. G. H., E. Andersson, M. Fisher, and L. Isaksen, 2007: Observing-system impact assessment using a data assimilation ensemble technique: Application to the ADM-Aeolus wind profiling mission. *Quart. J. Roy. Meteor. Soc.,* **133,** 381–390.

Torn, R. D., and G. J. Hakim, 2008: Ensemble-based sensitivity analysis. *Mon. Wea. Rev.,* **136,** 663–677.

Trémolet, Y., 2007: First-order and higher-order approximations of observation impact. *Meteor. Z.,* **16,** 693–694.

——, 2008: Computation of observation sensitivity and observation impact in incremental variational data assimilation. *Tellus,* **60A,** 964–978.

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.