Comparing Two Approaches for Assessing Observation Impact

RICARDO TODLING

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

(Manuscript received 4 April 2012, in final form 20 September 2012)

ABSTRACT

Langland and Baker introduced an approach to assess the impact of observations on the forecasts. In that approach, a state-space aspect of the forecast is defined and a procedure is derived ultimately relating changes in the aspect with changes in the observing system. Some features of the state-space approach are to be noted: the typical choice of forecast aspect is rather subjective and leads to incomplete assessment of the observing system, it requires availability of a verification state that is in practice correlated with the forecast, and it involves the adjoint operator of the entire data assimilation system and is thus constrained by the validity of this operator. This article revisits the topic of observation impacts from the perspective of estimation theory. An observation-space metric is used to allow inferring observation impact on the forecasts without the limitations just mentioned. Using differences of observation-minus-forecast residuals obtained from consecutive forecasts leads to the following advantages: (i) it suggests a rather natural choice of forecast aspect that directly links to the data assimilation procedure, (ii) it avoids introducing undesirable correlations in the forecast aspect since verification is done against the observations, and (iii) it does not involve linearization and use of adjoints. The observation-space approach has the additional advantage of being nearly cost free and very simple to implement. In its simplest form it reduces to evaluating the statistics of observation-minus-background and observation-minus-analysis residuals with traditional methods. Illustrations comparing the approaches are given using the NASA Goddard Earth Observing System.

1. Introduction

Langland and Baker (2004, hereafter LB04) introduce a technique to examine the impact of observations on the short-range forecast. In that technique, an aspect of the forecast is defined and changes to the aspect are then associated to changes in the observing system. LB04 derive an expression to calculate observation impacts using the sensitivity (adjoint) operators of both the underlying forecast model and that of the analysis system. Errico (2007, hereafter E07) rederives the expression of LB04, as well as introduces other expressions, using a Taylor series expansion of the forecast aspect when the initial conditions are subject to infinitesimal changes. Gelaro et al. (2007, hereafter GZE07) use an early version of the National Aeronautics and Space Administration (NASA) fifth-generation Goddard Earth Observing System (GEOS-5), together with the adjoint of the GEOS-5 general circulation model (GCM), and the line-by-line adjoint of the gridpoint statistical interpolation (GSI) analysis system of Zhu and Gelaro (2008), to examine in detail the various approximations in E07. More recently, Daescu and Todling (2009, hereafter DT09) show that the expressions in E07, and higher-order accurate ones, can also be derived using a parametric approach in conjunction with numerical quadrature methods.

When the technique of LB04, E07, and DT09 is used to assess the impact of the observing system on the forecast one must keep in mind the assumptions and limitations of the method. First, the definition of the forecast aspect reflects whatever measure one judges to be relevant to describe features of interest. Much of the work done in this area has thus far focused on the impact of observations in the 24-h forecasts. The forecast aspect has typically been chosen to measure the 24-h forecast error, as defined by a *linearized total energy* norm. In addition, the works of LB04, GZE07, and DT09 have applied a projection operator to the forecast error measure that intentionally excludes errors, roughly, above 100 hPa. Other works have used very specific projection

Corresponding author address: Dr. Ricardo Todling, Global Modeling and Assimilation Office, NASA GSFC, Code 610.1, Greenbelt, MD 20771.
E-mail: ricardo.todling@nasa.gov

DOI: 10.1175/MWR-D-12-00100.1

operators as they are aimed at evaluating targeted observations (Langland 2005; Buizza et al. 2007; Cardinali et al. 2007; Kelly et al. 2007; Rabier et al. 2008). Undeniably, use of such measures restrict conclusions drawn about the observing system and its impact on forecasts in general.

Second, a further complication associated with the definition of typical forecast aspects is the need to use a verifying state. It is common practice to evaluate the quality of forecasts with respect to underlying analyses, as for example when calculating forecast skill scores. However, one must recognize the limitations of such a practice since correlations between the forecasts and the verifying analyses introduce unknown features to the results, particularly when relatively short-range forecasts are involved.

Third, the approximate observation impact formulas of LB04, E07, and DT09, are limited by the validity of the adjoint models they invoke. In particular, the validity of the underlying general circulation model adjoint restricts the studies of observation impact to those related to *short-range* forecasts *only*, though there are currently attempts to extend the validity of adjoints by defining optimal trajectories (see Barkmeijer and Stappers 2011). Lack of proper representation of full nonlinear processes introduces yet another uncertainty factor (e.g., some adjoints have poor or no representation of convective processes). Furthermore, the nonlinearity of many analysis systems (Trémolet 2007, 2008) introduces another layer of complexity and limitation to the adjoint-based approach. The ability to use adjoint-free diagnostics is advantageous, at least from the practical point of view. Liu and Kalnay (2008) have recently introduced an ensemble-based observation impact technique that, by construction, avoids the need for adjoints (see also Liu et al. 2009). Still, the work of Liu and Kalnay is based on a rather subjective forecast aspect, and continues to rely on the availability of verification states.

More generally, it is not necessarily clear how using information derived from observation impact evaluations can aid the development and enhancement of data assimilation systems. Since observation impact results tangle the use of observations through the data assimilation process with errors in the short-range forecast that are not part of the data assimilation process, it is hard to disentangle the information provided by observation impact studies to make sound decisions about the use of observations. It is thus the main point of the present work to argue that one might as well stick with traditional methods of evaluating the statistics of observation-minus-background (OMB) and, to some extent, observation-minus-analysis (OMA) residuals to allow one to infer the required information to improve upon the cycling data assimilation scheme and the use of observations (e.g., Hollingsworth and Lönnberg 1989; Daley 1992; Dee and da Silva 1999; Desroziers et al. 2005a,b; Chapnik et al. 2006; Lupu et al. 2011).

To build toward the main goal here, the present work examines the impact of observations on the forecasts by introducing measures based on weighted observation-minus-forecast (OMF) residuals directly. The weighted difference of the squared OMF residuals calculated from two forecasts issued from two consecutive analyses serves as a means to evaluate the impact of observations on forecasts. The limiting case in this approach is one in which the forecast is reduced to be the very short-range background, thus leading back to the more traditional residual-based techniques cited above. The advantages of this approach are: (i) it still allows for a complete assessment of the observing system; (ii) it avoids introducing undesirable corrections to the forecast aspect; (iii) it involves no approximations to the model and analysis operators, avoiding adjoints and consequently being devoid of their limitations; and finally, (iv) it requires only straightforward calculations involving quantities readily available in most practical data assimilation systems.

Arguing from the perspective of estimation theory, we examine the *statistics* of the various approaches to observation impact. It becomes evident that the *state-space* approach is, in principle, more encompassing than the *observation-space* approach. It is shown that, under idealized conditions, it is possible to choose a state-space measure that obtains the same *expected* observation impact as that obtained with a corresponding observation-space measure. It is further shown that optimality of the data assimilation system implies that the expected observation impacts are always negative in the expected mean sense, that is, assimilation of observations always amounts to improvements in the forecast. The analytic evaluation then goes on to examine the consequences of having to replace the unknown truth with an arbitrary verifying state, and the consequences resulting from when the verifying state is chosen to be the analysis. At this point it becomes rather clear that the observation-space approach is preferable over the state-space approach since the former uses the observations for verification and therefore does not introduce spurious correlation factors as does the latter. Ultimately, it is argued that if the objective of calculating observation impacts is to aid the underlying data assimilation procedure and help decide how best to use the available observations, employment of the observation-space approach, with the additional simplification of taking the fields involved in the measures difference to be the analysis and the background, suffices.

Illustration of the various points raised along the text is given by using GEOS-5. The dependency of the approaches on the definition of the forecast aspect is given by calculating observation impacts on the 24-h forecasts using three distinct measures based on (i) the tropospheric, linearized total energy norm used in LB04 and other works; (ii) a modified version of the linearized total energy norm that more evenly weighs the vertical (e.g., Errico et al. 2007); and (iii) the inverse of the observation error covariances used by the analysis system, projected onto state space. There are important differences among results obtained with these different measures. Though most differences can be well understood, the clear message is that the ranking of the observing system that almost naturally follows from these studies is a rather subjective matter. The theoretical considerations of section 2 suggest that the observation-space approach can be used in practice to try to assess the consequences of verifying forecasts against analyses. The experiments with GEOS-5 indicate these verifications tend to overestimate the impact of the observations. This is even more so when the state-space approach is used. Last, given the implicit subjectivity of the various measures of observation impacts some illustration is given to enforce the view that one might as well rely on the, readily available, usual OMB and OMA residuals to try to assess how observations are used in the assimilation process and what can be done with this information to make system improvements.

In what follows, section 2 compares the state- and observation-space approaches from the perspective of estimation theory. Section 3 shows results comparing observation impacts on the forecasts obtained using various measures and linking to the results of section 2. Closing remarks appear in section 4.

2. Measures of observation impact

a. Background

Let us write the expression describing a forecasting model as

$$\mathbf{x}^f_{k|k-m+1} = \mathbf{m}_{k,k-m+1}(\mathbf{x}^a_{k-m+1|k-m+1}), \tag{1}$$

where here, borrowing from the notation of estimation theory (e.g., see Cohn et al. 1994), the n-vector forecast state $\mathbf{x}^f_{k|k-m+1}$, at time t_k, is derived by integration of the model $\mathbf{m}$, from time t_{k-m+1} to time t_k, starting from an analysis state $\mathbf{x}^a_{k-m+1|k-m+1}$ calculated at time t_{k-m+1}. The subscript notation $i\,|\,j$ indicates that the estimate at time t_i is obtained by using observations up to and including observations at time t_j, for $j \le i$. In the linear case, this notation is associated with the more profound statement that the minimum variance estimate is the conditional mean (e.g., Cohn 1997). In the nonlinear case considered here caution must be exercised with this interpretation, but the notation is still informative.

In data assimilation, the analysis is an estimate that combines a model background field,

$$\mathbf{x}^b_{k-m+1|k-m} \equiv \mathbf{x}^f_{k-m+1|k-m} = \mathbf{m}_{k-m+1,k-m}(\mathbf{x}^a_{k-m|k-m}), \tag{2}$$

with the p_{k-m+1}-vector of observations $\mathbf{y}^o_{k-m+1}$, at time t_{k-m+1}, and can conveniently be written in the following form:

$$\mathbf{x}^a_{k-m+1|k-m+1} = \mathbf{x}^b_{k-m+1|k-m} + \tilde{\mathbf{K}}_{k-m+1|k-m} \times [\mathbf{y}^o_{k-m+1} - \mathbf{h}_{k-m+1}(\mathbf{x}^b_{k-m+1|k-m})], \tag{3}$$

where $\mathbf{h}_{k-m+1}$ is the p_{k-m+1} observation operator that transforms model states into observables: $\tilde{\mathbf{K}}_{k-m+1|k-m}$ is a general, not necessarily optimal, $n \times p_{k-m+1}$ matrix of weights used to update the background state given observations at time t_{k-m+1}; and $\mathbf{d}_{k-m+1|k-m} \equiv \mathbf{y}^o_{k-m+1} - \mathbf{h}_{k-m+1}(\mathbf{x}^b_{k-m+1|k-m})$ is the p_{k-m+1}-vector of residuals representing the difference between the actual and the model-predicted observations. The analysis expression (3) is representative of a three-dimensional variational formulation; for convenience, it is written here at time t_{k-m+1} rather than at the usual time t_k; taking $(k - m + 1) \to k$ converts it into its familiar form.

Studies of observation impact on the forecast introduce a scalar functional to measure the quality of the forecast and establish a common means to evaluate the effect of observations on the forecast. A typical quadratic scalar measure takes the following form:

$$e_{k|\ell} \equiv (\mathbf{x}^f_{k|\ell} - \mathbf{x}^t_k)^{\mathrm{T}}\mathbf{T}_k(\mathbf{x}^f_{k|\ell} - \mathbf{x}^t_k), \tag{4}$$

where the $n \times n$ symmetric positive semidefinite matrix $\mathbf{T}_k$ stands for the weight given to the forecast error $\epsilon^f_{k|\ell} \equiv \mathbf{x}^f_{k|\ell} - \mathbf{x}^t_k$, evaluated at time t_k, for $\ell < k$. In practice, one cannot calculate the weighted forecast error in (4) since it involves the *unknown* true state $\mathbf{x}^t_k$. Instead, a *verification* state $\mathbf{x}^v_k$ is used, so that the forecast error becomes $\epsilon^v_{k|\ell} \equiv \mathbf{x}^f_{k|\ell} - \mathbf{x}^v_k$, and the scalar measure (4) is replaced with

$$e^v_{k|\ell} \equiv (\mathbf{x}^f_{k|\ell} - \mathbf{x}^v_k)^{\mathrm{T}}\mathbf{T}_k(\mathbf{x}^f_{k|\ell} - \mathbf{x}^v_k). \tag{5}$$

Naturally, the verification state is usually taken to be an analysis, the consequences of which are discussed in what follows.

The impact of observations on the forecast can be evaluated by studying how the forecast error measure (5) changes with respect to changes in the initial condition in (1). These changes can be thought to be a result of changes due to the assimilation of observations. Assuming the changes to be infinitesimal, E07 and DT09 derive various approximations to the corresponding change in the forecast error measure. Relevant to the present work are the following first- and second-order (per DT09) accurate formulas:

$$\delta e_k^{v,1} = \mathbf{d}_{k-m+1|k-m}^{\mathrm{T}} \tilde{\mathbf{K}}_{k-m+1|k-m}^{\mathrm{T}} \nabla_{\mathbf{x}^b} e_{k|k-m}^{v}, \tag{6a}$$

$$\delta e_k^{v,2} = \frac{1}{2} \mathbf{d}_{k-m+1|k-m}^{\mathrm{T}} \tilde{\mathbf{K}}_{k-m+1|k-m}^{\mathrm{T}} \times [\nabla_{\mathbf{x}^b} e_{k|k-m}^{v} + \nabla_{\mathbf{x}^a} e_{k|k-m+1}^{v}], \tag{6b}$$

where the gradient vectors above can be written as

$$\nabla_{\mathbf{x}^g} e_{k|\ell}^{v} = 2\mathbf{M}_{g;k,k-m+1}^{\mathrm{T}} \mathbf{T}_k [\mathbf{x}_{k|\ell}^{f} - \mathbf{x}_k^{v}], \tag{7}$$

with the $n \times n$ matrix,

$$\mathbf{M}_{g;k,k-m+1} = \left. \frac{\partial \mathbf{m}_{k,k-m+1}(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}^g}, \tag{8}$$

standing for the Jacobian of the model $\mathbf{m}$ in (1) integrated from time t_{k-m+1} to t_k, and linearized about integrations started from either the background $\mathbf{x}^g = \mathbf{m}(\mathbf{x}_{k-m+1|k-m}^b)$, indexed $g = b$, or the analysis $\mathbf{x}^g = \mathbf{m}(\mathbf{x}_{k-m+1|k-m+1}^a)$, indexed $g = a$, respectively; in (7), $\ell = k - m$ or $\ell = k - m + 1$, depending whether the error (5) is evaluated for a forecast started from a background or an analysis, respectively. Equation (6b) first appeared in LB04. Figure 1 gives a schematic representation of the various relevant times involved in the calculation of the formulas above.

A proxy for the infinitesimal error change δe_k^v of the forecast error measure (5) is simply the difference between the error measure calculated for forecasts issued from two consecutive analyses:

$$\delta e_k^v \equiv e_{k|k-m+1}^v - e_{k|k-m}^v. \tag{9}$$

This proxy is used, for example, in the work of Liu and Kalnay (2008) to derive an expression for observation impact that can be more readily calculated in the context of ensemble data assimilation procedures. Similarly, the present work relies on the argument that much of what is obtained with approximations such as those in (6) can be obtained using differences such as in (9). Indeed, in many respects the error difference (9) is more revealing than its infinitesimal counterpart exactly because the former is capable of describing what finite-size changes in the initial conditions, due to the assimilation of observations, imply to the forecast.

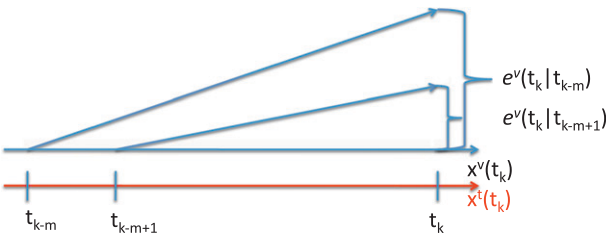

FIG. 1. Schematic representation of time line and the relevant forecast error definitions.

The approximations in (6) express the change in forecast error in the form of an inner product between a "sensitivity" vector and the observation-minus-background residual vector, $\mathbf{d}_{k-m+1|k-m}$. Consequently, it is possible to break down the inner product into its many elements—corresponding to the contributions from the various components of the observing system to the error change. These individual contributions define the corresponding individual observation impacts.

b. A simple alternative forecast error measure

In observation-space, we replace the state-space forecast error measure (5) with the observation-space forecast error defined as the weighted difference between the model-predicted observations $\mathbf{h}_k(\mathbf{x}_{k|\ell}^f)$ and the observation vector $\mathbf{y}_k^o$, at time t_k, for $\ell < k$, that is,

$$e_{k|\ell}^y \equiv [\mathbf{h}_k(\mathbf{x}_{k|\ell}^f) - \mathbf{y}_k^o]^{\mathrm{T}} \mathbf{C}_k [\mathbf{h}_k(\mathbf{x}_{k|\ell}^f) - \mathbf{y}_k^o] = \mathbf{d}_{k|\ell}^{\mathrm{T}} \mathbf{C}_k \mathbf{d}_{k|\ell}, \tag{10}$$

where $\mathbf{C}_k$ is a $p_k \times p_k$ positive semidefinite suitable weighting matrix, and the forecast error is calculated for a forecast started at time $t_\ell < t_k$. The first equality in the expression above is written to emphasize the analogy between this error expression and that in (4) or (5), suggesting that now the verification state is simply replaced with the observations.

The observation-space counterpart of the error difference in (9) is now

$$\delta e_k^y \equiv e_{k|k-m+1}^y - e_{k|k-m}^y. \tag{11}$$

The error-change δe_k^y is a scalar that can be calculated for any portion of the observing system. Error differences such as (9) and (11) identify the improvement (or degradation) in the prediction at time t_k from a forecast started at time t_{k-m+1} over that started at time t_{k-m}. In

other words, these differences are a measure of the impact that assimilating observations at time t_{k-m} has on the forecast at time t_k. Partitioning the impacts into individual contributions from the various components of the observing system is straightforward since (10) and (11) are calculated directly in observation space. On the other hand, when the error measure is defined in state space, as in (5), one must use approximate expressions such as those in (6) to be able to partition the impacts as desired. As long as the verifying observations amount to a robust set, such as being relatively homogenous in time, one should be able to derive statistically reliable observation impacts using the observation-space approach. The same comment applies to observation impacts derived with the state-space approach.

However, in the observation-space measure the observations are taken at verification time (by construction), rather than at initial time, as in the state-space measure. It is the view of the present work that, as long as the observing system is relatively homogenous when comparing the initial and verification times, this is not a serious issue since any statement based on residuals can only be made on the basis of statistics; the statistics must be robust for there to be any confidence in the results. For typical evaluations of, say, observation impact on the 24-h forecasts the verification times are aligned with the initial times and no issues are at stake; if impacts on the 18-h forecasts were to be calculated instead, the observing network differences between initial and verification time would be something to watch for.

A relationship between the error changes (6) and (11) can easily be established. If one expands the forecast error vectors in (10) following similar arguments to those used to derive (6), the residual-based forecast error change (11) can be approximated to first- and second-order accuracy as

$$\delta e_k^y \approx \delta e_k^{y,1} = -\mathbf{d}_{k-m+1|k-m}^{\mathrm{T}} \tilde{\mathbf{K}}_{k-m+1|k-m}^{\mathrm{T}} \nabla_{\mathbf{x}^b} e_{k|k-m}^y, \tag{12a}$$

$$\delta e_k^y \approx \delta e_k^{y,2} = -\frac{1}{2} \mathbf{d}_{k-m+1|k-m}^{\mathrm{T}} \tilde{\mathbf{K}}_{k-m+1|k-m}^{\mathrm{T}} \times [\nabla_{\mathbf{x}^b} e_{k|k-m}^y + \nabla_{\mathbf{x}^a} e_{k|k-m+1}^y], \tag{12b}$$

respectively, where the gradients are calculated as

$$\nabla_{\mathbf{x}^g} e_{k|\ell}^y = -2\mathbf{M}_{g;k,k-m+1}^{\mathrm{T}} \mathbf{H}_{k|\ell}^{\mathrm{T}} \mathbf{C}_k \mathbf{d}_{k|\ell}, \tag{13}$$

for $\mathbf{H}_{k|\ell}$ representing the $p_k \times n$ Jacobian matrix,

$$\mathbf{H}_{k|\ell} = \left. \frac{\partial \mathbf{h}_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_{k|\ell}^g}, \tag{14}$$

of the nonlinear observation operator $\mathbf{h}_k$ linearized about the state $\mathbf{x}_{k|\ell}^g$, for $\ell = k - m + 1$ when $g = b$, and $\ell = k - m$ when $g = a$. Higher-order expressions can be derived following the procedures in E07 or DT09. The approximations in (12) are useful to highlight the fact that both the *state-space* and *observation-space* approaches refer to the same initial perturbation to the forecast, that is, the transpose of the analysis increment vector $\delta \mathbf{x}_{k-m+1|k-m}^a = \tilde{\mathbf{K}}_{k-m+1|k-m} \mathbf{d}_{k-m+1|k-m}$. When using (11) to calculate observation impacts this initial perturbation appears only implicitly. Indeed, the difference being these approximations comes from the gradient vector representing how initial perturbations propagate within the time window starting at t_{k-m} and ending at the verification time t_k. For example, in the first-order approximation in (6a) the gradient vector is $\nabla_{\mathbf{x}^b} e_{k|\ell}^v = 2\mathbf{M}_{g;k,k-m+1}^{\mathrm{T}} \mathbf{T}_k \epsilon_{k|\ell}^v$, whereas in (12a) the gradient vector $\nabla_{\mathbf{x}^b} e_{k|\ell}^y = -2\mathbf{M}_{g;k,k-m+1}^{\mathrm{T}} \mathbf{H}_{k|\ell}^{\mathrm{T}} \mathbf{C}_k \mathbf{d}_{k|\ell}$. This is simply a consequence of the choice of error measures; the negative sign appears simply because $\mathbf{d}_{k|\ell}$ is a residual defined as observation-minus-field instead of a typical error defined as field-minus-observation.

c. *Relationship between the state- and observation-space error measures*

To more accurately establish a relationship between the state-space- and observation-space-based error measures we consider now the definition of the error measures introduced above in a probabilistic sense by referring to their expected values. This goes along with the thinking that any statement made in data assimilation must have a statistical basis. For the sake of argument, we simplify matters by considering the linear case, when the matrices $\mathbf{M}_{g;k,\ell}$, $\tilde{\mathbf{K}}_{k|\ell}$ and $\mathbf{H}_{k|\ell}$ become state independent and can be written more compactly as $\mathbf{M}_{k,k-m}$, $\tilde{\mathbf{K}}_k$ and $\mathbf{H}_k$, respectively. Taking the linear version of the analysis equation (3) at time t_{k-m} and substituting into the linear version of (2), the forecast error of a forecast issued from the analysis at time $t_a = t_{k-m+1}$, and valid at time $t = t_k$, can be expressed as

$$\epsilon_{k|k-m+1}^f = \epsilon_{k|k-m}^f + \mathbf{M}_{k,k-m+1} \tilde{\mathbf{K}}_{k-m+1} \mathbf{d}_{k-m+1|k-m}, \tag{15}$$

which also holds to first-order accuracy for the nonlinear case. The expression relates error in the forecasts valid at the same time, but issued from analyses one cycle apart. Since the only difference between forecasts issued from two consecutive analyses, at times $t_b = t_{k-m}$ and $t_a = t_{k-m+1}$, is the observations assimilated in the more recent analysis, the error in the forecasts differ by how much the incremental difference introduced by the

assimilated data propagates in time, that is, by how the term $\delta\mathbf{x}^a_{k-m+1|k-m}$ evolves under the dynamics $\mathbf{M}_{k,k-m+1}$.

Taking the expected mean of the true *state-space* measure introduced in (4) gives

$$\begin{aligned}\langle e_{k|k-m}\rangle &\equiv \langle(\epsilon^f_{k|k-m})^{\mathrm{T}}\mathbf{T}_k(\epsilon^f_{k|k-m})\rangle\\ &= \mathrm{Tr}\{\mathbf{T}_k[\langle(\epsilon^f_{k|k-m})(\epsilon^f_{k|k-m})^{\mathrm{T}}\rangle]\}\\ &= \mathrm{Tr}\{\mathbf{T}_k\mathbf{P}^f_{k|k-m}\},\end{aligned} \tag{16}$$

where $\mathbf{P}^f_{k|k-m} \equiv \langle\epsilon^f_{k|k-m}(\epsilon^f_{k|k-m})^{\mathrm{T}}\rangle$ is the forecast error covariance matrix, with $\langle\cdot\rangle$ representing the expectation operator, $\mathrm{Tr}(\cdot)$ stands for the trace operator, and we used the trace property $\mathrm{Tr}(\mathbf{A}^{\mathrm{T}}\mathbf{TB}) = \mathrm{Tr}(\mathbf{TBA}^{\mathrm{T}})$ for arbitrary matrices $\mathbf{A}$ and $\mathbf{B}$ of dimension $n \times p$. Therefore, the expected mean of the forecast error *change* calculated for forecasts from two consecutive analyses is

$$\langle\delta e_k\rangle = \mathrm{Tr}\{\mathbf{T}_k\Delta\mathbf{P}^f_k\}, \tag{17}$$

where $\Delta\mathbf{P}^f_k \equiv \mathbf{P}^f_{k|k-m+1} - \mathbf{P}^f_{k|k-m}$ is the difference between the two forecast error covariances corresponding to the two lagged forecasts.

As suggested above, the impact of observations can also be evaluated by examining the expectation of the *observation-space* measure, that is,

$$\begin{aligned}\langle e^y_{k|k-m}\rangle &\equiv \langle(\mathbf{d}_{k|k-m})^{\mathrm{T}}\mathbf{C}_k(\mathbf{d}_{k|k-m})\rangle\\ &= \mathrm{Tr}\{\mathbf{C}_k[\langle(\mathbf{d}_{k|k-m})(\mathbf{d}_{k|k-m})^{\mathrm{T}}\rangle]\}\\ &= \mathrm{Tr}\{\mathbf{C}_k\boldsymbol{\Gamma}_{k|k-m}\},\end{aligned} \tag{18}$$

where we introduce the $p_k \times p_k$ residual error covariance matrix:

$$\boldsymbol{\Gamma}_{k|k-m} \equiv \langle\mathbf{d}_{k|k-m}\mathbf{d}^{\mathrm{T}}_{k|k-m}\rangle = \mathbf{H}_k\mathbf{P}^f_{k|k-m}\mathbf{H}^{\mathrm{T}}_k + \mathbf{R}_k. \tag{19}$$

Using this observation-space measure the change in forecast error due to the assimilation of observations becomes

$$\langle\delta e^y_k\rangle = \mathrm{Tr}\{\mathbf{H}^{\mathrm{T}}_k\mathbf{C}_k\mathbf{H}_k\Delta\mathbf{P}^f_k\}, \tag{20}$$

for $\delta e^y_k \equiv \delta e^y_{k|k-m+1} - \delta e^y_{k|k-m}$.

After some algebra, the difference between the two forecast error covariances defining $\Delta\mathbf{P}^f_k$ is shown in appendix A to be

$$\Delta\mathbf{P}^f_k = \mathbf{M}_{k,k-m+1}(\mathbf{P}^a_{k-m+1|k-m+1} - \mathbf{P}^f_{k-m+1|k-m})\mathbf{M}^{\mathrm{T}}_{k,k-m+1}, \tag{21}$$

with no assumptions made on optimality and model error. Furthermore, recall that at any time t_k, the analysis error covariance can be written as [e.g., see Cohn et al. (1994), their Eq. (2.33)]:

$$\mathbf{P}^a_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}^f_{k|k-1} + \Delta\mathbf{P}^a_{k|k}, \tag{22}$$

where the increment matrix $\Delta\mathbf{P}^a_{k|k} = \Delta\mathbf{K}_k\boldsymbol{\Gamma}_k\Delta\mathbf{K}^{\mathrm{T}}_k$ incorporates all the suboptimality in the analysis error covariance, with $\Delta\mathbf{K}_k \equiv \tilde{\mathbf{K}}_k - \mathbf{K}_k$ being the difference between a general gain matrix $\tilde{\mathbf{K}}_k$ and the optimal Kalman gain matrix $\mathbf{K}_k$, with

$$\mathbf{K}_k = \mathbf{P}^f_{k|k-1}\mathbf{H}^{\mathrm{T}}_k\boldsymbol{\Gamma}^{-1}_k. \tag{23}$$

From (22) it follows that,

$$\mathbf{P}^a_{k|k} - \mathbf{P}^f_{k|k-1} = -\mathbf{K}_k\boldsymbol{\Gamma}_k\mathbf{K}^{\mathrm{T}}_k + \Delta\mathbf{P}^a_{k|k}, \tag{24}$$

is negative semidefinite in the optimal case, when $\Delta\mathbf{P}^a_{k|k} \overset{\mathrm{opt}}{=} \mathbf{0}$, and therefore so is the forecast error covariance difference in (21), $\Delta\mathbf{P}^f_k < \mathbf{0}$. Applying this result at time t_{k-m+1} and combining it with (17) and (21) shows that in the optimal case the expected forecast error change $\langle\delta e_k\rangle$ is guaranteed to be nonpositive,

$$\begin{aligned}\langle\delta e_k\rangle \overset{\mathrm{opt}}{=} -\mathrm{Tr}\{&\mathbf{T}_k\mathbf{M}_{k,k-m+1}\mathbf{K}_{k-m+1}\boldsymbol{\Gamma}_{k-m+1}\\ &\times\mathbf{K}^{\mathrm{T}}_{k-m+1}\mathbf{M}^{\mathrm{T}}_{k,k-m+1}\} \overset{\mathrm{opt}}{\leq} 0,\end{aligned} \tag{25}$$

since the kernel inside the trace operator is positive semidefinite, with equality holding when $\mathbf{T}_k = \mathbf{0}$. In the optimal case, and in the expected mean sense, assimilation of observations is guaranteed to *reduce* forecast errors.

Similarly, since in the optimal case $\Delta\mathbf{P}^f_k < \mathbf{0}$, the observation-space-expected forecast error change in (20) is also nonpositive,

$$\begin{aligned}\langle\delta e^y_k\rangle \overset{\mathrm{opt}}{=} -\mathrm{Tr}\{&\mathbf{H}^{\mathrm{T}}_k\mathbf{C}_k\mathbf{H}_k\mathbf{M}_{k,k-m+1}\mathbf{K}_{k-m+1}\boldsymbol{\Gamma}_{k-m+1}\\ &\times\mathbf{K}^{\mathrm{T}}_{k-m+1}\mathbf{M}^{\mathrm{T}}_{k,k-m+1}\} \overset{\mathrm{opt}}{\leq} 0,\end{aligned} \tag{26}$$

corroborating again that, in the expected mean sense, assimilation of observations leads to forecast error *reduction*.

A particular choice of weighting matrix $\mathbf{T}_k$, namely $\overline{\mathbf{T}}_k = \mathbf{H}^{\mathrm{T}}_k\mathbf{C}_k\mathbf{H}_k$, leads to

$$\langle\delta e_k(\mathbf{T}_k = \overline{\mathbf{T}}_k)\rangle \overset{\mathrm{opt}}{=} \langle\delta e^y_k\rangle, \tag{27}$$

that is, in the linear optimal case one can choose the forecast error measure in state space to obtain the same

total expected error as that obtained in observation space. Conversely, since the observation space is normally smaller than the state space, one cannot choose $\mathbf{C}_k$ to obtain the same total expected forecast error as obtained in state space.

d. The role of the verification

One can also inquire about the relationship between the change in the true expected forecast error $\langle \delta e_k \rangle$ and that defined with respect to a verification state $\langle \delta e_k^v \rangle$. Sticking with the linear, but not necessarily optimal, case we show in appendix B that the relationship between these expected errors can be written as

$$\langle \delta e_k^v \rangle = \langle \delta e_k \rangle - 2\mathrm{Tr}[\tilde{\mathbf{K}}_{k-m+1}^{\mathrm{T}} \mathbf{M}_{k,k-m+1}^{\mathrm{T}} \mathbf{T}_k \langle \epsilon_k^v \mathbf{d}_{k-m+1|k-m}^{\mathrm{T}} \rangle], \tag{28}$$

where $\epsilon_k^v \equiv \mathbf{x}_k^v - \mathbf{x}_k^t$ is the true error in the verification state. Replacement of the truth with a verification state results in a term involving the correlation between the error in the verification and the background residuals.

More specifically, when the verification state is taken to be the analysis, $\mathbf{x}_k^v = \mathbf{x}_{k|k}^a$, a typical choice in practical applications, the expression above can be worked out in more detail to give

$$\delta e_k^{v=a} = \langle \delta e_k \rangle - 2\mathrm{Tr}\left[\tilde{\mathbf{K}}_{k-m+1}^{\mathrm{T}} \mathbf{M}_{k,k-m+1}^{\mathrm{T}} \mathbf{T}_k \times \left(\mathbf{M}_{k,k-m+1} \Delta \mathbf{K}_{k-m+1} \mathbf{\Gamma}_{k-m+1} + \sum_{j=0}^{m-2} \mathbf{M}_{k,k-j} \tilde{\mathbf{K}}_{k-j} \langle \mathbf{d}_{k-j|k-j-1} \mathbf{d}_{k-m+1|k-m}^{\mathrm{T}} \rangle \right)\right]. \tag{29}$$

This relates the *perceived* error change, calculated using the analysis for verification, with the *actual* error change. To arrive at this result (see appendix C) one requires typical assumptions of linear filtering theory (e.g., Jazwinski 1970, chapter 7): that model errors be uncorrelated with observation errors, and that forecast errors be uncorrelated with observation errors for all times larger than the time the forecast begins. Two terms prevent the perceived error change $\langle \delta e_k^{v=a} \rangle$ from equaling the actual error change. One is the first term in the trace expression involving the difference $\Delta \mathbf{K}_{k-m+1}$ between the suboptimal and optimal gains. The other is the second term in the trace expression involving the cross covariances of the various observation-minus-background residuals between the analysis time t_{k-m+2} and all times up to the verification time t_k. It is only in the optimal case that both these terms vanish: the first, for obvious reasons, $\Delta \mathbf{K}_k = \mathbf{0}$; the second, because the sequence of observation-minus-background *residuals*, $\{\mathbf{d}_{k|k-1}\}$, becomes the sequence of *innovations*, which is white in time, and with all the time cross covariances becoming zero[1] (Kailath 1968; Daley 1992; see also Anderson and Moore 1979, their section 5.3). Therefore, in the *optimal case* and in the *expected mean sense*, verifying against the analysis is the same as verifying against the truth when it comes to evaluating the expected forecast error change under consideration.

Of more practical interest is the question of what happens when, using the observation-space measure (10), the observation vector is replaced with a verification state $\mathbf{x}_k^v$ projected onto the observation space by the observation operator, that is, when $\mathbf{y}_k^o$ is replaced with $\mathbf{h}_k(\mathbf{x}_k^v)$, or in the linear case, $\mathbf{H}_k \mathbf{x}_k^v$. Following similar steps to those taken in appendix B to derive (28), but now applying the rational to the observation-space measure, it can be shown that the observation-space error measure change takes the following form:

$$\langle \delta e_k^{y=v} \rangle = \langle \delta e_k^y \rangle - 2\mathrm{Tr}[\tilde{\mathbf{K}}_{k-m+1}^{\mathrm{T}} \mathbf{M}_{k,k-m+1}^{\mathrm{T}} \mathbf{H}_k^{\mathrm{T}} \mathbf{C}_k \mathbf{H}_k \times \langle \epsilon_k^v \mathbf{d}_{k-m+1|k-m}^{\mathrm{T}} \rangle], \tag{30}$$

when a verification state replaces the observations. As before, the perceived expected observation error change $\langle \delta e_k^{y=v} \rangle$ differs from the actual expected error change $\langle \delta e_k^y \rangle$ by a similar term to that found when examining this relationship in state space, that is, the cross correlation between the verification error and the observation-minus-background error. Analogously, as one might expect, when the verification is taken to be the underlying analysis, this relationship becomes

$$\langle \delta e_k^{y=a} \rangle = \langle \delta e_k^y \rangle - 2\mathrm{Tr}\left[\tilde{\mathbf{K}}_{k-m+1}^{\mathrm{T}} \mathbf{M}_{k,k-m+1}^{\mathrm{T}} \mathbf{H}_k^{\mathrm{T}} \mathbf{C}_k \mathbf{H}_k \times \left(\mathbf{M}_{k,k-m+1} \Delta \mathbf{K}_{k-m+1} \mathbf{\Gamma}_{k-m+1} + \sum_{j=0}^{m-2} \mathbf{M}_{k,k-j} \tilde{\mathbf{K}}_{k-j} \langle \mathbf{d}_{k-j|k-j-1} \mathbf{d}_{k-m+1|k-m}^{\mathrm{T}} \rangle \right)\right] \tag{31}$$

(this can be derived following the rational of appendix C). Analogously to (29), equality here only holds in the optimal case, when $\Delta \mathbf{K}_k = \mathbf{0}$ and the sequence of observation-minus-background residuals becomes the innovation sequence. Unlike the expressions (28) and (29),

[1] Note that the cross variances, the cross terms calculated for the same time, are not zero, but they also do not appear inside the summation sign.

the observation-space expressions (30) and (31) can, in principle, be verified in practice. Still, one thing to notice related to expression (30) is that it assumes the verifying observations to be fully independent of the analyses. In many practical data assimilation applications this assumption is violated because of procedures to perform observation bias correction.

e. General remarks

Both the state- and observation-space measures considered above involve what are in principle arbitrary weighting matrices $\mathbf{T}_k$ and $\mathbf{C}_k$, respectively. One particular choice has special meaning: when they become the inverse error covariance matrices of the corresponding underlying error. That is,

$$\mathbf{T}_k = [\mathbf{P}^f_{k|k-m}]^{-1}, \tag{32a}$$

$$\mathbf{C}_k = [\mathbf{H}_{k|k-m}\mathbf{P}^f_{k|k-m}\mathbf{H}^{\mathrm{T}}_{k|k-m} + \mathbf{R}_k]^{-1}. \tag{32b}$$

When $\mathbf{T}_k$ is defined as in (32a), the forecast error (4) becomes a measure of the accuracy of the fit of the forecasts to the truth. In particular, when $m = 1$, the forecast error covariance becomes the background error covariance, $\mathbf{B} \equiv \mathbf{P}^f_{k|k-1}$, and the forecast error measure (4) becomes the background error term present in the cost function of variational data assimilation methods (e.g.: Courtier and Talagrand 1987; Lewis et al. 2006, their chapter 20).

Similarly, when $\mathbf{C}_k$ is defined as in (32b), the forecast error (10) becomes a measure of the accuracy of the fit of the model-predicted observations to the actual observations. When $m = 1$, the weight matrix $\mathbf{C}_k$ becomes the inverse of the familiar residual (''innovation'') error covariance matrix used in, say, observation-space-based analysis procedures (e.g., Cohn et al. 1998). In practice, this weighting factor can come from prescribed statistics or from an estimate obtained directly from sample error covariances constructed from the observation residuals themselves; similar to, say, the treatment of consistent or inconsistent error statistics of Lupu et al. (2011).

The case $m = 1$ is important, and it is the only measure that can directly be associated with the data assimilation *cycle*. In this case, the observation-space measure in (10) reduces to

$$\begin{aligned}\delta e^y_k &= e^y_{k|k} - e^y_{k|k-1}\\ &= [\mathbf{h}_k(\mathbf{x}^a_{k|k}) - \mathbf{y}^o_k]^{\mathrm{T}}\mathbf{C}_k[\mathbf{h}_k(\mathbf{x}^a_{k|k}) - \mathbf{y}^o_k]\\ &\quad - [\mathbf{h}_k(\mathbf{x}^b_{k|k-1}) - \mathbf{y}^o_k]^{\mathrm{T}}\mathbf{C}_k[\mathbf{h}_k(\mathbf{x}^b_{k|k-1}) - \mathbf{y}^o_k],\end{aligned} \tag{33}$$

where $\mathbf{x}^f_{k|k-1} = \mathbf{x}^b_{k|k-1}$ and $\mathbf{x}^f_{k|k} = \mathbf{x}^a_{k|k}$. Thus, the expression above is simply the difference between the weighted OMA squared residual with the corresponding OMB residual. This quantity is readily available in any analysis system. The particularly convenient choice,

$$\mathbf{C}_k = \mathbf{R}^{-1}_k, \tag{34}$$

turns the difference above into the difference between the final and initial ''J_o'' term of the typical variational assimilation problem.

Furthermore, simple algebra shows that the expected mean of the measure in (33), with weights from (34), relates to a version of the degrees of freedom for signal (DFS) diagnostic that uses residuals based on posterior estimates (analysis residuals). That is,

$$\begin{aligned}\langle\delta e^y_k\rangle &\approx -f_k - \langle[\mathbf{h}_k(\mathbf{x}^a_{k|k}) - \mathbf{h}_k(\mathbf{x}^b_{k|k-1})]^{\mathrm{T}}\mathbf{R}^{-1}_k[\mathbf{y}^o_k - \mathbf{h}_k(\mathbf{x}^b_{k|k-1})]\rangle\\ &\approx -f_k - \langle\mathbf{d}^{\mathrm{T}}_{k|k-1}(\mathbf{\Gamma}^{-1}_{k|k-1}\mathbf{H}_{k|k-1}\mathbf{P}^f_{k|k-1}\mathbf{H}^{\mathrm{T}}_{k|k-1}\mathbf{R}^{-1}_k)\mathbf{d}_{k|k-1}\rangle + O(\Delta\mathbf{K}_{k|k-1})\\ &\overset{\text{opt}}{=} -f_k - \mathrm{Tr}(\mathbf{H}_{k|k-1}\mathbf{P}^f_{k|k-1}\mathbf{H}^{\mathrm{T}}_{k|k-1}\mathbf{R}^{-1}_k),\end{aligned} \tag{35}$$

where

$$f_k \equiv \langle[\mathbf{h}_k(\mathbf{x}^a_{k|k}) - \mathbf{h}_k(\mathbf{x}^b_{k|k-1})]^{\mathrm{T}}\mathbf{R}^{-1}_k[\mathbf{y}^o_k - \mathbf{h}_k(\mathbf{x}^a_{k|k})]\rangle \tag{36}$$

is the DFS as derived in Lupu et al. (2011) [see Eq. (10) in that work]; the second approximate expression in (35) follows from noticing that $\mathbf{h}_k(\mathbf{x}^a_{k|k}) - \mathbf{h}_k(\mathbf{x}^b_{k|k-1}) \approx \mathbf{H}_{k|k-1}\tilde{\mathbf{K}}_{k|k-1}[\mathbf{y}^o_k - \mathbf{h}_k(\mathbf{x}^b_{k|k-1})]$; and the third equality follows from applying the trace operator and invoking optimality, when $\Delta\mathbf{K}_{k|k-1} = \mathbf{0}$. Thus, the observation-space measure for $m = 1$ is directly related to a DFS-like diagnostic for a suitable choice of $\mathbf{C}_k$. This is not very surprising: one should expect various diagnostics involving the same quantities (i.e., OMBs and OMAs), to provide similar information. Further insight is gained by examining the last equality, optimal expression, in (35). Since the second term on the right-hand side is positive, we see that when the overall impact under the observation-space measure indicates that a set of observations degrades the forecasts, that is when $\langle\delta e^y_k\rangle$ is positive, the corresponding DFS will have a negative sign; observations with negative $\langle\delta e^y_k\rangle$, shown to contribute to

improve the forecast under the observation-space measure, may have DFS of either sign. In other words, if the sign of the measure is used to define impact, both measures agree when observations degrade the forecasts, but they might disagree on which observations contribute to improve forecasts. This apparent contradiction relates to the uncertainty associated in these statistical measures. If we accept both measures in question as reasonable ways of determining the impact of observations, then the degree to which their results can be trusted is measured by the second term on the right-hand side of (35). Essentially, the term $\mathrm{Tr}(\mathbf{H}_{k|k-1}\mathbf{P}^{f}_{k|k-1}\mathbf{H}^{\mathrm{T}}_{k|k-1}\mathbf{R}^{-1}_{k})$ provides an error bar for the reliability of impacts calculated with either method; impacts within the error bar can be considered neutral.

As hinted above, it is the view of this author that the only measures of observation impact that have direct implications to data assimilation are those calculated from (33) or, analogously, from (36). It is hard to see how general measures such as those in (9) and (11), for $m > 1$, can be used to aid the *cycling* of data assimilation procedures. The sensitivities of the forecasting model beyond the integration time window of a typical assimilation cycle do not participate in the formulation of data assimilation procedures. It is a basic fact of data assimilation, and generally estimation theory, that the statistics of the background- and analysis-observation residuals teach us what can be learned of the behavior of the data assimilation system and how it uses the observations. This is the reason why maximum likelihood procedures (e.g., Dee and da Silva 1999) and chi-square-based measures (e.g., Ménard et al. 2000) are solely based on these residuals (e.g., Desroziers et al. 2005b). This is not to say there is nothing to be learned from what has become known as ''observation impacts''; it is just that it is unclear how information obtained from such diagnostics relates to the cycling of data assimilation. Indeed, if we were to identify components of the observing system leading to particularly undesirable features to, say, the 24-h forecasts and a decision was made to change the manner in which these components are handled during the assimilation *cycle*, it is certain that such changes would lead to unpredictable results in the *cycle*. In other words, one would not be able to tell in advance if the changes would lead to improvement or degradation to the quality of the assimilation cycle, much less to the quality of future 24-h forecasts.

3. Illustrative results

This section presents a brief illustration of the main points made in the previous section. GEOS-5 (Rienecker et al. 2008) is used for this purpose. GEOS-5 assimilates observations using the incremental analysis update technique of Bloom et al. (1996). It consists of a global atmospheric model developed at Goddard and an analysis system developed jointly by the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Prediction (NCEP) and the NASA Global Modeling and Assimilation Office. The GEOS-5 GCM retains an updated version of the finite-volume hydrostatic dynamical core (Lin 2004) from its predecessor GEOS-4. The GEOS-5 GCM is built under the infrastructure of the Earth System Modeling Framework (Collins et al. 2005) used to couple together various physics packages including a modified version of the Relaxed Arakawa–Schubert convective parameterization scheme of Moorthi and Suarez (1992), the catchment-based hydrological model of Koster et al. (2000), the multilayer snow model of Stieglitz et al. (2001), and the radiative transfer model of Chou and Suarez (1999). Furthermore, the GCM is accompanied by its adjoint model (ADM), which is essentially the ADM of the finite-volume dynamical core of GEOS-4 (Giering et al. 2005), with added vertical diffusion and a polar filter (Errico et al. 2007).

The GEOS-5 analysis component consists of the GSI system. The GSI implements a three-dimensional variational data assimilation (3D-Var) using the incremental approach of Courtier et al. (1994) for minimization and the preconditioning strategy of Derber and Rosati (1989). The background error covariance is implemented as a series of recursive filters producing nearly Gaussian and isotropic correlation functions (Wu et al. 2002). Satellite radiances are processed using the Community Radiative Transfer Model (CRTM; Kleespies et al. 2004) and the online bias-correction procedure of Derber and Wu (1998). Furthermore, the version of GSI used in the experiments here includes the adjoint capability of Trémolet (2007, 2008). This adjoint differs from its previous incarnation in Zhu and Gelaro (2008) in that it is not a line-by-line adjoint, but rather it is derived from a swap of operations used in the forward GSI. Combining the GSI adjoint with the GCM adjoint, GEOS-5 has all the ingredients necessary to calculate the adjoint-based approximations to observation impacts discussed earlier (see also Gelaro et al. 2010).

The main experiment period considered here is the same as that considered to obtain some of the results in DT09. Similarly to other related works, we study mainly the impact of observations on the 24-h forecasts. Twenty-four-hour forecasts and 30-h forecasts are issued for all 0000 and 1800 UTC times, respectively, for the month of August 2007, and second-order accurate observation impacts are calculated over 0000 UTC times. The assimilation, forecasts and, when applicable, adjoint

integrations, are performed on a horizontal grid of 2.5° × 2° resolution, with 72 hybrid vertical levels. Todling and Trémolet (2008) have shown that the *relative* impact among the various observing systems is largely independent of resolution, rendering conclusions from the experiments here general.

a. Role of norm

1) STATE-SPACE APPROACH

To illustrate the role played by the choice of weight matrices in the forecast error measures of the previous section, we consider three choices of norms when using the state-space approach. The first two are total energy measures. Following Errico et al. (2007), the inner product between two vectors, $\mathbf{x}_1$ and $\mathbf{x}_2$, involving perturbations in the zonal and meridional components of the wind, temperature, and surface pressure, is calculated using either one of the following expressions:

$$e_t \equiv \mathbf{x}_1^{\mathrm{T}}\mathbf{T}_t\mathbf{x}_2 = \frac{1}{2}\sum_{i,j,k}\Delta H_{i,j}\Delta\sigma_{i,j,k} \times\left[u_1'u_2' + v_1'v_2' + \frac{c_p}{T_r}T_1'T_2' + \frac{RT_r}{p_r^2}p_{s1}'p_{s2}'\right]_{i,j,k}, \tag{37a}$$

$$e_v \equiv \mathbf{x}_1^{\mathrm{T}}\mathbf{T}_v\mathbf{x}_2 = \frac{1}{2}\sum_{i,j,k}\Delta H_{i,j}\Delta z_{i,j,k} \times\left[u_1'u_2' + v_1'v_2' + \frac{c_p}{T_r}T_1'T_2' + \frac{RT_r}{p_r^2}p_{s1}'p_{s2}'\right]_{i,j,k}, \tag{37b}$$

where $\Delta H_{i,j}$ is a horizontal gridbox weight and the distinction between the two norms is in how they weight the fields in the vertical, with $\Delta\sigma_{i,j,k}$ and $\Delta z_{i,jk}$ being fractional weights, respectively, defined as

$$\Delta\sigma_{i,j,k} = \frac{\Delta p_{i,j,k}}{p_{s,i,j} - p_t}, \tag{38a}$$

$$\Delta z_{i,j,k} = \frac{\Delta \ln p_{i,j,k}}{\ln p_{s,i,j} - \ln p_t}. \tag{38b}$$

For consistency with Errico et al. (2007), we refer to (37a) and (37b) as the ET-norm and EV-norm, respectively. The physical scaling coefficients $c_p = 1004.6$ J kg^{-1} K^{-1}, $R = 287.04$ J kg^{-1} K^{-1}, $T_r = 280$ K, and $p_r = 1000$ hPa, are the specific heat at constant pressure, the gas constant of dry air, and a reference temperature and pressure, respectively. Detailed discussions of the applicability of these two flavors of the total energy norm appear in Lewis et al. (2001) and in Errico et al. (2007). Here, we simply note that the energy in perturbations in the troposphere are emphasized when the fractional mass weights in (38a) are used; whereas the energy of perturbations mainly concentrated in the midtroposphere and stratosphere are emphasized when the fractional distance weights in (38b) are used instead. This is illustrated in Fig. 2 where the weights in (38) are displayed as a function of the 72 vertical levels of GEOS-5 for a grid point where $p_s = 1000$ hPa. The thin curve is for the fractional mass weight (38a), and the thick curve is for the fractional distance weight (38b). In the former, the weights have a stepwise increase in the troposphere up to 300 hPa, above which level they decrease rapidly. The opposite happens in case of the latter fractional weights, where they are comparatively small in the troposphere and increase rapidly in the stratosphere and mesosphere.

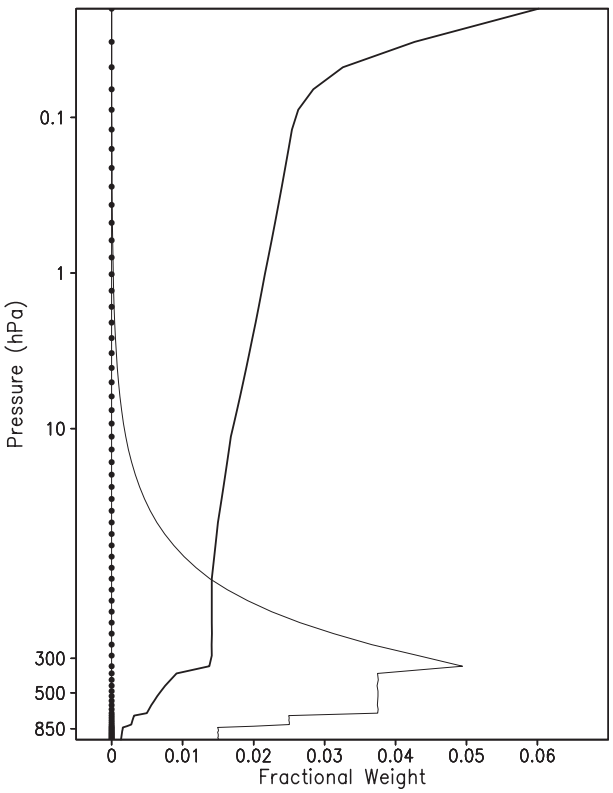

FIG. 2. The fractional vertical weights $\Delta\sigma$ (thin curve) and Δz (thick curve) used for calculating the ET- and EV-norms, respectively. The dotted vertical line indicates the model levels. All calculated at a point where $p_s = 1000$ hPa; the model-top pressure is 0.01 hPa. [Similar to Fig. 1 of Errico et al. (2007)].

The works of LB04, E07, GZE07, and DT09 all use the ET-norm to calculate observation impacts. Moreover, as mentioned in the introduction, these works also use a projection operator that is unit for all grid points roughly below 100 hPa, and zero above that. We have

compared state-space observation impacts using the ET-norm with and without this projection operator. The impacts change somewhat, but as expected, given that the ET-norm weights decay rapidly above the top of the projection operator box, 100 hPa, the differences in observation impact are not significant (not shown). In the present work we use a projection operator that only amounts to excluding the five top layers, mainly to avoid letting the EV-norm include anything too close to the model top, knowing that this changes essentially nothing when using the ET-norm.

The third norm considered when calculating observation impacts with the state-space approach follows from our considerations regarding the observation-space approach. That is, (6b) is evaluated for weighting matrices $\mathbf{T}_k$ of the following form:

$$\mathbf{T}_k = \mathbf{H}_k^{\mathrm{T}} \mathbf{R}_k^{-1} \mathbf{H}_k . \tag{39}$$

Contrary to the ET- and EV-norms, this amounts to a time-dependent choice of weights. We have no indication that this affects in any significant way the conclusions drawn from the study that follows, and time dependency is consistent with the arguments of the previous section. This choice of norm is a simple attempt to have the state-space and the observation-space approaches use a similar weighting factor, when the latter employs (34).

Notice that using the weight matrices in (37) lead to observation impacts that have units of energy, while using the weights in (39) or, as we will consider later, those in (34), lead to nondimensional impacts. To facilitate comparing the various norms and approaches we mostly show results as, nondimensional, *fractional* observation impacts. The fractional impact $\delta e^m_{\%}(i)$ of the ith observation set, in the measure m, is defined by

$$\delta e^m_{\%}(i) = 100 \times \frac{\delta e^m(i)}{\sum_i \delta e^m(i)}, \tag{40}$$

where the summation runs over all observing sets in the figure under consideration.

Figure 3 shows fractional observation impacts for each measure of interest and various observing systems: (Fig. 3a) the state-space ET-norm, (Fig. 3b) the state-space EV norm, (Fig. 3c) the state-space $\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}$ weights (labeled R-norm), and (Fig. 3d) the observation-space $\mathbf{R}^{-1}$ weights (labeled R-omf). The breakdown of the observing system used during the period of the experiment is described in the figure caption. A preliminary version of the results in this figure appeared in Todling (2009). Let us concentrate first on Figs. 3a–c. These provide a direct assessment of the effect of the choice of norm in the "traditional" state-space observation impact calculation. Between Figs. 3a,b, the most striking difference shows in the role played by the radiosonde (RaobDsnd) and satellite observations: under the ET-norm radiosondes visibly have a comparable role to that played by the Advanced Microwave Sounding Unit-A (AMSU-A) and the Atmospheric Infrared Sounder (AIRS), whereas under the EV-norm AMSU-A dominates, while the role of radiosondes and AIRS becomes comparable. The fractional impact of the High Resolution Infrared Radiation Sounder (HIRS) and the aircraft is flipped under these two norms, with the aircraft being significant in a largely tropospheric measure and HIRS becoming more dominate under a stratospheric measure. When the weights are based on the inverse of the observation error covariance matrix (Fig. 3c), the global fractional impacts seem to resemble more closely those obtained with the tropospheric ET-norm, though there are differences: aircraft observations show as being slightly more significant than AMSU-A and AIRS, and almost comparable to radiosondes.

Interestingly, the seemingly negligible fractional impact of AMSU-B when using either the ET- or EV-norms, is no longer so under the R-norm. This is attributed to the fact that both the ET- and EV-norms in (37) are dry measures and consequently provide a zero input gradient in specific humidity to the adjoint (forecast) sensitivity integration. The lack of any convective parameterization in the model adjoint results in the integrated sensitivity to remain zero. Therefore, the tiny fractional impact appearing in Figs. 3a,b for AMSU-B are simply a result of how sensitive the moist channels are to changes in the temperature field within the analysis itself. On the other hand, under the R-norm, the input gradient to the forecast model adjoint is no longer zero in its specific humidity term; the model adjoint simply advects the initial gradient 24 h backward and generates a nonzero forecast sensitivity in specific humidity that is then fed into the analysis adjoint; this, in turn, results in a nonnegligible contribution from the AMSU-B, moist-sensitive, channels (Fig. 3c).

These three measures of observation impact on the 24-h forecasts corroborate the, not surprising, significance of AMSU-A found elsewhere (LB04; E07; GZE07; Gelaro et al. 2010). Though we stress that the aim of the present work is not to evaluate and rank the observing system, it is still instructive to look more closely at AMSU-A to see exactly where, for example, the dramatic increase in fractional impact comes from when using the EV-norm. Figure 4 shows the breakdown of the fractional impacts for this instrument under the various norms. Results include accumulated impacts from *NOAA-15*, *-16*, and *-18*, as well as from *Aqua*. With the

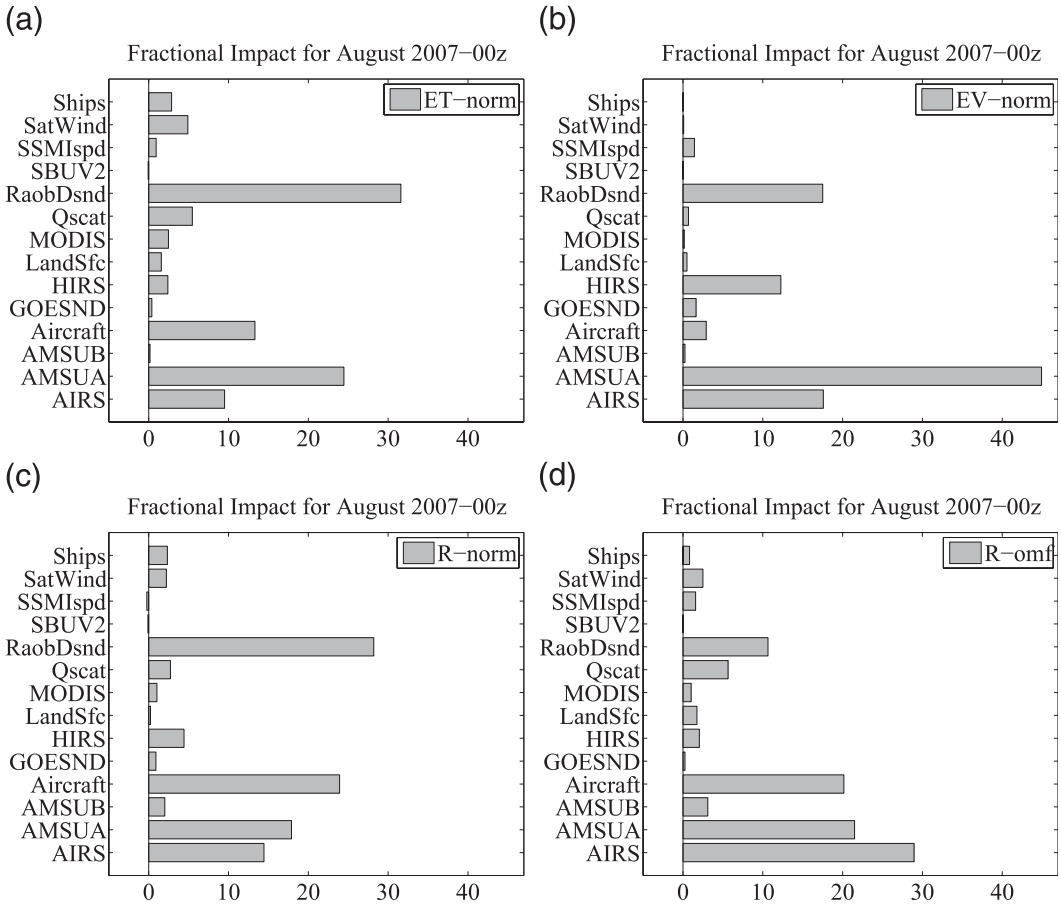

FIG. 3. The fractional observation impact (%) for various instruments for all 31, 24-h forecasts of GEOS-5 valid at 0000 UTC August 2007. The fractions are calculated with respect to total 31-day impact as defined in each case: (a) adjoint-based, total tropospheric energy norm; (b) adjoint-based, total stratospheric energy norm; (c) adjoint-based with forecast errors normalized with inverse of observation error covariance; and (d) OMF-based approach. Results from (a) and (b) first appeared in Todling (2009). The abbreviations along the vertical axis stand for observations of ships and buoy temperature, winds specific humidity, and near-surface pressure (Ship); cloud-drift winds (SatWind); Special Sensor Microwave Imager wind speeds (SSMIspd); Backscatter Ultraviolet Instrument total column ozone (SBUV2); radiosonde and dropsonde temperature, winds, specific humidity (RaobDsnd); Scatterometer winds (Qscat); Moderate Resolution Imaging Spectroradiometer (MODIS) clear-sky and water vapor winds; land observations of temperature, winds, surface pressure, and specific humidity (LandSfc); radiances from the High Resolution Infrared Radiation Sounder (HIRS) in (3) from *NOAA-16* and *-17*; radiances from the Geostationary Operational Environmental Satellites (GOESND); aircraft temperature and winds (Aircraft); radiances from the Advanced Microwave Sounding Unit-A (AMSU-A) on the *NOAA-15*, *-16*, and *-18*; Advanced Microwave Sounding Unit-B (ANSU-B) from *NOAA-15*, *-16*, and *-17*, as well as on *Aqua*; and NASA *Aqua*, Atmospheric Infrared Sounders radiances (AIRS).

EV-norms (Fig. 4b) the stratospheric channels 13 and 14 have a very large fractional impact, and contribute considerably to the large accumulated fractional impact of AMSU-A seen in Fig. 3b. We should not be deceived by the seemingly small fractional impacts of the mid-tropospheric channels implied in Fig. 4b; the contribution from these channels is still rather significant under this norm (see below). Examination of Fig. 4c, showing results when the R-norm is employed seem to indicate the use of the stratosphere-peaking channels to be significant, though results for the troposphere-peaking channels are comparable to when the ET-norm is used.

To stress the point that when using the EV-norm the impact of channels peaking in the midtroposphere is similar to the impact seen by other norms, we show in Fig. 5 the fractional impact on the 24-h forecasts of channel 6 on *NOAA-18*, summed over all 0000 UTC analyses of August 2007. Results for all measures are displayed in the figure. Neutral fractional impacts are shaded gray, with red indicating regions of positive

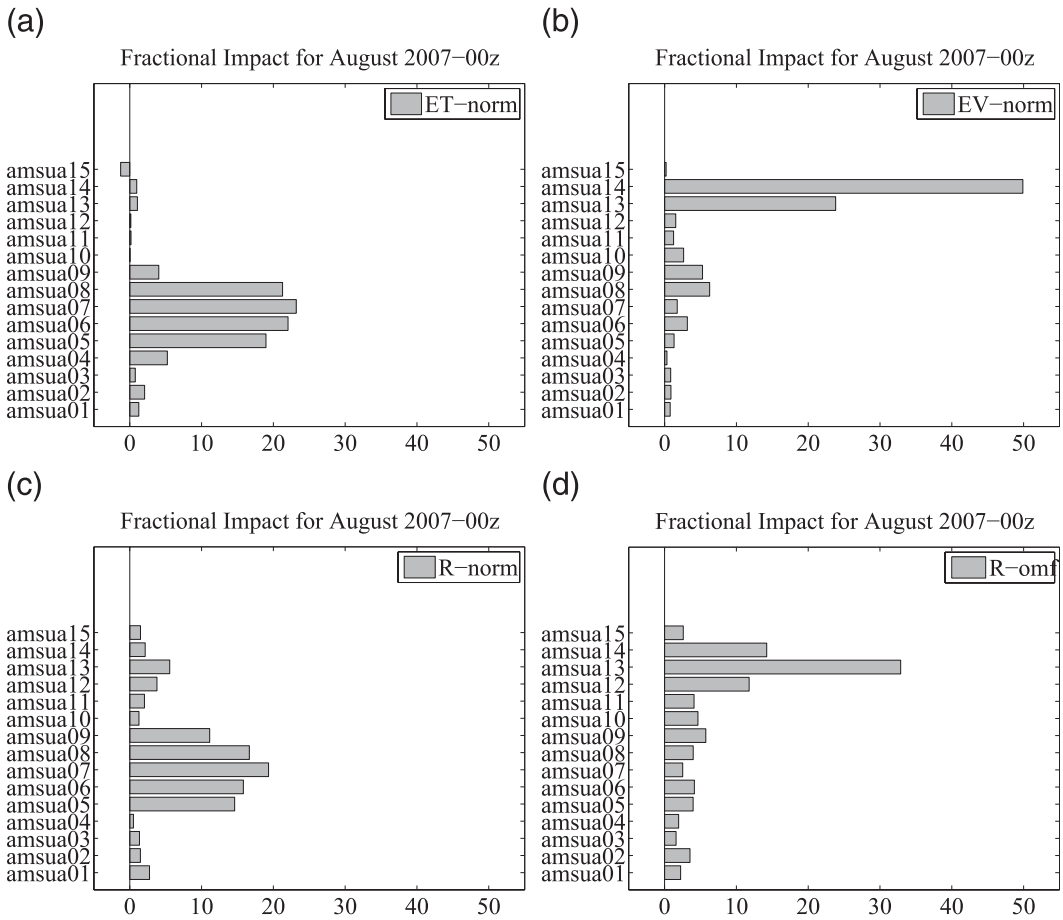

FIG. 4. As in Fig. 3, but for each of the 15 channels of AMSU-A on the *NOAA-15*, *-16*, *-18* and on the NASA *Aqua* satellites; fractions are now calculated with respect to total 31-day impact of all AMSU-A channels for each case.

impact and blue the regions where impact is negative. Figures 5a–c light up similar areas of the globe: the eastern Pacific Ocean is a particular area where all three measures agree about the significance of the observations. There are also areas of disagreement, such as the southwestern Pacific Ocean, which the EV-norm indicates the observations to not be doing so well whereas results seem quite neutral when evaluated with either the ET- or R-norm. Clearly, there is a level of arbitrariness in what we choose to highlight in the plots, which is another way of saying that when comparing different measures of observation impacts there is considerable uncertainty involved until statistical significance is assigned to the results (left for future work).

2) Observation-space approach

We now turn our attention to the observation-space approach and impact calculations based on (11) when using weighting factors defined by the inverse of the observation error covariance matrix, as in (34). Results of observation impacts on the 24-h forecast are seen in Figs. 3d, 4d, and 5d.

Looking again at the summary plots in Fig. 3 and focusing on how the ''traditional'' state-space ET-normalized observation impacts (Fig. 3a) compare with the observation-space impacts (Fig. 3d), the most noticeable difference is in the fractional impact of radiosondes. With the observation-space measure radiosondes seem much less significant, whereas aircraft, AMSU-A, and AIRS, seem to show roughly similar results in both approaches and norms, though their ranking changes. The reduced fractional impact of radiosondes measured with the observation-space approach is compensated for by the increased fractional impacts of AMSU-A and AIRS. A more direct comparison follows by contrasting Figs. 3c,d, since the norms in these two cases are closely related, and the comparison more faithfully addresses the difference in approaches rather than norms. However, given the similarity between results with the ET-norm and those with the R-norm the same conclusions

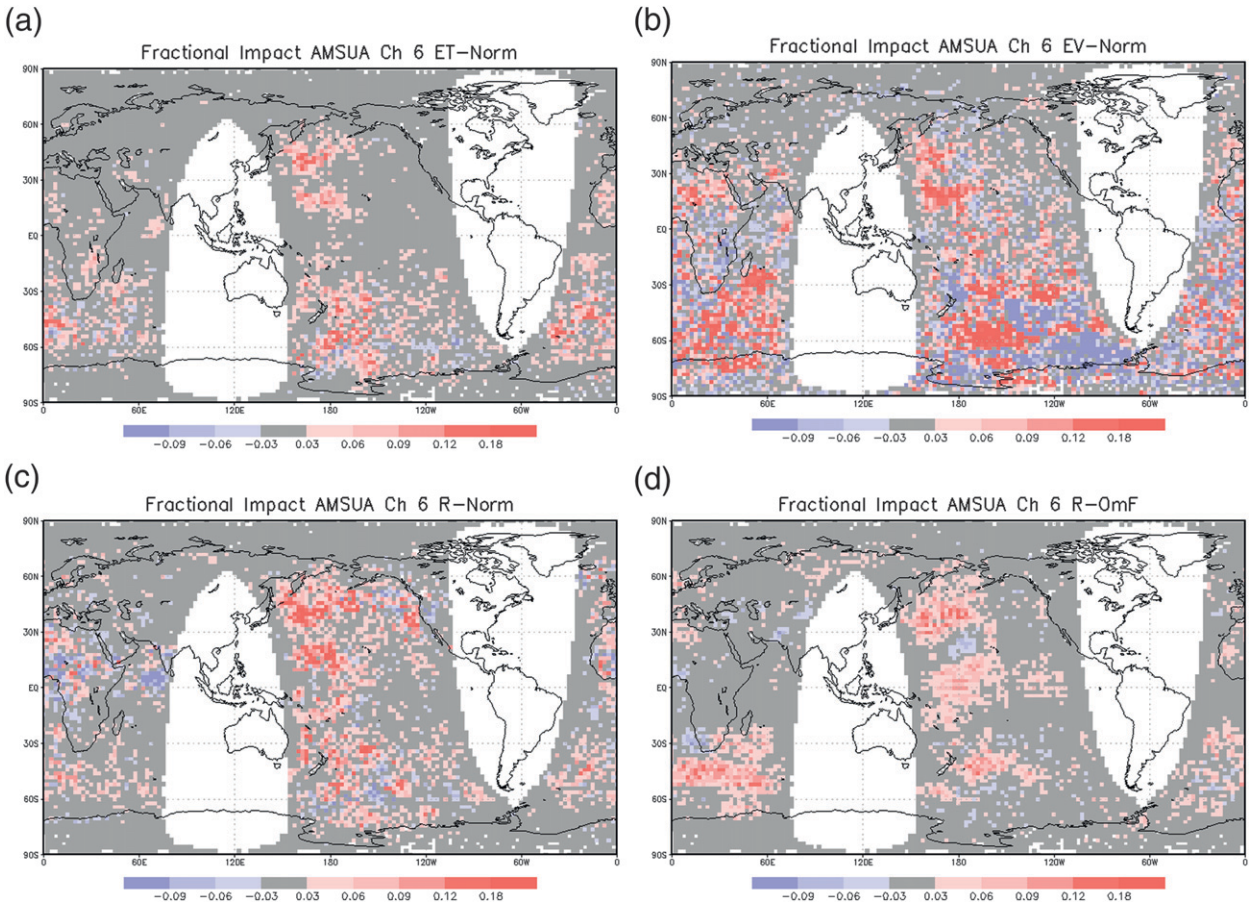

FIG. 5. Fractional impacts on the 24-h forecasts from channel 6 of AMSU on *NOAA-18* for the state-space approach using (a) the ET-norm, (b) the EV-norm, and (c) the R-norm, and (d) the observation-space approach using the R-norm.

follow: the dominant observing systems under the observation-space measure are aircraft, AIRS, and AMSU-A. The reduced role of the radiosonde observations can be understood by recalling that the observation-space-based measure compares how the 30- and 24-h forecasts fit the observations 30 and 24 h later, respectively. This measure does not explicitly incorporate the sensitivities of forecasts; it simply compares the quality of the forecasts directly. For the case of observation impacts on the 0000 UTC forecasts, the observation-space measure is saying that the forecasts from 1800 UTC are not as bad in predicting the radiosonde observations 6 h later as one might be inclined to think. Taking the difference in (9) as a proxy for the difference in forecast quality, one would be led to think the 30- and 24-h forecasts do not predict the radiosondes as well when evaluation is done with the state-space measures. Indeed, what this measure is saying is that there are larger sensitivities in the 30-h forecasts than in the 24-h forecasts when predicting the radiosonde observations (due to the considerable difference in the radiosonde network between 0000 and 1800 UTC). Unfortunately, this is not information the cycling analysis system can make use of.

Examining the results for the various channels of AMSU-A displayed in Fig. 4, we see the fractional impacts from the observation-space measure (Fig. 4d) resemble more closely those obtained with the state-space approach when it uses the EV-norm (Fig. 4b). The stratospheric channels, 12–14, have a large role. Here again, we emphasize that this should not lead us to think that the role of the troposphere-peaking channels is not significant; it is simply that 6 h is enough time for the model stratosphere to reset its own errors, stressing the relevance of radiance assimilation at these levels. This is not unreasonable since there are plenty of other data in the troposphere and only those few channels in the stratosphere. Still the relevance of assimilating radiances in the troposphere is noticeable in the observation-space measure. This can be seen in Fig. 5d, where similar regions of significant impact from channel-6 brightness temperatures are highlighted when compared with other measures.

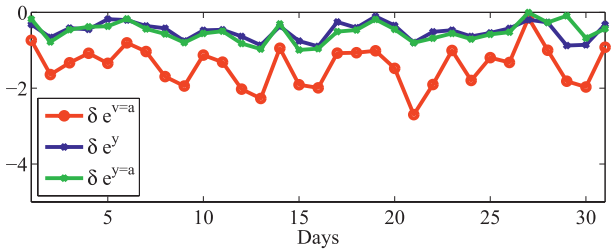

FIG. 6. Time series of total forecast error reduction calculated from the state-space approach in (6b), with the norm in (39) (red curve); the observation state approach using (11), with the weighting matrix in (34), and verified against the observations (blue curve) or against the analysis (green curve).

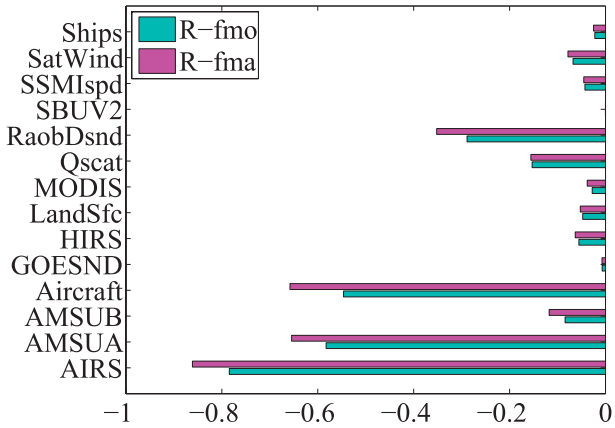

FIG. 7. Breakdown of observation impacts on the 24-h forecasts calculated with the observation-space approach. The blue bars show impacts when observations are used for verification; the magenta are for when the analyses are used for verification. A preliminary version of this result appeared in Todling (2009).

b. *Role of verification*

Let us now turn our attention to the role of the verification in light of the discussion in section 2d. As we saw, the observation-space approach provides a way to examine the matter by explicitly calculating, for example, the first term on the right-hand side of (31) and the left-hand side of the same equation. Using the GEOS-5 illustration considered here, Fig. 6 displays the time series of total observation impacts, for each 0000 UTC analysis of August 2007 calculated in observation space using the weights in (34), but when the verification is taken to be (i) the observations [first term on right-hand side of (31), δe^y; blue curve]; and (ii) the analyses converted to observation locations [left-hand side of (31), $\delta e^{y=a}$; green curve]. Note that the sum of all observation impacts used to calculate the fractional impacts Fig. 3d amounts to the sum of all impacts included in the time series displayed by the blue curve in the current figure. As a whole, it seems that the nearness of the blue and green curves is confirmation of the hope that it is reasonable to verify against the analysis, even at such short ranges as 1 day. Also, from the discussion in section 2d, the proximity of the curves is indication that the system is not so far from optimality—the last term in (31) does not appear to be very dominant—although there are exceptions, such as seen on day 29 of the time series in Fig. 6.[2]

These conclusions, however, cannot be taken too far. Another, closer, look at the differences between verifying against the observations versus verifying against the analysis is presented in Fig. 7, where the total observation impact for the 0000 UTC analyses of August 2007 is partitioned into the different components of the observing system. The blue bars are for when verifying against the observations and the magenta bars are for when the projected analyses are used for verification. Verification against the analysis provides an overestimate of the impacts for each observing system.

As pointed out in section 2d, similar calculations cannot easily be done in state space, but we can employ approximations such as those in (6). In practice, calculations from these approximations automatically involve verifying against the analysis. The red curve in Fig. 6 shows the time series of total impact per analysis time when the weights in (39) and the state-space-based second-order formula in (6b) are used. The state-space approach seems to overestimate the impact of the observations in the 1-day forecasts beyond what one gets with the observation-space approach. Even when the norms in the two approaches are made to resemble each other, it is still the case that their corresponding errors live in largely different spaces resulting in considerably different total impacts.

c. *OMB-based observation impacts*

Though there are some clear similarities among results obtained with the state- and observation-space approaches and various norms, the truth is that the closer one looks, the more differences one finds. These differences become rather hard to reconcile with the idea that results from observation impact on the forecasts are to aid our ability to make improvements in the underlying data assimilation system. As pointed out at the end of section 2, it is the opinion of this author that measures directly linked to the data assimilation cycle can naturally reduce some of the redundancy that

[2] It is important not to forget that here, as in most works, the expectation operator in expressions such as (31) is replaced with a simple summation (average) as only a single realization is available in practice.

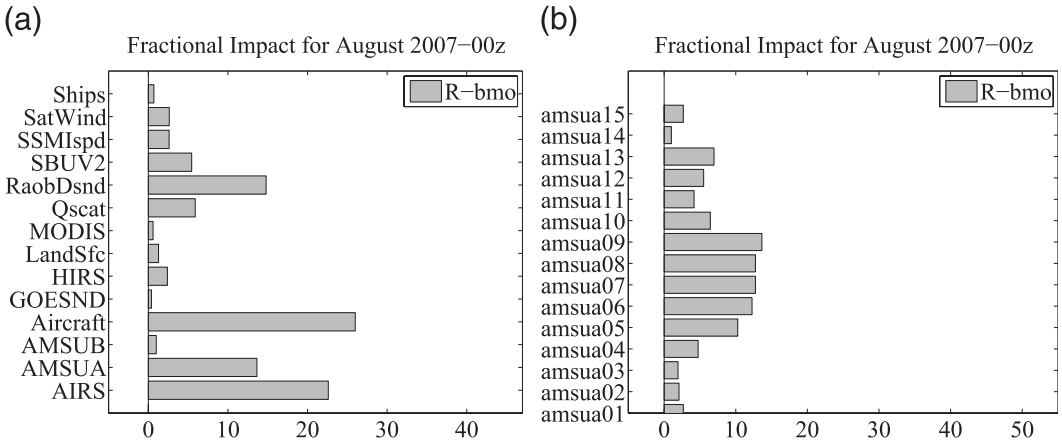

FIG. 8. Observations impacts on the 0-h forecast (analysis) calculated with the observation-space approach, when weighting is based on the inverse of the observation error covariances. (a) A summary similar to those appearing in Fig. 3; and (b) similar to those appearing in Fig. 4, and provides a breakdown of result for all platforms carrying AMSU-A instruments in August 2007.

norm-dependent approaches bring. Fortunately, the by-products of any data assimilation system already provide most of what is needed to diagnose its own reliability, including how effectively various components of the observing system are used. The statistics of the time series of observation-minus-background (OMB) and observation-minus-analysis (OMA) residuals provide all needed information to allow such investigation to take place. This is not to say the task is easy. There is plenty of redundant information in the observing system (and therefore in the OMBs and OMAs), there are also observation and model biases that interfere with the robustness of these statistics, and there are spatial and temporal gaps in the observing system that make the problems quite arduous. What we know from estimation theory is that most of the information is found in these residuals.

In this final section we then turn to the straightforward calculation indicated in section 2e, resulting from considering the observation-space approach with $m = 1$. We construct so-called observation impacts from simply the differences of normalized squared OMA and OMB, as in (33). These are essentially the impact of observations on the analyses (referred to below as "0 h" impact). The panels in Fig. 8 show results analogous to those in Figs. 3 and 4, but here the 0-h fractional impacts are evaluated in observation space. Concentrating first on the summary plot of Fig. 8a we see considerable resemblance to the observation-space-based impacts on the 24-h forecasts shown in Fig. 3d. A noticeable difference being that radiosondes and aircraft seem to play a relatively larger role affecting the analysis cycle than the 24-h forecasts; the roles of AMSU-A and AIRS seem to be more dominant in impacting the 24-h forecast as compared to radiosondes and aircraft. Turning to Fig. 8b, the breakdown of fractional impacts for AMSU-A shows a slightly dominant contribution coming from the largely tropospheric channels 5–9. When comparing with the 24-h observation-space fractional impacts in Fig. 3d, the stratospheric channels show less impact affecting the cycling data assimilation than affecting the 24-h forecasts. This is consistent with the view that when set free the model tends to go back to the stratosphere it likes, while during the assimilation cycle the model does not have the opportunity to restore its stratospheric behavior (e.g., Bloom et al. 1996).

To provide further corroboration for the results above, an experiment covering the more recent period of December 2011 is briefly examined. The experiment now takes the analysis to be at its present full $0.625° \times 0.5°$ horizontal resolution, and uses a more recent version of GEOS-5 capable of handling the latest components of the observing system, such as refractivity derived from radio occultation observations from various platforms of the Global Positioning System (GPSRO), the Microwave Humidity Sounder (MHS), the Infrared Atmospheric Sounding Interferometer (IASI), and the Advanced Scatterometer (ASCAT) on board the *European Meteorological Operational* (*MetOp-A*) satellite. As before, observation impacts are only shown for 0000 UTC. The comparison that follows is not meant to be exhaustive. It is simply aimed at adding to the main points of the present contribution.

Figure 9, is similar to Fig. 3, but now for December 2011, and it shows summaries of observation impacts on the 24-h forecasts derived with the state-space

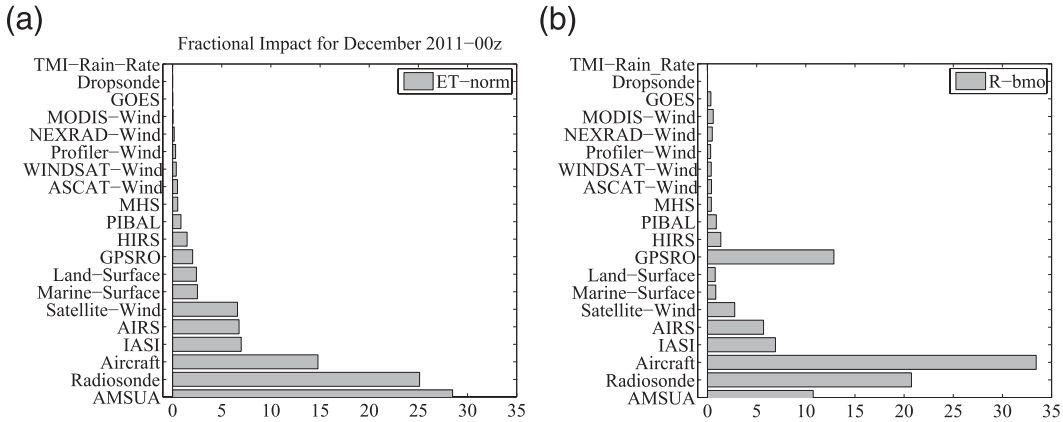

FIG. 9. (a) State-space, ET-norm, observation impact on the 24-h forecast and (b) observation-space, R-norm, observation impact on the 0-h forecast (analysis); as in (a) in Figs. 3 and 8, respectively, but for December 2011.

approach using the ET-norm (Fig. 9a), and observation impacts on the analyses derived with the observation-space approach combining (33) with the weight in (34). To avoid being repetitive, results from other norms and approaches are not considered here. Caution should be exercised when comparing results from Figs. 3a and 9a. The first figure covers a summer month while the second covers a winter month. And, more importantly, some major components of the observing system change quite dramatically. Beyond the new instrument types already mentioned above, we also point out that *NOAA-16* available in August 2007 is no longer available in December 2011, while *NOAA-19* is only available during this latter period; AMSU-B has basically been replaced with MHS; and a decision was made not to use window channels of AMSU-A in this more recent version of GEOS-5. This is not a comprehensive list of changes in the observing system, but cites the most important for the present work. Even with all these differences, there are clear similarities between the two time periods in terms of which observing systems dominate. AMSU-A, radiosondes, and aircrafts are still the ones impacting most the 24-h forecasts (see Fig. 9a). Next to them, AIRS is quite dominant, too, but now shares the role with IASI and the satellite winds; the lesser role played by the satellite winds in the August 2007 experiment might be associated with the considerably lower analysis resolution in that experiment.

The most relevant comparison for the present work is made by looking at the two panels in Fig. 9. When it comes to influencing the analysis cycle (Fig. 9b), the role of AMSU-A is not as dramatic as indicated by the 24-h impact results (this is similar to what is found for August 2007; see discussion of results for Fig. 8). However, in December 2011 the role of radiosonde is markedly larger than that of AMSU-A. This reduced impact of AMSU-A on the analysis is seen here as due to the addition of IASI and GPSRO to the present period. In a sense these new data types ''steal'' from the influence of AMSU-A; indeed, even AIRS, which in August 2007 was seen as a major contributor to the 0-h impact, has its role downplayed now, likely due to the same reason. Such compensatory interplay among the various components of the observing system has been noticed in the observing system experiments of Gelaro and Zhu (2009). Comparing the 24- and 0-h impacts, we also see some ranking differences among the various observing systems. One of the most noticeable changes is seen for GPSRO: 24-h impacts lead us to think GPSRO contributes little, but the role of GPSRO in aiding the analyses is undeniable. This significant contribution of GPSRO to the analysis is consistent with results from the work of Cucurull (2010).

Last, we look at the percentage of observations contributing positively to the observation impacts. These are calculated by simply counting the number of observations whose impacts are negative (beneficial) for each instrument of interest, then dividing each result by the corresponding total number of observations for that instrument type, and multiplying the result by 100. Figure 10 shows results for both the state-space- and observation-space-based impacts on the 24- and 0-h forecasts, respectively, as displayed in Fig. 9. The panel on the right is analogous to the result shown by Gelaro et al. (2010, their Fig. 5), though for a different time period and observing system selection here. Gelaro et al. (2010) remind us that scalar theoretical analysis in idealized settings (M. Fisher 2006, personal communication; Ehrendorfer 2007) indicate that roughly 60%–65% of the observations should contribute positively to the background when the accuracy of both backgrounds

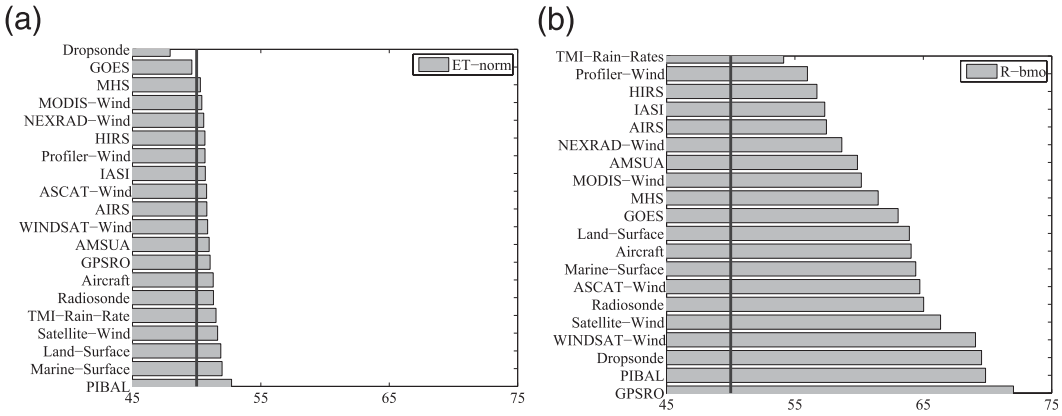

FIG. 10. As in Fig. 9, but showing percentage of observations contributing positively to the (a) 24- and (b) 0-h forecasts. The vertical line in both panels indicates the 50% beneficial mark for reference.

and observations are comparable. Results such as the one in Fig. 10a, suggesting that only roughly 50% of the observations lead to positive impact on the 24-h forecast, has led Gelaro et al. (2010) to the conclusion that operational systems cannot be expected to perform at theoretical levels. Results displayed in Fig. 10b state otherwise: operational systems do corroborate the scalar theoretical analysis. The bulk of the observing system contributes at exactly the expected theoretical range, that is, 60%–65% of the assimilated observations contribute positively. This is further corroboration of what was concluded while examining the result of Fig. 6, that is, global measures indicate that current operational data assimilation systems run near optimality. This is not to say further improvements are not needed or possible. As results from Fig. 7 have indicated, there is still plenty of room for improvements when we start looking more closely. Similarly, results from Fig. 10 suggest work needs to be done to bring performance of some observing systems to theoretical levels.

4. Conclusions

Studies of observation impact on the forecast have relied on the approach put forward by Langland and Baker (2004). A number of works have followed since. The present work is a contribution that provides insight on basic issues behind the technique. A few limitations and difficulties associated with the basic approach have been highlighted here, namely: (i) the need to rely on a norm not directly linked to the underlying data assimilation cycle; (ii) the need to rely on a verifying state; (iii) the need to rely on the model adjoint; and last, (iv) the added computational expense. Though (iii) has been tackled in the work of Liu and Kalnay (2008) by generalizing the approach of Langland and Baker to work within the context of ensemble data assimilation procedures, the other issues still remain. More complex is the idea of using observation impacts derived from these available techniques as an aid to improve on the use of observations in the corresponding cycling data assimilation system.

The present work identifies two approaches to observation impacts. The ''traditional'' method works in state space, while an alternative is to define measures of observation impact directly in observation space. Arguing that results from observation impact studies must be interpreted statistically, the present work recasts the problem in the language of estimation theory. This allows studying more closely the assumptions involved in the methodology. In particular, it becomes clear that a state-space approach is more encompassing than an observation-space approach, simply because observations span a smaller space than the full state space. But this advantage quickly disappears when realizing that the state-space approach requires a verification state normally not available in practice. Under certain conditions, the consequences of choosing the analysis for verification are investigated, showing explicitly how the corresponding observation impacts carry undesirable correlations with the verification. The observation-space approach, on the other hand, allows verification to be made against the observations therefore, in principle, avoiding such undesirable correlations. Furthermore, the observation-space approach permits evaluating what is obtained when the analyses, instead of the observations, are used for verification. It is shown that only under optimality can a system be indifferent to whether verification is done against the observations or the analyses. It is recognized in the present work that some of this advantage disappears in practice since observations are usually bias corrected, thus making the observations also

correlated with the analysis. The effect of observation bias correction on impacts derived from observation-space approaches remains to be studied in future work.

The fifth generation of the Goddard Earth Observing System (GEOS-5) is used to illustrate the main points of the theoretical part of the present work. First, the role of the norm used to defined the forecast aspect of interest is investigated. Results from the, usual, largely tropospheric measure are compared with those from when the norm weighs the vertical evenly. Observation-space considerations suggest investigating results from a third norm based on the inverse of the observation error covariances projected onto state space. This is later compared with a similar norm, but directly defined in observation space. Observation impacts on the 24-h forecasts obtained with these different norms show considerable similarities, but also show differences that might be difficult to sort out if results are to be used to try to revisit the observing system used in the data assimilation system to try to improve its cycled performance.

Second, GEOS-5 is also used to investigate the role of the verification, and in particular the effect of using the analysis for verification. Within the observation-space approach, using the analysis instead of the observations for verification seems to make only a small difference, at least when it comes to the impact of the entire observing system on the 24-h forecasts. From the theoretical discussion, this result is interpreted as indicative of the assimilation system operating near optimality. However, it is also shown that this interpretation cannot be pushed too far. A closer look at the observation impact for the individual components of the observing system indicates impacts tend to be overestimated when verification is done against the analysis. Moreover, when the state-space approach uses a similar norm to calculate observation impacts on the 24-h forecasts, the impacts are an even larger overestimate than when derived in observation space.

Ultimately, an argument is made that observation impacts might as well be defined on the basis of the time series of the readily available observation-minus-background and observation-minus-analysis residuals. The argument follows directly from theoretical considerations. The main advantage of using these residuals to define observation impacts is that whatever is learned from them is immediately related to the cycling data assimilation system. They reflect how observations affect the analyses. They relate directly to the traditional use of residuals going back to the works of Kailath (1968), and largely explored by many works in the meteorological and oceanic data assimilation literature. It is even shown that our present data assimilation systems corroborate back-of-the-envelope calculations for how much the assimilation of observations should in principle, positively, impact the backgrounds.

The observation-space approach definition of observation impact used in the present work, namely the difference between an observation-space error measure calculated for two consecutive analyses, was simply employed for convenience of the mathematical analysis and to facilitate comparison with the state-space approach. Evaluation of the impact of observations through methods based on information content and degrees of freedom for signal should be just as well suitable (e.g., Chapnik et al. 2006; Lupu et al. 2011). The main point being that observation-space approaches relying on observation-minus-background and observation-minus-analysis residuals are the only ones directly relating the effectiveness of the observations to the quality of the data assimilation cycle.

Acknowledgments. The author would like to thank Stephen E. Cohn for fruitful discussions at different stages of this work, and for having gone over an early version of this manuscript. The author also thanks Dacian N. Daescu for discussions at an early stage of this investigation. Thanks are also due to two anonymous reviewers for providing insightful comments and helping to improve the arguments presented in this study, and to Herschel Mitchell for helping improve readability. The computational work was carried out on the Linux Explore and Discover Systems through cooperation with the NASA Center for Computational Sciences at Goddard Space Flight Center. This research was partially supported by the Atmospheric Data Assimilation Development component of the NASA Modeling, Analysis and Prediction Program (MAP/04-0000-0080).

APPENDIX A

Proof of Equation (21)

In this appendix we derive expression (21), for linear dynamics, making use of the typical assumption of time-uncorrelated model errors. We start by making use of the following recursion:

$$\epsilon^f_{k|k-m} = \mathbf{M}_{k,k-m}\epsilon^a_{k-m|k-m} - \sum_{j=0}^{m-1} \mathbf{M}_{k,k-j}\epsilon^q_{k-j}, \tag{A1}$$

expressing the forecast error at time t_k initialized from an analysis obtained at time t_{k-m}, where ϵ^q_k represents the n-vector model error. Consequently, the expected value of the outer product of this error vector with itself defines the forecast error covariance:

$$\mathbf{P}^f_{k|k-m} = \left\langle \mathbf{M}_{k,k-m}\epsilon^a_{k-m|k-m}(\epsilon^a_{k-m|k-m})^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}_{k,k-m}\right\rangle + \sum_{i=0}^{m-1}\sum_{j=0}^{m-1}\left\langle \mathbf{M}_{k,k-i}\epsilon^q_{k-i}(\epsilon^q_{k-j})^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}_{k,k-j}\right\rangle$$
$$= \mathbf{M}_{k,k-m}\mathbf{P}^a_{k-m|k-m}\mathbf{M}^{\mathrm{T}}_{k,k-m} + \sum_{j=0}^{m-1}\mathbf{M}_{k,k-j}\mathbf{Q}_{k-j}\mathbf{M}^{\mathrm{T}}_{k,k-j}, \tag{A2}$$

where the first equality is obtained by noticing that the analysis error at time t_{k-m} is uncorrelated with model errors in its future, that is, for $t_k > t_{k-m}$, and the second equality makes use of the fact that model errors are uncorrelated in time and assumed to have covariance $\mathbf{Q}_k = \langle \epsilon^q_k(\epsilon^q_k)^{\mathrm{T}}\rangle$.

Therefore, the forecast error difference $\Delta\mathbf{P}_k$ becomes,

$$\begin{aligned}\Delta\mathbf{P}_k &\equiv \mathbf{P}^f_{k|k-m+1} - \mathbf{P}^f_{k|k-m}\\ &= \mathbf{M}_{k,k-m+1}(\mathbf{P}^a_{k-m+1|k-m+1} - \mathbf{Q}_{k-m+1})\mathbf{M}^{\mathrm{T}}_{k,k-m+1} - \mathbf{M}_{k,k-m}\mathbf{P}^a_{k-m|k-m}\mathbf{M}^{\mathrm{T}}_{k,k-m}\\ &= \mathbf{M}_{k,k-m+1}(\mathbf{P}^a_{k-m+1|k-m+1} - \mathbf{Q}_{k-m+1})\mathbf{M}^{\mathrm{T}}_{k,k-m+1} - \mathbf{M}_{k,k-m+1}\mathbf{M}_{k-m+1,k-m}\mathbf{P}^a_{k-m|k-m}\mathbf{M}^{\mathrm{T}}_{k-m+1,k-m}\mathbf{M}^{\mathrm{T}}_{k,k-m+1}\\ &= \mathbf{M}_{k,k-m+1}(\mathbf{P}^a_{k-m+1|k-m+1} - \mathbf{P}^f_{k-m+1|k-m})\mathbf{M}^{\mathrm{T}}_{k,k-m+1},\end{aligned} \tag{A3}$$

where the last equality is obtained after application of the usual one-time forecast error covariance propagation expression:

$$\mathbf{P}^f_{k-m+1|k-m} = \mathbf{M}_{k-m+1,k-m}\mathbf{P}^a_{k-m|k-m}\mathbf{M}^{\mathrm{T}}_{k-m+1,k-m} + \mathbf{Q}_{k-m+1}, \tag{A4}$$

at t_{k-m+1}. This completes proof of (21).

APPENDIX B

Proof of Equation (28)

To derive an expression relating the perceived and actual forecast error changes notice first that by adding and subtracting the true state at time t_k from (5), the expected perceived error can decomposed into

$$\langle e^v_{k|\ell}\rangle = \langle(\epsilon^f_{k|\ell} - \epsilon^v_k)^{\mathrm{T}}\mathbf{T}_k(\epsilon^f_{k|\ell} - \epsilon^v_k)\rangle = \mathrm{Tr}[\mathbf{T}_k(\mathbf{P}^f_{k|\ell} + \mathbf{P}^v_k)] - 2\langle(\epsilon^v_k)^{\mathrm{T}}\mathbf{T}_k\epsilon^f_{k|\ell}\rangle, \tag{B1}$$

where $\epsilon^v_k \equiv \mathbf{x}^v_k - \mathbf{x}^t_k$ is the error in the verification $\mathbf{x}^v_k$, and $\mathbf{P}^v_k \equiv \langle\epsilon^v_k(\epsilon^v_k)^{\mathrm{T}}\rangle$ is its corresponding error covariance. Inserting this in expression (9), the expected perceived error change becomes

$$\begin{aligned}\langle\delta e^v_k\rangle &= \mathrm{Tr}[\mathbf{T}_k(\mathbf{P}^f_{k|k-m+1} + \mathbf{P}^v_k)] - 2\langle(\epsilon^v_k)^{\mathrm{T}}\mathbf{T}_k\epsilon^f_{k|k-m+1}\rangle - \mathrm{Tr}[\mathbf{T}_k(\mathbf{P}^f_{k|k-m} + \mathbf{P}^v_k)] + 2\langle(\epsilon^v_k)^{\mathrm{T}}\mathbf{T}_k\epsilon^f_{k|k-m}\rangle\\ &= \mathrm{Tr}[\mathbf{T}_k\Delta\mathbf{P}^f_k] - 2\langle(\epsilon^v_k)^{\mathrm{T}}\mathbf{T}_k(\epsilon^f_{k|k-m+1} - \epsilon^f_{k|k-m})\rangle\\ &= \langle\delta e_k\rangle - 2\langle\epsilon^v_k\mathbf{T}_k\mathbf{M}_{k,k-m+1}\tilde{\mathbf{K}}_{k-m+1}\mathbf{d}^{\mathrm{T}}_{k-m+1|k-m}\rangle\\ &= \langle\delta e_k\rangle - 2\mathrm{Tr}[\tilde{\mathbf{K}}^{\mathrm{T}}_{k-m+1}\mathbf{M}^{\mathrm{T}}_{k,k-m+1}\mathbf{T}_k\langle\epsilon^v_k\mathbf{d}^{\mathrm{T}}_{k-m+1|k-m}\rangle],\end{aligned} \tag{B2}$$

where the equality before last is obtained after using (15) and (17), and we remind the reader again that though the notation $i \mid j$ resembles the estimation theory notation for conditional means, the use here is symbolic—this is the reason why the matrices in the expression above can be moved outside of the expectation operation $\langle\cdot\rangle$. This completes the proof of (28).

APPENDIX C

Proof of Equation (29)

When the verification is chosen to be the underlying analysis, that is, $\mathbf{x}^v_k = \mathbf{x}^a_{k|k}$, the relationship between the expected perceived and actual errors derived in appendix B can be explored farther. In this case, the error

cross-covariance matrix inside the trace term on the right-hand side of (28) involves the analysis error $\epsilon_k^v = \epsilon_{k|k}^a$. To evaluate this cross covariance we notice first that the residual vector $\mathbf{d}_{k|\ell}$ can be approximated to first order as

$$\mathbf{d}_{k|\ell} = \mathbf{y}_k^o - \mathbf{h}_k(\mathbf{x}_{k|\ell}^f) = \epsilon_k^o + \mathbf{h}_k(\mathbf{x}_k^t) - \mathbf{h}_k(\mathbf{x}_{k|\ell}^f) \approx \epsilon_k^o - \mathbf{H}_{k|\ell}\epsilon_{k|\ell}^f. \tag{C1}$$

Furthermore, assuming the forecast error propagates nearly linearly as in the propagation of analysis error from time t_{k-m+1} to time t_k follows the recursion:

$$\epsilon_{k|k}^a = \mathbf{M}_{k,k-m+1}\epsilon_{k-m+1|k-m+1}^a + \sum_{j=0}^{m-2}\mathbf{M}_{k,k-j}(\tilde{\mathbf{K}}_{k-j}\mathbf{d}_{k-j,k-j-1} - \epsilon_{k-j}^q), \tag{C2}$$

and relates the analysis error at time t_k with the analysis error from time t_{k-m+1} all the way to time t_k, plus a term that amounts to the propagation of analysis increments from all times between times t_{k-m+2} to t_k, and the propagated contribution from all model error components, ϵ_{k-j}^q for $j = 0, \ldots, m-2$. This recursion is applicable within the validity of piecewise linearizations of the various time intervals between the times t_{k-m+1} and t_k.

Substituting (C2) in the cross-covariance term in (28), and using that $\varepsilon_{k-m+1|k-m+1}^a = \varepsilon_{k-m+1|k-m}^f + \tilde{\mathbf{K}}_{k-m+1}\mathbf{d}_{k-m+1|k-m}$, it follows that

$$\begin{aligned}\left\langle \epsilon_k^{v=a}\mathbf{d}_{k-m+1|k-m}^{\mathrm{T}}\right\rangle &= \Big\langle [\mathbf{M}_{k,k-m+1}\epsilon_{k-m+1|k-m}^f + \mathbf{M}_{k,k-m+1}\tilde{\mathbf{K}}_{k-m+1}\mathbf{d}_{k-m+1,k-m} \\ &\quad + \sum_{j=0}^{m-2}\mathbf{M}_{k,k-j}(\tilde{\mathbf{K}}_{k-j}\mathbf{d}_{k-j,k-j-1} - \epsilon_{k-j}^q)]\mathbf{d}_{k-m+1|k-m}^{\mathrm{T}}\Big\rangle, \\ &= \mathbf{M}_{k,k-m+1}(-\mathbf{P}_{k-m+1|k-m}^f\mathbf{H}_{k-m+1|k-m}^{\mathrm{T}} + \tilde{\mathbf{K}}_{k-m+1}\mathbf{\Gamma}_{k-m+1}) \\ &\quad + \sum_{j=0}^{m-2}\mathbf{M}_{k,k-j}\tilde{\mathbf{K}}_{k-j}\left\langle \mathbf{d}_{k-j,k-j-1}\mathbf{d}_{k-m+1|k-m}^{\mathrm{T}}\right\rangle = \mathbf{M}_{k,k-m+1}\Delta\mathbf{K}_{k-j}\mathbf{\Gamma}_{k-m+1} \\ &\quad + \sum_{j=0}^{m-2}\mathbf{M}_{k,k-j}\tilde{\mathbf{K}}_{k-j}\left\langle \mathbf{d}_{k-j,k-j-1}\mathbf{d}_{k-m+1|k-m}^{\mathrm{T}}\right\rangle,\end{aligned} \tag{C3}$$

for observations $\mathbf{y}_k^o = \mathbf{h}_k(\mathbf{x}_k^t) + \epsilon_k^o$, and noticing that the observation error ϵ_k^o is uncorrelated with the model error ϵ_k^q, for all times t_k and t_ℓ, as well as it is uncorrelated with all forecast errors $\epsilon_{k|\ell}^f$ calculated for forecasts issue for $\ell < k$. This concludes the proof of (29).

REFERENCES

Anderson, B. D. O., and J. B. Moore, 1979: *Optimal Filtering.* Prentice-Hall, 357 pp.

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.,* **124,** 1256–1271.

Buizza, R., C. Cardinali, G. Kelly, and J.-N. Thépaut, 2007: The value of observations. II: The value of observations located in singular-vector-based target areas. *Quart. J. Roy. Meteor. Soc.,* **133,** 1817–1832.

Cardinali, C., R. Buizza, G. Kelly, M. Shapiro, and J.-N. Thépaut, 2007: The value of observations. III: Influence of weather regimes on targeting. *Quart. J. Roy. Meteor. Soc.,* **133,** 1833–1842.

Chapnik, B., G. Desroziers, F. Rabier, and O. Talagrand, 2006: Diagnosis and tuning of observational error in a quasi-operational data assimilation setting. *Quart. J. Roy. Meteor. Soc.,* **132,** 543–565.

Chou, M.-D., and M. J. Suarez, 1999: A shortwave radiation parameterization for atmospheric studies. Tech. Rep. NASA-TM-104606, Vol. 15, NASA, 40 pp.

Cohn, S. E., 1997: An introduction to estimation theory. *J. Meteor. Soc. Japan,* **75,** 257–288.

——, N. Sivakumaran, and R. Todling, 1994: A fixed-lag Kalman smoother for retrospective data assimilation. *Mon. Wea. Rev.,* **122,** 2838–2867.

——, A. da Silva, J. Guo, M. Sienkiewicz, and D. Lamich, 1998: Assessing the effects of data selection with the DAO physical-space statistical analysis system. *Mon. Wea. Rev.,* **126,** 2913–2926.

Collins, N., and Coauthors, 2005: Design and implementation of components in the Earth System Modeling Framework. *Int. J. High Perform. Comput. Appl.,* **19,** 341–350.

Courtier, P., and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.,* **113,** 1311–1328.

——, J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.,* **120,** 1367–1387.

Cucurull, L., 2010: Improvement in the use of an operational constellation of GPS radio occultation receivers in weather forecasting. *Wea. Forecasting,* **25,** 749–767.

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., 1992: The lagged-innovation covariance: A performance diagnostic for data assimilation. *Mon. Wea. Rev.,* **120,** 178–196.

Dee, D. P., and A. M. da Silva, 1999: Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. *Mon. Wea. Rev.,* **127,** 1822–1834.

Derber, J. C., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.,* **19,** 1333–1347.

——, and W.-S. Wu, 1998: The use of TOVS could-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.,* **126,** 2287–2299.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli, 2005a: Diagnosis of observation, background and analysis-error statistics in observation space. *Quart. J. Roy. Meteor. Soc.,* **131,** 3385–3396.

——, P. Brousseau, and B. Chapnik, 2005b: Use of randomization to diagnose the impact of observations on analyses and forecasts. *Quart. J. Roy. Meteor. Soc.,* **131,** 2821–2837.

Ehrendorfer, M., 2007: A review of issues in ensemble-based Kalman filtering. *Meteor. Z.,* **16,** 795–818.

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

——, R. Gelaro, E. Novakovskaia, and R. Todling, 2007: General characteristics of stratospheric singular vectors. *Meteor. Z.,* **16,** 621–634.

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.

——, R. H. Langland, S. Pellerin, and R. Todling, 2010: The THORPEX observation impact intercomparison experiment. *Mon. Wea. Rev.,* **138,** 4009–4025.

Giering, R., T. Kaminski, R. Todling, R. Errico, R. Gelaro, and N. Winslow, 2005: Generating tangent linear and adjoint versions of NASA/GMAO's FORTRAN-90 global weather forecast model. *Lecture Notes in Computational Science and Engineering,* H. M. Bücker et al., Eds., Vol. 50, Springer, 275–284.

Hollingsworth, A., and P. Lönnberg, 1989: The verification of objective analyses: Diagnostics of analysis system performance. *Meteor. Atmos. Phys.,* **40,** 3–27.

Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory.* Academic Press, 376 pp.

Kailath, T., 1968: An innovations control approach to least square estimation—Part I: Linear filtering in additive white noise. *IEEE Trans. Automat. Contrib.,* **AC-13,** 646–655.

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.

Kleespies, T. J., P. van Delst, L. M. McMillin, and J. Derber, 2004: Atmospheric transmittance of an absorbing gas. 6. OPTRAN status report and introduction to the NESDIS/NCEP Community Radiative Transfer Model. *Appl. Opt.,* **43,** 3103–3109.

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *J. Geophys. Res.,* **105** (D20), 24 809–24 822.

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.

Lewis, J. M., K. D. Raeder, and R. M. Errico, 2001: Vapor flux associated with return flow over the Gulf of Mexico: A sensitivity study using adjoint modeling. *Tellus,* **53A,** 74–93.

——, S. Lakshmivarahan, and S. K. Dhall, 2006: *Dynamic Data Assimilation: A Least Squares Approach.* Encyclopedia of Mathematics and Its Applications Series, Vol. 104, Cambridge University Press, 647 pp.

Lin, S.-J., 2004: A ''vertically Lagrangian'' finite-volume dynamical core for global models. *Mon. Wea. Rev.,* **132,** 2293–2307.

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.

——, ——, T. Miyoshi, and C. Cardinali, 2009: Analysis sensitivity calculation in an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc.,* **135,** 1842–1851.

Lupu, C., P. Gauthier, and S. Laroche, 2011: Evaluation of the impact of observations on analyses in 3D- and 4D-Var based on information content. *Mon. Wea. Rev.,* **139,** 726–737.

Ménard, R., S. E. Cohn, L.-P. Chang, and P. M. Lyster, 2000: Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part I: Formulation. *Mon. Wea. Rev.,* **128,** 2654–2671.

Moorthi, S., and M. Suarez, 1992: Relaxed Arakawa–Schubert: A parameterization of moist convection for general circulation models. *Mon. Wea. Rev.,* **120,** 978–1002.

Rabier, F., and Coauthors, 2008: An update on THORPEX-related research in data assimilation and observing strategies. *Nonlinear Processes Geophys.,* **15,** 81–94.

Rienecker, M. M., and Coauthors, 2008: The GEOS-5 Data Assimilation System—Documentation of versions 5.0.1, 5.1.0, and 5.2.0. Tech. Rep. NASA-TM-104606, Technical Report Series on Global Modeling and Data Assimilation, Vol. 27, NASA, 101 pp.

Stappers, R. J. J., and J. Barkmeijer, 2011: Optimal linearization trajectories. *Ninth Int. Workshop on Adjoint Model Applications in Dynamic Meteorology,* Cefalu, Italy, NASA, 31 pp. [Available online at http://gmao.gsfc.nasa.gov/events/adjoint_workshop-9/presentations/Barkmeijer.pdf.]

Stieglitz, M., A. Ducharne, R. Koster, and M. Suarez, 2001: The impact of detailed snow physics on the simulation of snow cover and subsurface thermodynamics at continental scales. *J. Hydrometeor.,* **2,** 228–242.

Todling, R., 2009: An approach to assess observation impact based on observation-minus-forecast residuals. *Proc. ECMWF Workshop on Diagnostics of Data Assimilation System Performance,* Reading, United Kingdom, ECMWF, 199–202. [Available online at http://www.ecmwf.int/publications/library/do/references/list/20091231.]

——, and Y. Trémolet, 2008: The GMAO 4DVAR and its adjoint tools. *16th Int. TOVS Study Conf. (ITSC-16),* Angra dos Reis, Brazil, International TOVS Working Group. [Available online at http://cimss.ssec.wisc.edu/itwg/itsc/itsc16/proceedings/A30_Todling.pdf.]

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.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spacially inhomogeneous covariances. *Mon. Wea. Rev.,* **130,** 2905–2916.

Zhu, Y., and R. Gelaro, 2008: Observation sensitivity calculations using the adjoint of the Gridpoint Statistical Interpolation (GSI) analysis system. *Mon. Wea. Rev.,* **136,** 335–351.