1 Introduction

The adjoint approach to observation sensitivity has been introduced in numerical weather prediction (NWP) by Baker and Daley (2000) for the analysis and design of observation targeting strategies. Subsequently, advances in both theoretical aspects and software capabilities at NWP centers have shown that adjoint-data assimilation system (adjoint-DAS) techniques provide effective tools for the analysis and optimization of the DAS performance. Practical applications include monitoring the impact of data provided by the global observing system to reduce short-range forecast errors, data quality diagnostics and guidance for optimal satellite channel selection, and adaptive observation targeting (Langland and Baker 2004, Cardinali 2009, Baker and Langland 2004, Langland and Baker 2009, Gelaro and Zhu 2009, Gelaro et al. 2010).

The adjoint-DAS applications may be extended to include the forecast sensitivity to the specification of error covariance parameters. Daescu (2008) derived the error-covariance sensitivity equations in a four-dimensional variational (4DVAR) DAS. The practical ability to obtain the forecast sensitivity to observation and background error covariance weight parameters was illustrated by Daescu and Todling (2010) using a simplified version of the NASA Goddard Earth Observing System (GEOS-5) atmospheric DAS and its adjoint developed at the Global Modeling and Assimilation Office.

This work presents theoretical aspects of error covariance sensitivity analysis and recent results obtained with the adjoint versions of the Naval Research Laboratory Atmospheric Variational Data Assimilation System - Accelerated Representer (NAVDAS-AR) (Xu et al. 2005, Rosmond and Xu 2006) and the Navy Operational Global Atmospheric Prediction System (NOGAPS) (Hogan and Rosmond 1991).

2 Theoretical aspects

Variational data assimilation (VDA) provides an analysis \(\mathbf{x}^a \in \mathbb{R}^n\) to the true state \(\mathbf{x}^f\) of the atmosphere by minimizing the cost functional

\[
J(\mathbf{x}) = J^b + J^o = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2} [\mathbf{h}(\mathbf{x}) - \mathbf{y}]^T \mathbf{R}^{-1} [\mathbf{h}(\mathbf{x}) - \mathbf{y}] \tag{1}
\]

where \(\mathbf{x}^b \in \mathbb{R}^n\) is a prior (background) state estimate, \(\mathbf{y} \in \mathbb{R}^p\) is the vector of observational data, and \(\mathbf{h}\) is the observation operator that maps the state into observations. Statistical information on the background error \(\mathbf{e}^b = \mathbf{x}^b - \mathbf{x}^f\) and observational error \(\mathbf{e}^o = \mathbf{y} - \mathbf{h}(\mathbf{x}^f)\) is used to specify the weighting matrices \(\mathbf{B} \in \mathbb{R}^{n \times n}\) and \(\mathbf{R} \in \mathbb{R}^{p \times p}\) that are representations in the DAS of the background and observation error covariances \(\mathbf{B}_t = \mathbb{E}(\mathbf{e}^o \mathbf{e}^o^T)\) and \(\mathbf{R}_t = \mathbb{E}(\mathbf{e}^b \mathbf{e}^b^T)\) respectively, where \(\mathbb{E}(\cdot)\) denotes the statistical expectation operator. For a linear observational operator, \(\mathbf{b}(\mathbf{x}) = \mathbf{Hx}\), the analysis is expressed as

\[
\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}[\mathbf{y} - \mathbf{Hx}^b] \tag{2}
\]

where the gain matrix \(\mathbf{K}\) is defined as

\[
\mathbf{K} = [\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}]^{-1} \mathbf{H}^T \mathbf{R}^{-1} = \mathbf{B}^T [\mathbf{HB}^T + \mathbf{R}]^{-1} \tag{3}
\]

In this context, Baker and Daley (2000) derived the sensitivity equations of a scalar forecast aspect \(e(\mathbf{x}^a)\) with respect to observations and background

\[
\frac{\partial e}{\partial \mathbf{y}} = \mathbf{K}^T \frac{\partial e}{\partial \mathbf{x}^a} \tag{4}
\]

\[
\frac{\partial e}{\partial \mathbf{x}^b} = [\mathbf{I} - \mathbf{H}^T \mathbf{K}^T] \frac{\partial e}{\partial \mathbf{x}^a} \tag{5}
\]

where \(\mathbf{I}\) denotes the \(n \times n\) identity matrix and \(\mathbf{K}^T\) is the adjoint-DAS operator. In the appendix we provide a self-contained derivation of the forecast sensitivity equations to the DAS input \([\mathbf{y}, \mathbf{R}, \mathbf{x}^b, \mathbf{B}]\) for
a linear analysis scheme. We refer to Daescu (2008) for the extension to a nonlinear 4DVAR DAS and to Trémolet (2008) for the simplifying assumptions that are necessary to estimate the observation sensitivity in an incremental VDA. The forecast $\mathbf{B}$- and $\mathbf{R}$-sensitivity is expressed in terms of the $\mathbf{x}^b$- and $\mathbf{y}$-sensitivity respectively, as

$$\frac{\partial e}{\partial \mathbf{B}} = \frac{\partial e}{\partial \mathbf{x}^b} (\mathbf{x}^a - \mathbf{x}^b)^\top \mathbf{B}^{-1}$$  \hspace{1cm} (6)$$

$$\frac{\partial e}{\partial \mathbf{R}} = \frac{\partial e}{\partial \mathbf{y}} [\mathbf{h}(\mathbf{x}^a) - \mathbf{y}]^\top \mathbf{R}^{-1}$$  \hspace{1cm} (7)

Estimation of $\mathbf{B}$- and $\mathbf{R}$- sensitivities associated to a short-range forecast error measure may be used to assess the optimality of the DAS error covariance model ($\mathbf{B}, \mathbf{R}$) and to provide guidance to error covariance tuning procedures. Proper weighting between the background information and the information provided by uncorrelated observing system components $\mathbf{y}_i, i \in I$, may be investigated through a parametric specification

$$\mathbf{B}(s^b) = s^b \mathbf{B}, \quad \mathbf{R}_i(s^o_i) = s^o_i \mathbf{R}_i, \quad i \in I$$  \hspace{1cm} (8)

which is a common representation used to perform error covariance tuning (Desroziers et al. 2009). The forecast sensitivity to the covariance weight factors $s^b > 0$, $s^o_i > 0$ is expressed in terms of the observation sensitivity and the observed-minus-analysis residual as (Daescu and Todling 2010)

$$\frac{\partial \mathbf{e}(\mathbf{x}^a)}{\partial s^b} = [\mathbf{y} - \mathbf{h}(\mathbf{x}^a)]^\top \frac{\partial \mathbf{e}(\mathbf{x}^a)}{\partial \mathbf{y}}$$  \hspace{1cm} (9)$$

$$\frac{\partial \mathbf{e}(\mathbf{x}^a)}{\partial s^o_i} = [\mathbf{h}(\mathbf{x}^a) - \mathbf{y}_i]^\top \frac{\partial \mathbf{e}(\mathbf{x}^a)}{\partial \mathbf{y}_i}, \quad i \in I$$  \hspace{1cm} (10)

where $\mathbf{x}^a$ denotes the analysis state provided by the reference DAS configuration ($\mathbf{B}, \mathbf{R}$) with all weight parameters in (8) set to 1. In conjunction with the observation sensitivity, the observed-minus-analysis (o-m-a) information allows the identification of the input components where improved estimates of the error statistics have a potentially large impact on the forecast error reduction. In particular, the steepest descent direction

$$\mathbf{d} = -\frac{\partial \mathbf{e}(\mathbf{x}^a)}{\partial \mathbf{s}}$$  \hspace{1cm} (11)

identifies the direction of small variations $\delta \mathbf{s}$, from the current DAS configuration, that will be of largest forecast benefit and provides a first order error covariance diagnostic.

### 3 Numerical results

Sensitivities derived from the adjoint of the NAVDAS-AR analysis and the adjoint of the NOGAPS forecast model are used to assess the optimality of the information weighting in the data assimilation system and first order diagnostics for error covariance tuning. The forecast score is defined as a 24-hour forecast error measure

$$e(\mathbf{x}^a) = (\mathbf{x}_f^a - \mathbf{x}_f^b)^\top \mathbf{E} (\mathbf{x}_f^a - \mathbf{x}_f^b)$$  \hspace{1cm} (12)

where $\mathbf{x}_f^b = M_{t_f \to t_f}(\mathbf{x}^a)$ is the model forecast at verification time $t_f$ initiated from $\mathbf{x}^a$, $\mathbf{x}_f^b$ is the verifying analysis at $t_f$ and serves as a proxy to the true state $\mathbf{x}_f^a$, and $\mathbf{E}$ is a diagonal matrix of weights that gives (12) units of energy per unit mass. The adjoint-DAS software tools developed at NRL for observation impact estimation facilitate the error covariance sensitivity analysis. The observation impact measure introduced by Langland and Baker (2004)

$$e(\mathbf{x}^a) - e(\mathbf{x}^b) \approx (\delta \mathbf{y})^\top \mathbf{K}^\top \left[ \frac{1}{2} \frac{\partial e}{\partial \mathbf{x}^b} + \frac{1}{2} \frac{\partial e}{\partial \mathbf{x}^a} \right]$$  \hspace{1cm} (13)

is defined in terms of the innovation vector $\delta \mathbf{y} = \mathbf{y} - \mathbf{h}(\mathbf{x}^b)$ and forecast gradients evaluated along the analysis and background trajectories. The sensitivities (9) and (10) to each covariance weight factor are defined in terms of the observation-minus-analysis (residual) vector $\mathbf{y} - \mathbf{h}(\mathbf{x}^a)$ and the forecast gradient evaluated along the analysis trajectory. The computations that are necessary to evaluate (9), (10), and (13) share the same adjoint tools and may be performed simultaneously to obtain complementary information that is necessary to optimize the DAS performance. The observation impact estimation provides an assessment of the contribution of each observing system component $\mathbf{y}_i$ to the forecast error reduction, given the DAS error covariance model ($\mathbf{B}, \mathbf{R}$). The ($\mathbf{B}, \mathbf{R}$)-sensitivities provide guidance on the variations in the error covariance parameters that are necessary to reduce the forecast error, given the observing system configuration $\mathbf{y}$.

The observation impact (13) and sensitivities (9), (10) are evaluated for the 24-h forecasts initiated at every analysis time 00 UTC for the time period July 15 - August 15 of 2010 (hereafter summer period) and for the time period September 29 - October 26 of 2010 (hereafter fall period). The summer period incorporates 30 data sets (data for July 28 and August 1 was not incorporated in this study) whereas the fall period incorporates 27 data sets (data for October 15 was not incorporated in this study).
The total amount of data at 00UTC assimilated during each time period is displayed in Fig. 1. The fall period incorporates a significantly larger amount of radiance data and in addition, it incorporates Global Positioning Satellite (GPS) refractivity data from atmospheric sounding instruments on COSMIC, METOP-A (GRAS), and GRACE-A. The contribution of each data type to the forecast error reduction, as derived from (13), is shown in Fig. 2.

A recent study on observation impact assessment at various NWP centers is provided in the work of Gelaro et al. (2010). We discuss the additional adjoint-DAS ability to obtain error-covariance sensitivity information and its relevance to establish DAS diagnostics and guidance to forecast error reduction. By analogy with the observation impact estimation, the adjoint-DAS approach provides a detailed error covariance sensitivity information and allows the analysis of each observation type, instrument, and data location in the time-space domain. Daily-averaged forecast sensitivities to the observation error covariance weight factor (10) are displayed for each observation type and for various observing instruments in Fig. 3 and Fig. 4, respectively. For comparison, the values of the sensitivity to the background error covariance weight (9) are also shown.

Each of the $s$-sensitivities $\partial e / \partial s$ provides a first order guidance on the forecast impact as a result of variations in the corresponding error covariance weight coefficient. Forecast error reduction may be obtained by taking a steepest descent update

$$s^* = 1 - \alpha \frac{\partial e}{\partial s} \quad (14)$$

where $\alpha > 0$ is an appropriate step length. A negative sensitivity $\frac{\partial e}{\partial s} < 0$ implies $s^*_i > 1$ and identifies
the DAS components whose error covariance inflation is of potential benefit to the forecasts, whereas a positive sensitivity $\frac{\partial e}{\partial s_i} > 0$ implies $s_i^* < 1$ and identifies the DAS components whose error covariance deflation is of potential benefit to the forecasts.

A large negative $s^b$-sensitivity is noticed in both experiments and indicates that background error covariance inflation is of potential benefit to the forecasts (alternatively, the information provided by the observing system as a whole is under-weighted in the DAS). In Figs. 3 and 4 it is noticed that a large fraction of the sensitivity to the radiosondes error covariance weight is explained by the sensitivity associated to the relative humidity and points to a suboptimal assimilation of this data. Additional information is shown in Fig. 5 as daily-averaged observation impact and forecast sensitivity to the error covariance weight for the radiosonde data between various pressure levels during the summer period.

The combined sensitivity analysis in Fig. 3-5 indicates that tuning the observing system through a single weight coefficient ($I = 1$) is suboptimal. Diagnostics and guidance to the parameterization (9), (10) that is necessary to optimize the use of data may be obtained by systematically monitoring the sensitivity to error covariance parameters along with the observation impact estimates.

Noticeable in Fig. 3 and Fig. 4 is the variation from the summer period to the fall period in the forecast sensitivity to the error covariance weight associated to the radiosonde data. Valuable insight is obtained from a time-series analysis of each satellite instrument that allows the identification of tendencies in the forecast sensitivities to the information weight, as illustrated in Fig. 6 for the Advanced Microwave Sounding Unit (AMSU)-A and for the Infrared Atmospheric Sounding Interferometer (IASI).

Caution must be exercised in the interpretation of the sensitivity analysis for tuning DAS error covariance parameters and this information must be considered in conjunction with other diagnostics such as the methods of Desroziers and Ivanov (2001) and Desroziers et al. (2005). Bormann and Bauer (2010) and Bormann et al. (2010) implemented various diagnostics to estimate observation error statistics in the ECMWF system. Their findings suggest that observation-error estimates for sounder radiances are lower than the observation errors currently assigned in the DAS and indicate a too-conservative use of the AMSU-A instrument. In their study it was also found that AMSU-A shows little spatial and interchannel error correlations and that larger interchannel error correlations are present for surface-sensitive temperature-sounding or window channels for IASI. This is in agreement with the results in Fig. 6 where positive sensitivities indicate that data provided by the AMSU-A instrument appears to be under-weighted in the DAS. In Fig. 6, negative sensitivities associated to IASI suggest that further inflation of the assigned observation error variances for
this instrument may be of benefit to the forecasts. Since observation error correlations are not modeled in current data assimilation systems, the sensitivity guidance may be interpreted as an attempt to compensate for misspecified observation error correlations in the DAS through inflation of the assigned error variances.

4 Conclusions and future work

Modeling of the observation and background error covariances is a key ingredient in the implementation of atmospheric data assimilation techniques and an area of intensive research in NWP. The adjoint-DAS approach is a routine procedure used to monitor the observation performance on short-range forecasts and, for the first time in an operational DAS, we presented the ability to obtain derivative-based forecast sensitivities to error covariance parameters. This information may be obtained along with the observation impact assessment and may be used as an additional diagnostic tool and to provide guidance to error covariance tuning procedures. The simplicity of the parametric representation (8) facilitated the analysis presented in this study based on the data assimilation products developed on a routine basis at NRL. Equations (6) and (7) allow the extension of the sensitivity analysis to key ingredients in the error covariance modeling, the specification of the information error standard deviation and the information error correlation model, $R = \Sigma a^c \Sigma^c$, $B = \Sigma a^b \Sigma^b$. Our future work will investigate theoretical aspects and practical applications of these additional adjoint capabilities to optimize the DAS performance as well as the incorporation of the model error covariance sensitivity analysis in a weak-constraint 4DVAR.

Acknowledgements

The work of D.N. Daescu was supported by the NRL Atmospheric Effects, Analysis, and Prediction Program under award N00173-10-1-G032 and by the National Science Foundation under award DMS-0914937.

Appendix

For a linear observation operator, the cost functional (1) may be expressed as

$$J(x) = \frac{1}{2} (\Gamma x - z)^T S^{-1} (\Gamma x - z)$$

where

$$z = \begin{bmatrix} x^b \\ y \end{bmatrix} \in \mathbb{R}^{n+p}, \quad \Gamma = \begin{bmatrix} I \\ H \end{bmatrix} \in \mathbb{R}^{(n+p) \times n}$$

denote the information vector and the information operator respectively, and

$$S = \begin{bmatrix} B & 0_{n \times p} \\ 0_{p \times n} & R \end{bmatrix}$$

is the DAS model of the information error covariance. The first-order optimality condition to (15)

$$\Gamma^T S^{-1} (\Gamma x^a - z) = 0$$

is used to express the first order analysis variation $\delta x^a$ and, using chain rule differentiation, the first order forecast variation $\delta e$ in terms of perturbations in each DAS input component $^1$.

Sensitivity to the information vector

The first order variation $\delta x^a$ associated to a perturbation $\delta z$ is expressed from (18) as

$$\Gamma^T S^{-1} \delta x^a = \Gamma^T S^{-1} \delta z$$

It is noticed that the Hessian matrix of the cost (15) defines a self-adjoint (symmetric) operator on $\mathbb{R}^n$

$$\Gamma^T S^{-1} \Gamma = B^{-1} + H^T R^{-1} H \in \mathbb{R}^{n \times n}$$

Let $\lambda \in \mathbb{R}^n$ defined as the solution to the adjoint problem

$$\Gamma^T S^{-1} \Gamma \lambda = \frac{\partial e}{\partial x^a}$$

By taking the inner product of (19) with $\lambda$ and using (21) and the symmetry of the Hessian matrix (20),

$$\left\langle \frac{\partial e}{\partial x^a}, \delta x^a \right\rangle_{\mathbb{R}^n} = \left\langle \lambda, \Gamma^T S^{-1} \delta z \right\rangle_{\mathbb{R}^n} = \left\langle S^{-1} \Gamma \lambda, \delta z \right\rangle_{\mathbb{R}^{n+p}}$$

The left side of (22) is the first order variation of the functional aspect $e(x^a)$,

$$\delta e = \left\langle \frac{\partial e}{\partial x^a}, \delta x^a \right\rangle_{\mathbb{R}^n}$$

$^1$For a perturbation $\delta u$ in the input component $u$, the first order variation $\delta x^a$ is the differential of $x^a$ at $u$ in the direction $\delta u$, $\delta x^a = \delta x^a(u) \delta u$. Chain rule differentiation provides $\delta e = de[x^a(u)]/\delta x^a = de[x^a(u)]/\delta x^a(u)\delta u$. 
and from (22) and (23) the forecast sensitivity to the information vector is expressed as

$$\frac{\partial e}{\partial z} = S^{-1} \Gamma \lambda \in \mathbb{R}^{n+p} \quad (24)$$

Using matrix properties, it is noticed that (5) and (4) are a componentwise representation to (24)

$$\frac{\partial e}{\partial z} = \left[ \begin{array}{c} \frac{\partial e}{\partial \xi^0} \\ \frac{\partial e}{\partial y} \end{array} \right] = \left[ \begin{array}{c} B^{-1} \lambda \\ R^{-1} H A \end{array} \right] \quad (25)$$

**Sensitivity to the error covariance model**

The first order variation $\delta x^a$ associated to a perturbation $\delta S$ is expressed from (18) as

$$\Gamma^T S^{-1}(\Gamma x^a - z) + \Gamma^T S^{-1} \delta x^a = 0 \quad (26)$$

where

$$\delta S^{-1} = -S^{-1} \delta SS^{-1} \quad (27)$$

is the first order variation in the inverse of $S$ induced by the perturbation $\delta S$. Next, we replace (27) in (26) and separate the terms to obtain

$$\Gamma^T S^{-1} \delta x^a = \Gamma^T S^{-1} \delta SS^{-1}(\Gamma x^a - z) \quad (28)$$

Taking the inner product of (28) with $\lambda$ defined as in (21) and using the symmetry of the Hessian matrix,

$$\left\langle \frac{\partial e}{\partial x^a}, \delta x^a \right\rangle_{\mathbb{R}^n} = \left\langle \lambda, \Gamma^T S^{-1} \delta SS^{-1}(\Gamma x^a - z) \right\rangle_{\mathbb{R}^n} = \left\langle \frac{\partial e}{\partial z}, \delta SS^{-1}(\Gamma x^a - z) \right\rangle_{\mathbb{R}^{n+p}} \quad (29)$$

Equation (24) is used to establish the last equality in (29). The left side of (29) is the first order variation $\delta e$ in the forecast aspect whereas the right side of (29) is the Frobenius inner product on the vector space of $(n+p) \times (n+p)$-dimensional matrices defined in terms of the matrix trace operator $\text{Tr}$:

$$\delta e = \left\langle \frac{\partial e}{\partial z}, \delta SS^{-1}(\Gamma x^a - z) \right\rangle_{\mathbb{R}^{n+p}} = \text{Tr} \left[ \delta SS^{-1}(\Gamma x^a - z) \left( \frac{\partial e}{\partial z} \right)^T \right]$$

$$= \left\langle \delta S, \frac{\partial e}{\partial z}(\Gamma x^a - z)^T S^{-1} \right\rangle_{\mathbb{R}^{(n+p) \times (n+p)}} \quad (30)$$

Therefore, the forecast $S$-sensitivity is a rank-one matrix that is expressed in terms of the sensitivity to the information vector as

$$\frac{\partial e}{\partial S} = \frac{\partial e}{\partial \Gamma} (\Gamma x^a - z)^T S^{-1} \quad (31)$$

For a linear observational operator, (31) implies (6) and (7).

**References**


