Tellus
SERIES A
DYNAMIC METEOROLOGY AND OCEANOGRAPHY
PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

Tellus (2009), 61A, 179–193

TELLUS

Examination of observation impacts derived from observing system experiments (OSEs) and adjoint models

By RONALD GELARO[1]* and YANQIU ZHU[2], [1]*Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA;* [2]*Science Applications International Corporation, Beltsville, MD 20705, USA*

(Manuscript received 24 July 2008; in final form 2 December 2008)

ABSTRACT

With the adjoint of a data assimilation system, the impact of any or all assimilated observations on measures of forecast skill can be estimated accurately and efficiently. The approach allows aggregation of results in terms of individual data types, channels or locations, all computed simultaneously. In this study, adjoint-based estimates of observation impact are compared with results from standard observing system experiments (OSEs) using forward and adjoint versions of the NASA GEOS-5 atmospheric data assimilation system. Despite important underlying differences in the way observation impacts are measured in the two approaches, the results show that they provide consistent estimates of the overall impact of most of the major observing systems in reducing a dry total-energy metric of 24-h forecast error over the globe and extratropics and, to a lesser extent, over the tropics. Just as importantly, however, it is argued that the two approaches provide unique, but complementary, information about the impact of observations on numerical weather forecasts. Moreover, when used together, they reveal both redundancies and dependencies between observing system impacts as observations are added or removed from the data assimilation system. Understanding these dependencies appears to pose an important challenge in making optimal use of the global observing system for numerical weather prediction.

1. Introduction

Until recently, the impacts of observations on numerical weather forecasts have been measured mainly through observing system experiments (OSEs), in which subsets of observations are removed from a data assimilation system and the resulting forecasts compared against a control set that includes all observations. OSEs are performed intermittently at most operational centres (e.g. English et al., 2004; Kelly et al., 2004; Lord et al., 2004) but, because of their expense (multiple executions of the data assimilation system are required), usually involve a relatively small number of independent experiments, each considering relatively large subsets of observations. Nevertheless, OSEs have been extremely useful for quantifying the relative ‘importance’ of the various components of the observing system and have made clear, for example, the increasing benefit from satellite observations on the quality of weather forecasts in the Southern Hemisphere.

*Corresponding author.
e-mail: ron.gelaro@nasa.gov
DOI: 10.1111/j.1600-0870.2008.00388.x

During the last few years, the adjoint of a data assimilation system has also emerged as an accurate and efficient tool for estimating observation impacts on short-range weather forecasts (Baker and Daley, 2000; Langland and Baker, 2004). The impacts of any or all observations can be computed simultaneously based on a single execution of the adjoint system. In addition, the results can be easily aggregated by data type, location, channel, etc., making this technique especially attractive for examining the impacts of new hyper-spectral satellite instruments such as the Atmospheric Infrared Sounder (AIRS) and Infrared Atmospheric Sounding Interferometer (IASI), and for conducting regular, even near-real time, monitoring of the entire observing system. Other potential applications of the adjoint method include helping design intelligent strategies for data selection and thinning in data assimilation, improving the effectiveness of adaptively deployed (targeted) observations, and identifying future observing system needs.

While various aspects of the adjoint-based results have been studied in detail recently (Errico, 2007; Gelaro et al., 2007), comparisons between observation impacts derived from OSEs and the adjoint method have been mostly anecdotal up to now. This is due in part to the much more limited use of the latter at present, but also because of fundamental differences in the way observation impact is measured in the two approaches. These differences relate to, among other things, whether observations are included versus removed as a means of measuring their impact, and the treatment of information in the background state.

In this study, we compare observation impacts derived from OSEs and the adjoint of the GEOS-5 atmospheric data assimilation system developed at the NASA Global Modeling and Assimilation Office (GMAO). In contrast with most other observation impact studies, our experiments are designed with the primary intent of facilitating comparisons between the two methodologies rather than attempting to make definitive statements about which observing systems provide the most or least benefit for numerical weather forecasts.

Section 2 briefly describes the methodologies for computing observation impacts in the adjoint and OSE contexts, highlighting underlying differences relevant to the interpretation of results in later sections. Section 3 describes the design of the OSE and adjoint experiments conducted for this study. In Section 4, we present an overview of the results of these experiments in terms of metrics native to each approach and make general, mostly qualitative, comparisons between them. In Section 5, we make direct quantitative comparisons between the OSEs and adjoint results, based primarily on the fractional contributions from various observing systems to the total reduction of forecast error as measured by each method. In Section 6, we show how the combined application of OSEs and the adjoint technique can be used to examine complex dependencies between observing system impacts as observations are added or removed from the data assimilation system. Conclusions are presented in Section 7.

2. Estimation of observation impact

We consider the impact of observations on a measure of forecast error defined as

$$e = (\mathbf{x}^{\mathrm{f}} - \mathbf{x}^{\mathrm{t}})^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{C}\,\mathbf{P}\,(\mathbf{x}^{\mathrm{f}} - \mathbf{x}^{\mathrm{t}})\,, \tag{1}$$

where $\mathbf{x}^{\mathrm{f}}$ is a forecast state, $\mathbf{x}^{\mathrm{t}}$ is a verification state (considered 'truth'), $\mathbf{C}$ is a diagonal matrix of weights that defines a norm, $\mathbf{P}$ is a spatial projection operator that measures e only within a specified region of interest and the superscript T denotes the transpose operation. The forecast is produced by a non-linear model integration

$$\mathbf{x}^{\mathrm{f}} = \mathbf{m}(\mathbf{x}^0)\,, \tag{2}$$

where $\mathbf{m}$ is the model and $\mathbf{x}^0$ is the initial state. Both $\mathbf{x}^0$ and $\mathbf{x}^{\mathrm{t}}$ are based on atmospheric analyses produced at the initial and verification times, respectively, by a cycling data assimilation system. The assimilation system optimally combines observations $\mathbf{y}^{\mathrm{o}}$ with a prior model forecast (or background state) $\mathbf{x}_{\mathrm{b}}$ in the form

$$\mathbf{x}_{\mathrm{a}} = \mathbf{x}_{\mathrm{b}} + \mathbf{K}\delta\mathbf{y}\,, \tag{3}$$

where $\mathbf{K}$ is a matrix that determines the weight given to each observation and $\delta\mathbf{y}$ is the innovation vector

$$\delta\mathbf{y} = \mathbf{y}^{\mathrm{o}} - \mathbf{h}(\mathbf{x}_{\mathrm{b}}), \tag{4}$$

where $\mathbf{h}$ is an observation operator that relates the model state to the observations.

Following Errico (2007), we express changes in e due to changes in $\mathbf{x}^0$ in terms of a Taylor series approximation

$$\delta e = \sum_i \delta x_i^0 \left(\frac{\partial e}{\partial x_i^0} + \frac{1}{2}\sum_j \frac{\partial^2 e}{\partial x_i^0 \partial x_j^0}\,\delta x_j^0 + \frac{1}{6}\sum_{j,k} \frac{\partial^3 e}{\partial x_i^0 \partial x_j^0 \partial x_k^0}\,\delta x_k^0\,\delta x_j^0 + \cdots \right), \tag{5}$$

where δx_i^0 is a sufficiently small change to x_i^0 so that a truncated approximation is a good one. To examine the impact of observations on changes in e, an appropriate choice for δx_i^0 is the analysis increment

$$\delta\mathbf{x}^0 = \mathbf{x}_{\mathrm{a}} - \mathbf{x}_{\mathrm{b}} = \mathbf{K}\delta\mathbf{y}. \tag{6}$$

Written this way, the analysis equation may be easily interpreted as a transformation between a perturbation $\delta\mathbf{y}$ in observation space and a perturbation $\delta\mathbf{x}^0$ in physical space. Note that when $\delta\mathbf{y} = \mathbf{0}$, $\delta e = 0$, and there is no impact. With this choice, various-order approximations of δe can be expressed in terms of the innovations $\delta\mathbf{y}$ (Errico, 2007). Owing to the quadratic nature of (1), an approximation beyond first order is required to obtain a relatively accurate estimate of δe (Gelaro et al., 2007). In this study, we use the third-order approximation

$$\delta e \approx (\delta\mathbf{y})^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\left[\mathbf{M}_{\mathrm{b}}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{C}\,\mathbf{P}\,\left(\mathbf{x}_{\mathrm{b}}^{\mathrm{f}} - \mathbf{x}^{\mathrm{t}}\right) + \mathbf{M}_{\mathrm{a}}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{C}\,\mathbf{P}\,\left(\mathbf{x}_{\mathrm{a}}^{\mathrm{f}} - \mathbf{x}^{\mathrm{t}}\right)\right]\,, \tag{7}$$

where $\mathbf{x}_{\mathrm{b}}^{\mathrm{f}}$ and $\mathbf{x}_{\mathrm{a}}^{\mathrm{f}}$ are forecasts initialized from $\mathbf{x}_{\mathrm{b}}$ and $\mathbf{x}_{\mathrm{a}}$, $\mathbf{M}_{\mathrm{b}}^{\mathrm{T}}$ and $\mathbf{M}_{\mathrm{a}}^{\mathrm{T}}$ are the transposed resolvent matrices of the tangent linear version of the forecast model (usually referred to as the adjoint of the forecast model) evaluated along those trajectories and $\mathbf{K}^{\mathrm{T}}$ is the adjoint of the data assimilation system. This expression is equivalent to the impact measure derived originally by Langland and Baker (2004) using an alternative approach.

Equation (7) provides an estimate of the change, typically a reduction, in the error of $\mathbf{x}_{\mathrm{a}}^{\mathrm{f}}$ compared with that of $\mathbf{x}_{\mathrm{b}}^{\mathrm{f}}$ resulting from assimilation of the complete set of observations. It has the form of a weighted sum of the innovations, $\delta e \approx (\delta\mathbf{y})^{\mathrm{T}}\tilde{\mathbf{g}}$, where $\tilde{\mathbf{g}} = \mathbf{K}^{\mathrm{T}}[\mathbf{M}_{\mathrm{b}}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{C}\,\mathbf{P}\,(\mathbf{x}_{\mathrm{b}}^{\mathrm{f}} - \mathbf{x}^{\mathrm{t}}) + \mathbf{M}_{\mathrm{a}}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{C}\,\mathbf{P}\,(\mathbf{x}_{\mathrm{a}}^{\mathrm{f}} - \mathbf{x}^{\mathrm{t}})]$ is a vector in observation space. The impact of a particular subset of observations may be estimated by summing only those terms involving the corresponding elements of $\delta\mathbf{y}$. Although $\tilde{\mathbf{g}}$ is implicitly a function of $\delta\mathbf{y}$ (through its dependence on $\mathbf{x}_{\mathrm{a}}$), and therefore partial sums generally depend on observations outside the subset in question, we can ignore this effect for problems of interest in this study (Gelaro et al., 2007). The computation of $\tilde{\mathbf{g}}$ is done only once and includes the complete set of observations. Therefore, the impact of any subset of observations is determined with respect to *all other observations assimilated simultaneously*, so that the results accurately reflect the weights, $\mathbf{K}$, used to produce the analysis. The impact of any or all observations may be determined, therefore, from a single data assimilation experiment that includes all observations assimilated routinely (hereafter referred to as the control experiment). Application of (7) is subject to the same limitations as the adjoint model itself, including the validity of the tangent linear assumption and possible simplifications or deficiencies in the adjoint model compared to the non-linear model used to produce the forecast.

The impact of observations on (1) can also be assessed in an OSE context by simply computing differences in e between the control forecast and forecasts in which selected observations have been removed from the data assimilation system. In this case, (1) is evaluated in physical space from the non-linear forecasts and verifying analysis fields directly for each experiment. Results obtained with this approach contrast with those obtained with the adjoint approach in a number of important ways. They provide a measure of the change, typically an *increase*, in forecast error resulting from the *removal* of observations from the system. Moreover, the removal of observations generally increases the weights of the remaining observations which, assuming the system is optimally tuned for the complete set of observations, may further degrade the analysis and subsequent background state. As a result, the assimilation systems for the perturbed members of the OSE differ from the control as well as from one another. These effects may be significant since OSEs typically involve the removal of large subsets of observations, such as an entire satellite observing system, or observations that receive large weight in the analysis. Another important difference is that OSEs reflect the removal of observational information from both the background and analysis in a cumulative manner, whereas (7) measures the impact of observations in each analysis cycle separately and the background contains all previous information from the complete observing system.

3. Experimental design

The impacts of observations on 24-h forecasts from 00 UTC are examined for the months of July 2005 and January 2006 using forward and adjoint versions of the GEOS-5 atmospheric data assimilation system (Rienecker et al., 2007). The GEOS-5 system is comprised of the GMAO atmospheric model and the Gridpoint Statistical Interpolation (GSI) analysis scheme (Wu et al., 2002) developed at the National Centers for Environmental Prediction (NCEP). The model includes a finite volume dynamical core (Lin, 2004) and an advanced physics package. The GSI is a three-dimensional variational (3DVAR) analysis scheme which uses an incremental formulation involving successive minimizations, or outer loops, to solve a non-linear analysis problem. Analyses are produced at 0.5° horizontal resolution with 72 vertical levels using a 6-h assimilation cycle that includes all conventional observations and satellite radiances assimilated operationally during the study periods.

To evaluate (7), 24-h forecast trajectories are produced at 1.0° horizontal resolution with 72 vertical levels and full physics from the 00 UTC analysis ($\mathbf{x}_a$) and corresponding background state ($\mathbf{x}_b$) for each day of the study period. Adjoint forecasts along these trajectories are produced at this same horizontal and vertical resolution but with simplified dry physics. The starting conditions for the adjoint forecasts are $\partial e/\partial \mathbf{x}^f$, computed as indicated on the right-hand side of (7). The matrix $\mathbf{C}$ is defined such that (1) measures the forecast error in terms of the dry energy norm (Talagrand, 1981), including horizontal wind, temperature and surface pressure. This norm provides a simple way of combining mass-weighted, squared errors of disparate dynamical fields (Errico, 2000). Using the spatial projection operator $\mathbf{P}$ in (1), we define separate response functions and conduct separate adjoint experiments for e measured over the entire globe, the Northern Hemisphere (20°N–80°N), the Southern Hemisphere (20°S–80°S) and the tropics (20°N–20°S). All response functions extend vertically from the surface to approximately 130 hPa. The output fields from the model adjoint forecasts are then used as the starting conditions for the GSI adjoint. The GSI adjoint is an exact line-by-line adjoint (Zhu and Gelaro, 2008) run at the same resolution and with the same observation set used to produce $\mathbf{x}_a$. A postprocessing procedure is then applied to compute the inner product between the output of the GSI adjoint and the innovations $\delta\mathbf{y}$ to obtain the observation impact.

The outer loops of the GSI affect the impact of observations, particularly those whose operators are non-linear (e.g. radiances). As shown by Trémolet (2008), to account fully for this effect requires the second-order adjoint of the data assimilation system, which is not practical for any modern system. We use two outer loops of the GSI to produce the analyses in the present study, but account only partially for this in the adjoint calculation. As shown by Gelaro et al. (2007), the sum of all impacts estimated by (7) yields approximately 85% of the total determined by the difference $e(\mathbf{x}_b^f) - e(\mathbf{x}_a^f)$. This is comparable to the accuracy reported at other centres (e.g. Langland and Baker, 2004).

For comparison with the adjoint results, OSEs are conducted for July 2005 and January 2006 in which the following observation sets are individually excluded from the data assimilation system:

(1) all AMSU-A radiances from a single satellite (NOAA-16): ‘no amsua1’;

(2) all AMSU-A radiances from two satellites (NOAA-15, NOAA-16): ‘no amsua2’;

(3) all AMSU-A radiances from three satellites (NOAA-15, NOAA-16, Aqua): ‘no amsua3’;

(4) all AIRS radiances: 'no airs';
(5) all rawinsonde observations: 'no raob';
(6) all satellite winds (atmospheric motion vectors): 'no satwind';
(7) all aircraft observations: 'no aircraft';
(8) all scatterometer winds from QuikSCAT: 'no qkscat'.

Forecasts are produced from the 00 UTC analysis for each day and values of e are computed over the same global and regional verification domains used to define the adjoint response functions. In all cases, the verification state is the control analysis containing the complete set of observations. To allow some adjustment of the system to the initial removal of a given set of observations, the assimilation cycles for the OSEs are begun several days before the collection of forecast statistics begins on the first of each month.

4. Observation impact results

4.1. Adjoint results

In this section, we examine the impacts of various observing systems on 24-h forecasts from the 00 UTC analyses during January 2006 and July 2005 in terms of accumulated monthly values of δe. This requires, for each day of the study period, evaluation of the two model gradient terms on the right-hand side of (7), followed by application of $\mathbf{K}^{\mathrm{T}}$ and $\delta\mathbf{y}^{\mathrm{T}}$. The model gradient terms are evaluated along the forecast trajectories initialized from the 18 UTC background state and 00 UTC analysis state, respectively, and can thus be interpreted as the sensitivity of the energy-weighted 24-h forecast errors with respect to each of these initial states (e.g. Rabier et al., 1996). Before examining the accumulated values of δe, it is instructive to examine the monthly averaged values of the forecast error sensitivity patterns themselves. We emphasize that these monthly averaged sensitivities are not used in calculating the accumulated values of δe, which are based instead on the values for individual days.

As examples, Figs. 1 and 2 show the average 24-h global forecast error sensitivity patterns for January 2006 and July 2005, respectively, corresponding to the forecasts from the background (top panels) and analysed (middle panels) states, as well as the differences between these fields (bottom panels). The quantity shown is the time-averaged, vertically integrated energy of the sensitivity field at each horizontal gridpoint (e.g. Gelaro et al., 2002), which combines in a physically meaningful way the sensitivities with respect to wind, temperature and surface pressure. The areas of greatest sensitivity in the background states are concentrated near the major baroclinic zones of the winter hemisphere. During January, the maximum sensitivity is located over the central North Pacific, with secondary maxima over east Asia and the eastern North Atlantic. However, there is also significant sensitivity over the southern oceans, especially south of Africa and Australia. These results are consistent with those of previous

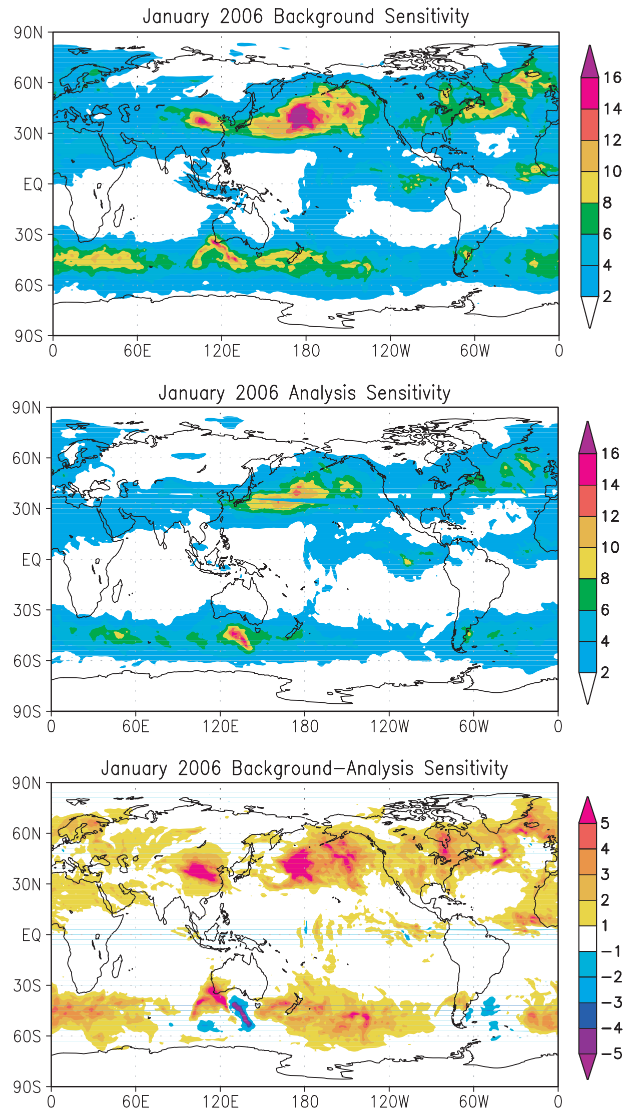

Fig. 1. Monthly averaged 24-h global forecast error sensitivity to initial conditions for January 2006 00 UTC in terms of vertically integrated energy for the background (upper), analysis (middle) and background minus analysis (lower). The units are 10^{-3} J kg^{-1}.

studies (e.g. Reynolds and Gelaro, 2001). During July, there is strong sensitivity extending across much of the mid-latitudes in the Southern Hemisphere and weak sensitivity over the entire Northern Hemisphere, especially compared with the results for January.

The effect of observations on reducing the forecast error sensitivity is clearly seen by comparing the upper and middle or lower panels of Figs. 1 and 2. On average, the observation impact is overwhelmingly beneficial in reducing the sensitivity in all but a few locations during both months. During January, there are substantial reductions in sensitivity over both land and sea, suggestive of the combined importance of both conventional land-based and oceanic or space-based observing systems

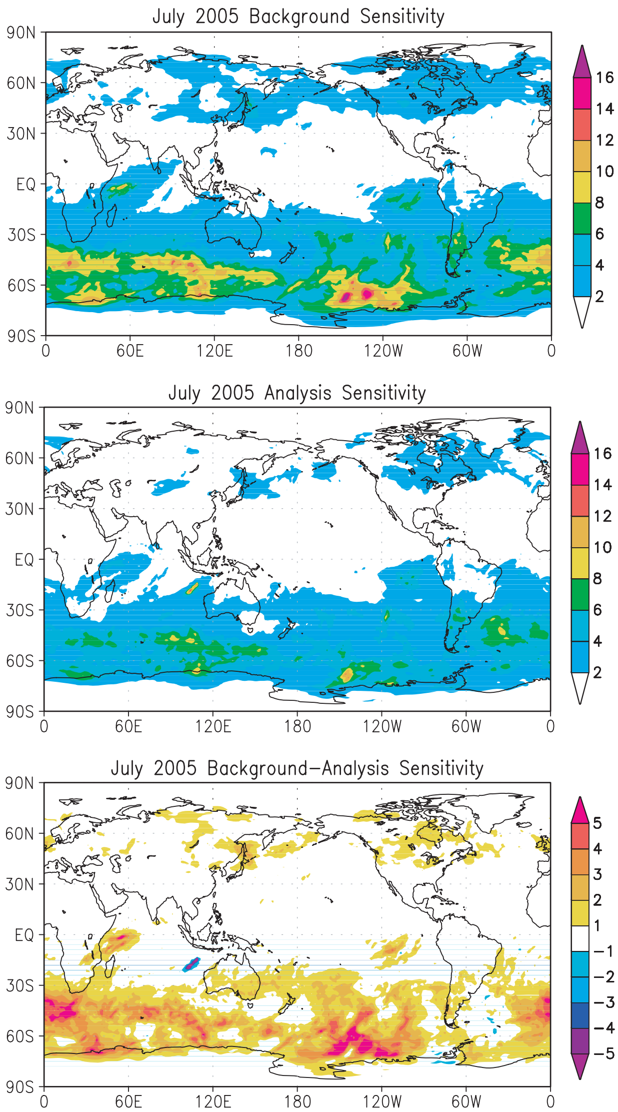

Fig. 2. As in Fig. 1, except for July 2005.

during northern winter. During July, the results suggest that oceanic or space-based observing systems play a crucial role in reducing forecast errors. Further investigation is required to explain the increased sensitivity in the analysed fields south of Australia during January.

The observation impacts implied by the (physical space) results in Figs. 1 and 2 are made explicit by the (observation space) results in Figs. 3 and 4, which show accumulated values of δe for the major observing systems assimilated in GEOS-5 at 00 UTC during the months of January 2006 and July 2005. The observing systems shown include those identified in Section 3 for the OSEs (except 'amsua1' and 'amsua2', which are discussed later) plus all radiances from AMSU-B ('amsub'), the High Resolution Infrared Radiation Sounder ('hirs'), geostationary sounders ('goes') and the Microwave Sounding Unit ('msu'), surface observations from ships, buoys and land ('surface') and wind speeds from the Special Sensor Microwave/Imager ('ssmisp').

The four panels in each figure show results for the globe, Northern Hemisphere, Southern Hemisphere and tropics, based on the separate forecast error response functions for these regions defined in Section 3. In all cases, the entire global set of observations is included in the calculations so that, for example, observations in the tropics can contribute to the forecast error reduction in the Northern or Southern Hemisphere. This is analogous to the conduct of most OSEs, including those in the present study, in which regional impacts on forecast skill are measured in response to the removal of entire observing systems from the data assimilation system.

Figures 3 and 4 show that, with only a few exceptions, the observing systems examined have an accumulated beneficial impact ($\delta e < 0$) on the 24-h forecast. The contributions from the various observing systems vary significantly depending on the season and verification region. Globally, both rawinsondes and AMSU-A radiances have the largest beneficial impact during both seasons, with smaller, but still significant, contributions from AIRS radiances, aircraft observations and satellite winds. The rawinsondes dominate the impact in the Northern Hemisphere while AMSU-A is dominant in the Southern Hemisphere, with both observing systems having the largest impact during the winter season in each case.

In the tropics, the relative impacts of the various observing systems are similar to the impacts globally, although the magnitude of the error reduction is smaller. It is interesting to note that, unlike in the Northern and Southern Hemispheres, the impact of AMSU-A in the tropics is somewhat larger in January than July while the impact of rawinsondes is larger in July than January. It should be kept in mind that, owing to the definition of $\mathbf{C}$ used in the present study, (1) does not measure errors in moisture explicitly, although it is affected by moisture indirectly. As a result, the impact of observations in the tropics, where moist processes play a significant role, may be underrepresented in these results. Although the impacts of all observing systems may be affected to some degree, those systems that observe moisture explicitly (e.g. AMSU-B) may be more significantly affected.

More generally, the fact that some observing systems have only a small positive, or even slightly negative ($\delta e > 0$), impact in Figs. 3 and 4 does not necessarily imply that they are 'bad' observations and should be excluded from the assimilation system. A different response function might yield different results. While in some cases it may be true that the assimilation system is not extracting a significant amount of information from a particular observation type, it is also possible that the background forecast is already quite accurate with respect to these observations and that little correction is required. In the same way, these observations may themselves significantly improve the quality of the background over many assimilation cycles but do not have a large impact on each analysis cycle separately. In that case, their benefit to the assimilation system might appear larger in

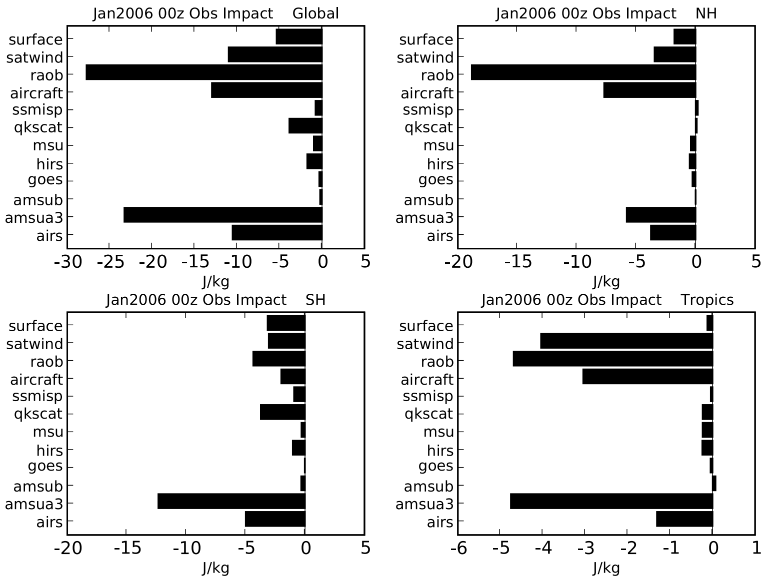

Fig. 3. Accumulated values of δe for various observing systems assimilated at 00 UTC during January 2006 using response functions for the globe (upper left-hand side), Northern Hemisphere (upper right-hand side), Southern Hemisphere (lower left-hand side) and tropics (lower right-hand side).

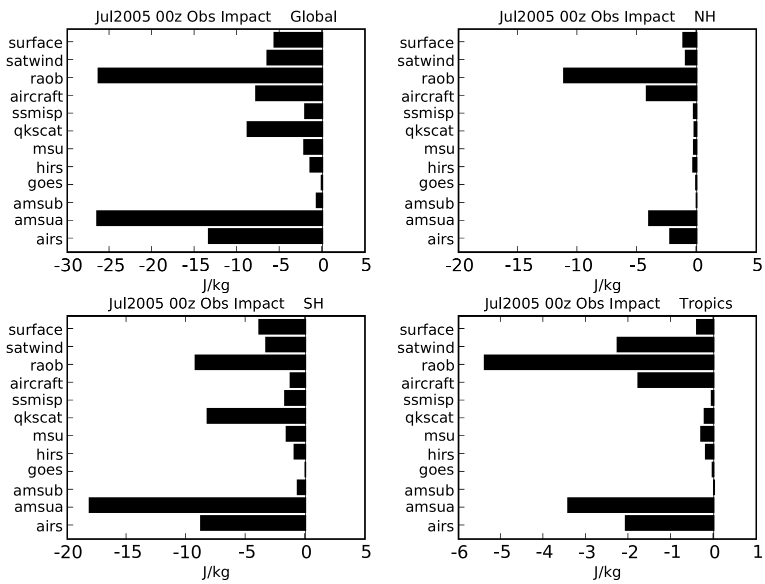

Fig. 4. As in Fig. 3, except for July 2005.

the context of an OSE. It is also possible that the background or observation errors are improperly specified, or that the methodology used to assimilate certain observation types needs to be improved. In the case of the slightly negative impact of SSM/I winds speeds over the Northern Hemisphere during January, for example, ancillary evidence shows that the procedure currently used to estimate the near-surface wind speed from the model background state introduces a significant low bias, especially in regions where the boundary layer is highly unstable. Map views of the impact of these observations (not shown) confirm that significant degradations occur in the vicinity of the Gulf Stream and Kuroshio currents, where cold continental air flows over much warmer ocean water.

4.2. *OSE results*

Figures 5 and 6 show results for a subset of the OSEs described in Section 3. For clarity, results are shown here only for the 'no raob', 'no amsua3' and 'no satwind' experiments, in addition to the control, for both January 2006 and July 2005. Each panel shows the growth of e for these experiments as a function of time for forecast days 1–5 for each verification region. While some qualitative comparisons are made with the adjoint results in Figs. 3 and 4, the results presented here are primarily intended to show that the OSEs in the present study yield reasonable results that are representative of the behaviours observed in similar experiments conducted at other centres. More detailed

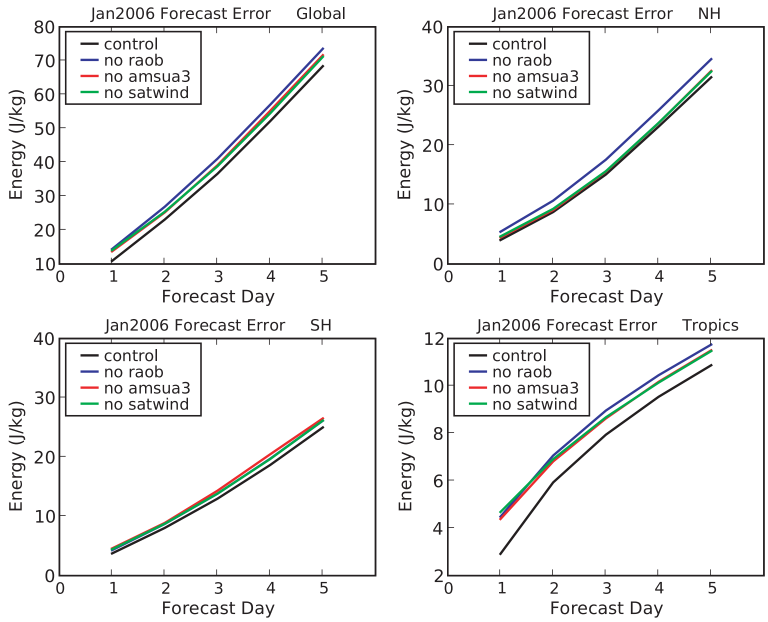

Fig. 5. Average values of *e* over the globe (upper left-hand side), Northern Hemisphere (upper right-hand side), Southern Hemisphere (lower left-hand side) and tropics (lower right-hand side) as a function of forecast length for various OSEs during January 2006.

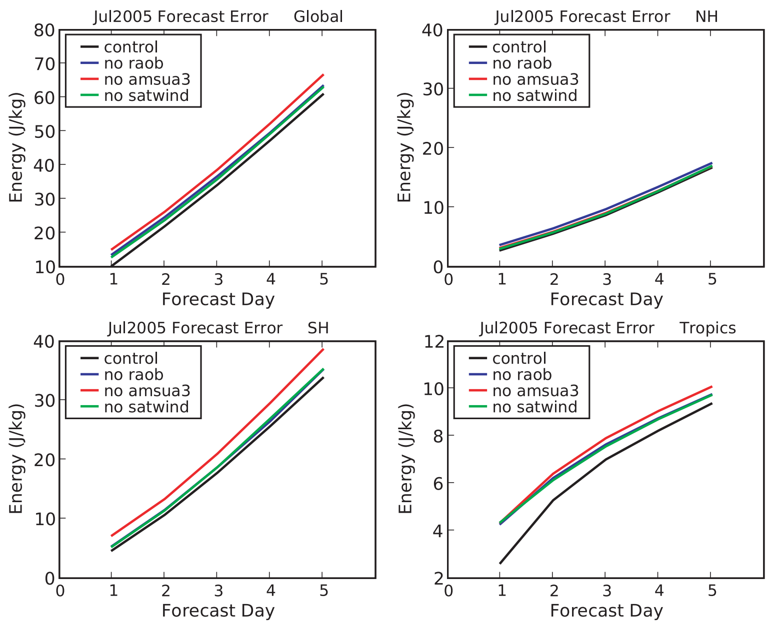

Fig. 6. As in Fig. 5, except for July 2005.

comparisons between the OSE and adjoint results are made in Sections 5 and 6.

As expected, the magnitudes and growth rates of the errors in Figs. 5 and 6 vary both seasonally and regionally. The largest seasonal differences occur in the Northern and Southern Hemispheres, with the largest magnitudes and growth rates occurring in the corresponding winter season. During January, the removal of rawinsonde observations has the largest impact globally, especially beyond day 1 of the forecast, and in the Northern Hemisphere at all forecast times. During July, the removal of all AMSU-A radiances has the dominant impact at all forecast times both globally and in the Southern Hemisphere especially. These results are consistent with the adjoint-based observation impacts in Figs. 3 and 4 which show the largest error reductions associated with these same observing systems for the corresponding months and verification regions. Comparisons between the adjoint and OSE results for other observing systems are more difficult based on inspection of these figures alone and are deferred to Sections 5 and 6.

The picture is more complex in the tropics. The forecast errors are not only smaller in magnitude but grow in a manner that more closely resembles a square root, as opposed to exponential, function of time.[1] At the same time, the relative impacts of the various observing systems change markedly as a function of time. During January, the removal of any of these observing systems results in a significant and near equal degradation of the forecast at day 1, with the removal of rawinsondes having the largest impact at later times. The results at day 1 appear more or less consistent with the adjoint impacts in Fig. 3 showing the large error reductions associated with these same observing systems. During July, there appear to be larger qualitative differences between the OSEs and adjoint results in Fig. 4, even at day 1. In the OSEs, the removal of any of these observing systems again degrades the forecast significantly at day 1. Later in the forecast, however, the removal of AMSU-A radiances has the largest impact. The adjoint results, on the other hand, indicate that the rawinsonde observations play the single largest role in reducing the 24-h forecast error in the tropics during July. Together, these results highlight the difficulty of attempting to draw general conclusions about the 'importance' of a given observing system based on a specific metric or forecast interval.

5. Direct comparison of OSEs and adjoint results

In this study, quantitative comparisons between the OSEs and adjoint results are restricted to the 24-h forecast since this is the only time at which the latter are formally valid. Even so, comparisons between the OSE and adjoint results are complicated by the fact that changes in e based on evaluation of (1) in the OSE context are not directly comparable to values of δe based on (7) in the adjoint context. We can, however, define for each approach a measure of the fractional impact of an observing system, j, on the total change in e as measured by that approach. Any such measures will of course reflect the same assumptions and underlying differences in interpretation as the native measures themselves, but will nonetheless provide a quantitative basis for comparison of results.

In the adjoint context, an obvious choice for measuring the fractional impact of a given observing system is

$$F_j(\mathrm{ADJ}) = \delta e_j / \delta e\,, \tag{8}$$

where δe_j represents the partial sum over only the elements in (7) involving observing system j. $F_j(\mathrm{ADJ})$ thus measures the fractional change in e with respect to the 24-h forecast from the background state resulting from the assimilation of observing system j. Note that for any partitioning of the complete observing system into j components, we have $\sum_j F_j(\mathrm{ADJ}) = 1$, since the impact of any or all observations in the adjoint context is assessed within a single experiment.

The choice of measure is less obvious for OSEs due, at least in part, to the fact that separate experiments employing different degradations of the assimilation system are required to assess each component of the observing system. At the same time, changes in e resulting from the removal of observations in these experiments can be measured with respect to forecasts from either the background or analysed states. Since OSEs typically involve comparisons with respect to the latter (as in Figs. 5 and 6), we focus first on impacts measured in terms of the differences between forecasts from the control and perturbed analyses. We define the fractional impact of a given observing system in the OSE context as

$$F_j(\mathrm{OSE}) = \left(e^{\mathrm{a}}_{j*} - e^{\mathrm{a}}_{\mathrm{ctl}}\right) / e^{\mathrm{a}}_{\mathrm{ctl}}\,, \tag{9}$$

where e^{a}_{j*} is the error measure of the 24-h forecast from the analysed state without observing system j and $e^{\mathrm{a}}_{\mathrm{ctl}}$ is the error measure of the 24-h forecast from the analysed control state. $F_j(\mathrm{OSE})$ measures the fractional change in e with respect to the 24-h forecast from the analysed control state resulting from the removal of observing system j from the assimilation system. Unlike $F_j(\mathrm{ADJ})$, there is no expectation that for a given partitioning of the observing system the sum of $F_j(\mathrm{OSE})$ will equal one since the contribution of each observing system is determined from a separate experiment and different permutation of the data assimilation system. If, for example, the forecast is highly sensitive to the removal of one or more types of observations, then the sum of the contributions may be much larger than one. In that case, as shown below, comparisons between $F_j(\mathrm{ADJ})$ and $F_j(\mathrm{OSE})$ are more restricted.

Figs. 7 and 8 show average values of $F_j(\mathrm{ADJ})$ and $F_j(\mathrm{OSE})$ during January 2006 and July 2005 for the eight observation sets removed in the OSEs. The four panels in each figure show results for the same verification regions as in previous figures. Over the global domain and extratropics, there is considerable quantitative agreement between $F_j(\mathrm{ADJ})$ and $F_j(\mathrm{OSE})$ for most observing systems during both months. The global results for January show good agreement for all observing systems except the satellite winds and, to a lesser extent, aircraft observations. In the Northern Hemisphere, there is good agreement for all observing systems. In the Southern Hemisphere, the impacts for AMSU-A are somewhat larger in the adjoint results, especially for a single satellite. In addition, the impact of satellite winds is larger in the OSE results, as in the global results. The results for July over the globe and extratropics are comparable with those for January in terms of the level of agreement between the adjoint and OSE results, with the notable exception that the removal of three AMSU-A instruments has a significantly larger impact in the OSE results both globally and over the Southern Hemisphere than does the combined impact of the three AMSU-A instruments in the adjoint results. There is, in contrast, good

[1] Similar behaviour has been reported by Kelly et al. (2004) in terms of the growth of root mean squared forecast errors in upper tropospheric tropical winds. Some studies suggest that this behaviour may indicate the dominance of model error over initial condition uncertainty in the growth of forecast errors (Orrell et al., 2001).

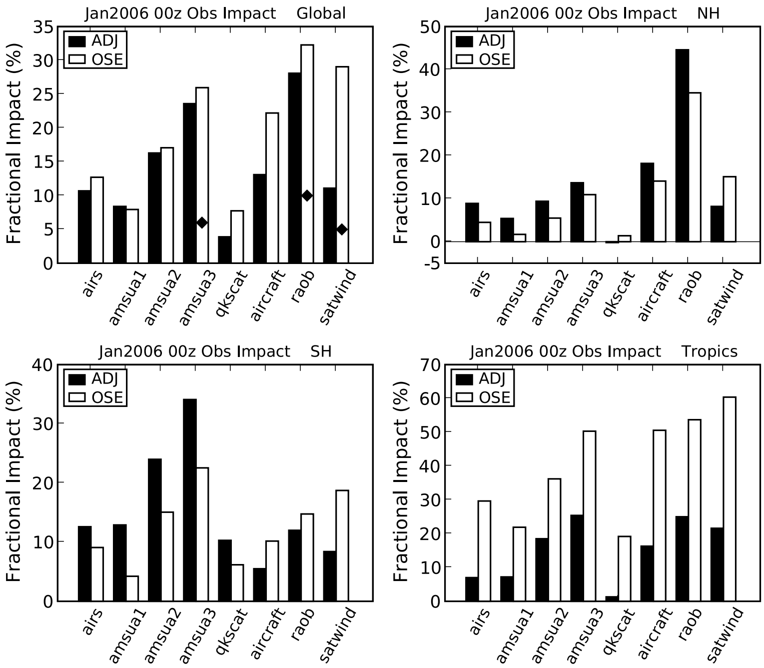

Fig. 7. Adjoint (black) and OSE (white) based fractional impacts of various observing systems on the change in 24-h forecast error over the globe (upper left-hand side), Northern Hemisphere (upper right-hand side), Southern Hemisphere (lower left-hand side) and tropics (lower right-hand side) during January 2006. The black diamonds in the upper left-hand panel denote results for supplemental experiments in which observations are removed but the control background state is used to compute the analysis. See text for details.

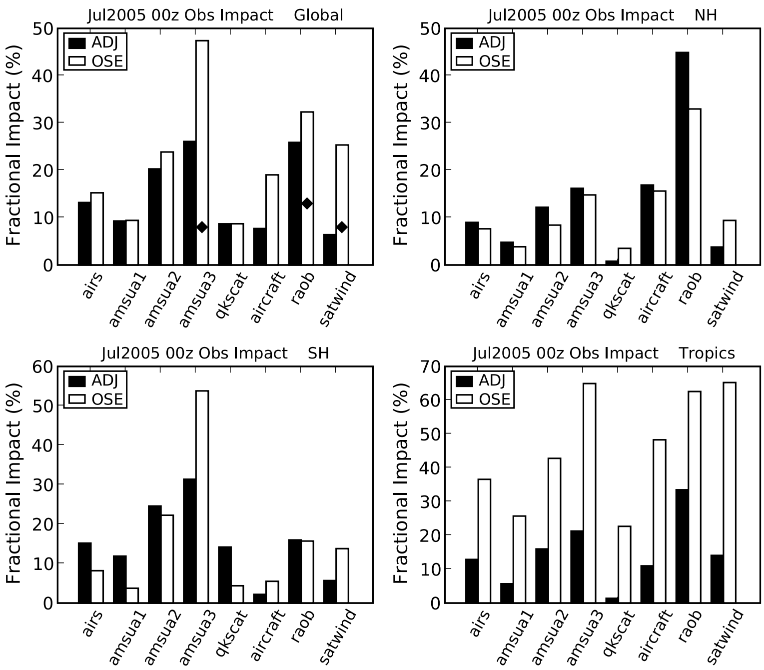

Fig. 8. As in Fig. 7, except for July 2005.

agreement between the adjoint and OSE results in the Northern Hemisphere during July for all AMSU-A permutations tested.

Assimilation of AMSU-A radiances plays a critical role in constraining the large-scale flow in the Southern Hemisphere (e.g. Kelly et al., 2004). The present results indicate that there is indeed an important difference between the inclusion versus complete removal of these observations in terms of quantifying their impact on the assimilation system. The removal of all AMSU-A radiances can reduce the quality of the assimilation system in the Southern Hemisphere, especially during winter, to the point where the results no longer represent a perturbation of the control state but rather a markedly different state. Figure 9 shows time-series of the 24-h forecast error norm over the Southern Hemisphere during January and July for the control experiment and the 'no amsua1', 'no amsua2' and 'no amsua3' experiments. Not only are the relative impacts of removing

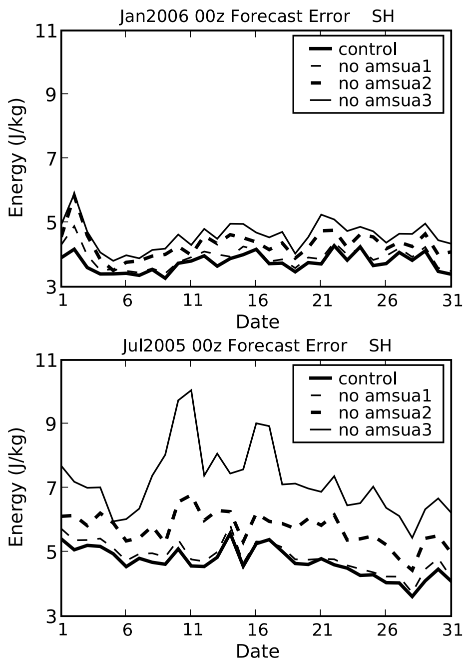

Fig. 9. Daily time-series of the 24-h forecast error norm over the Southern Hemisphere during January 2006 (upper) and July 2005 (lower) for selected OSEs: control (thick solid), 'no amsua1' (thin dash), 'no amsua2' (thick dash) and 'no amsua3' (thin solid) experiments.

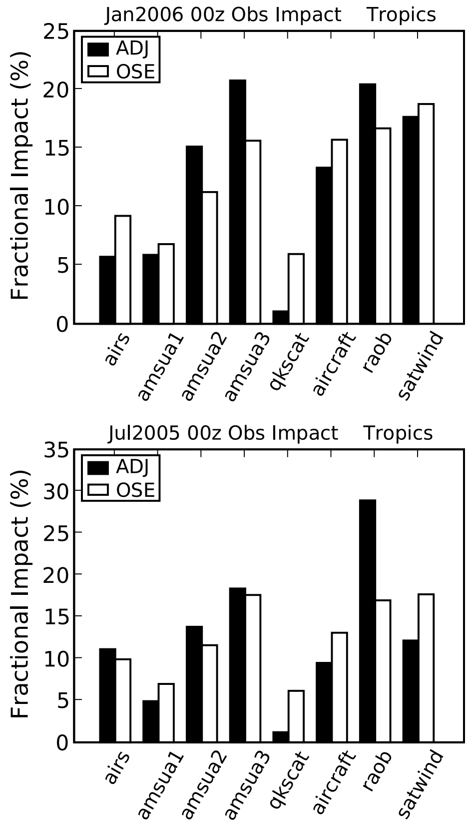

Fig. 10. Normalized adjoint (black) and OSE (white) based fractional impacts of various observing systems on the change in 24-h forecast error over the tropics during January 2006 (upper) and July 2005 (lower).

additional AMSU-A observations increasingly larger overall during July, but the removal of all AMSU-A observations produces large excursions in the time-series in which the values of e grow to as large as twice those of the control. These are by far the largest degradations in skill observed in any of the OSEs performed in this study.

Results for the tropics in Figs. 7 and 8 show greater overall disagreement between the magnitudes of F_j(ADJ) and F_j(OSE) during both January and July. Values of F_j(OSE) are much larger than those of F_j(ADJ) for all observing systems, with the former exceeding 50% for several observing systems during both months. The results are consistent with the large degradations with respect to the control experiment at day 1 in the OSEs noted earlier in Figs. 5 and 6, and reflect the generally high degree of sensitivity of the tropical forecasts, at least initially, to the removal of observations. In the adjoint results, it is of course impossible to have such large fractional contributions from several observing systems simultaneously. Nonetheless, the relative amplitudes of the various observing system contributions are consistent in the two sets of results. This can be seen more clearly in Fig. 10, in which the values of F_j(ADJ) and F_j(OSE) in the tropics for both January and July have been normalized such that they sum to one for the eight experiments shown. The normalized results show good overall agreement during both months. Only the contribution from QuikSCAT remains disproportionately larger in the OSEs after normalization, although it is among the smallest impacts overall. The contribution from satellite winds, on the other hand, is not disproportionately larger in the OSE results in the tropics, although it does remain larger in the global results even after normalization (not shown). The normalized adjoint contributions differ only slightly from their non-normalized counterparts since the latter are already constrained to be one for the complete observing system. In the same way, the normalization produces only minor changes in the OSE results over the globe and extratropics (not shown) where the non-normalized values are already in good overall agreement with the adjoint results.

The smaller impact of satellite winds in the adjoint results for the extratropics during both January and July cannot be fully explained based on the current results alone. The majority of these observations are in the tropics, which raises the possibility that their impact may be underestimated due to the absence of moist physical processes in the current adjoint model. However, this explanation is tempered by the results in Fig. 10 which show that in the tropical region itself the impact of the satellite winds is not disproportionately underrepresented in the adjoint results relative to other observing systems.

The different treatment of the background information in the OSEs and adjoint technique also may be an important factor. To explore this further, modified versions of the 'no satwind', 'no amsua3' and 'no raob' experiments were conducted in which the 00 UTC analysis for each day was computed using the control background state while each of these observation sets was individually excluded from the data assimilation system. The use of the control background state was intended to provide greater consistency with the adjoint-based calculations. Forecast statistics for these experiments were computed for the global verification domain. The results are shown in the upper left-hand panels of Figs. 7 and 8 by the black diamonds superimposed on the original OSE results for the same observing systems. It can be seen that the impacts in the modified OSEs are significantly smaller than those in the original OSEs in all cases. The results indicate that between two thirds and four fifths of the impact of removing these observing systems in the original OSEs comes from information accumulated in the background. The background contributions are proportionately largest for the satellite winds during January and for AMSU-A radiances during July, and smallest for the rawinsondes during both months but especially July. Compared with the original OSEs, the modified OSEs indeed show better agreement with the adjoint results in the case of the satellite winds, but much worse agreement in the case of the rawinsondes and AMSU-A radiances. Thus, while the impact of information in the background is clearly large in all cases, it does not appear to be of uniform consequence in explaining differences between the OSEs and adjoint results.

Finally, changes in e with respect to the analysis forecast as measured by F_j(OSE) can be related to changes in e with respect to the background forecast via an analogue of F_j(ADJ) in the OSE context. To develop this analogue, it is instructive first to approximate (8) using differences in the form

$$F_j(\mathrm{ADJ}) \approx \frac{e_j^{\mathrm{a}} - e_{\mathrm{ctl}}^{\mathrm{b}}}{e_{\mathrm{ctl}}^{\mathrm{a}} - e_{\mathrm{ctl}}^{\mathrm{b}}}, \tag{10}$$

where e_j^{a} is the error measure of the 24-h forecast from the analysed state in which only observing system j has been assimilated[2] and $e_{\mathrm{ctl}}^{\mathrm{b}}$ is the error measure of the 24-h forecast from the control background state. The quantities on the right-hand side of (10) are shown schematically in Fig. 11 for a case in which the assimilation of observing system j accounts for 20% of the error reduction obtained from assimilating the entire set of observations. Note that the numerator of (10) is negative, as is the denominator but with larger magnitude, so that F_j(ADJ) is a positive number equal to 0.2.

In the OSE context, we have not e_j^{a}, but e_{j*}^{a}, so that an appropriate analogue of (10) is

$$F_j'(\mathrm{OSE}) = 1 - \frac{e_{j*}^{\mathrm{a}} - e_{\mathrm{ctl}}^{\mathrm{b}}}{e_{\mathrm{ctl}}^{\mathrm{a}} - e_{\mathrm{ctl}}^{\mathrm{b}}} = \frac{e_{\mathrm{ctl}}^{\mathrm{a}} - e_{j*}^{\mathrm{a}}}{e_{\mathrm{ctl}}^{\mathrm{a}} - e_{\mathrm{ctl}}^{\mathrm{b}}}. \tag{11}$$

[2] This state typically would not exist in practice.

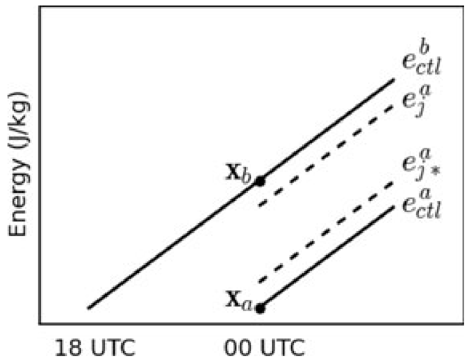

Fig. 11. Schematic representation of the error quantities used in the calculations of F_j(OSE) and F_j'(OSE) for a case in which assimilation (removal) of observing system j reduces (increases) e by 20%. See text for details.

The subtraction from one in (11) provides consistency with F_j(ADJ) in the sense that, for example, F_j'(OSE) also equals 0.2 if the removal of observing system j results in a 20% increase in forecast error with respect to the control. The measure e_{j*}^{a} for such a case is also shown in Fig. 11. As in (10), the numerator of (11) is a small negative number so that F_j'(OSE), like F_j(ADJ), is a small positive number less than one.

The relationship between F_j(OSE) and F_j'(OSE) can be expressed in terms of their ratio

$$\frac{F_j(\mathrm{OSE})}{F_j'(\mathrm{OSE})} = \frac{e_{\mathrm{ctl}}^{\mathrm{b}} - e_{\mathrm{ctl}}^{\mathrm{a}}}{e_{\mathrm{ctl}}^{\mathrm{a}}}. \tag{12}$$

Because this ratio depends only on measures of the control forecast, F_j(OSE) and F_j'(OSE) are proportional to each other by the same constant value for all j. In the present study, this value is close to 0.35 on average for both January and July. Thus, while F_j'(OSE) would appear nearly three times larger than F_j(OSE) if plotted in Figs. 7 and 8, normalized values of F_j'(OSE) would be identical to normalized values of F_j(OSE) such as those shown in Fig. 10. It is not surprising that the relative magnitudes of the various observing system impacts are the same whether measured in terms of changes in e with respect to the background or analysed control state. The fact that (12) is close to 1/3 may be understood by noting that, to first-order approximation, the numerator is equivalent in magnitude to $\mathbf{MK}\delta\mathbf{y}$; that is, the model evolution of the analysis increment. The denominator, on the other hand, is close to the evolution of the analysis error. The latter is expected to be larger than the evolution of the increment because the observations typically reduce only a small fraction (less than half) of the background error.

6. Combined use of OSEs and adjoints

The results presented thus far, whether based on OSEs or the adjoint method, measure only the net effects of observations on the forecast. They provide little insight into the behaviour of the data assimilation system itself, including likely dependencies and redundancies between observing system impacts as observations are added or removed from the system. Such information, which may be useful for improving data selection decisions, is implicitly available in an OSE in terms of the responses of the *remaining* observations when a given set of observations is removed. These responses can be measured through the combined use of OSEs and the adjoint method by applying the latter to the perturbed OSE members and comparing the impacts of the remaining observing systems with those in the control experiment. Changes in these impacts reveal changes in the effective influence (weighting) of the remaining observations that ultimately determine the outcome and interpretation of, for example, forecast skill scores, but would otherwise likely go undetected. Examples of these calculations are presented here for a subset of the OSEs performed in the present study.

Adjoint results based on (7) with the global error norm as the response function were produced for the following experiments in addition to the control for both January 2006 and July 2005: ‘no amsua3’, ‘no airs’, ‘no raob’, ‘no satwind’ and ‘no qkscat’. Of these, all but the ‘no qkscat’ experiment were found to produce substantial changes in the responses of one or more of the remaining observing systems with respect to the control experiment. It was deemed unnecessary for the purposes of this demonstration to conduct adjoint experiments for all OSEs and response functions examined in previous sections.

As an example, Fig. 12 compares the fractional impacts, computed as in (8), of the various observing systems in the control experiment with those in the ‘no amsua3’, ‘no airs’ and ‘no raob’ experiments for January 2006. For display purposes, the observing system removed in a given OSE is plotted with zero value. It can be seen that three of the five largest contributors in the control experiment, including AMSU-A, AIRS and aircraft observations, exhibit a high degree of sensitivity to the removal of one or more other observing systems in the OSEs. Specifically, the removal of all AMSU-A radiances results in a near doubling of the fractional impact of AIRS from just over 10% to almost 20%, as well as small increases in the impacts of rawinsondes and surface observations. At the same time, the contributions from other radiance data that have small impacts in the control experiment, including AMSU-B, HIRS and MSU, increase by several factors when AMSU-A radiances are removed, though they remain small overall. The removal of AIRS radiances, in turn, produces a reciprocal but smaller increase in the impact of AMSU-A by approximately one-third, from 24% to 33%. The compensating effects seen here between AIRS and AMSU-A in particular may help explain the fact that some operational centres reported only small to moderate gains in forecast skill even when high-quality, high-volume satellite observations such as AIRS were added to the observing system (Le Marshall et al., 2006). Finally, it can be seen that the removal of rawinsonde observations, which are given large weight in the assimilation system, significantly affects the impacts of several observing systems. Their removal nearly doubles the impact of aircraft observations from 13% to 22%, and increases the impact of AMSU-A by approximately the same amount as does the removal of AIRS. In addition, there are proportionately large increases in the impacts of surface observations and scatterometer winds from QuikSCAT.

Figure 13 provides examples of other types of data dependencies, in this case by comparing observation impacts in the tropics for the control, ‘no amsua3’, ‘no raob’ and ‘no satwind’ experiments during July 2005. Here, we show contributions from observations in the tropics only to the reduction of the global error norm in each of these experiments. This calculation is much easier to perform than rerunning the adjoint system with a different response function. Since the influence of observations on the analysis is mostly local, these values provide a reasonable approximation to the contributions from the global set of observations to the reduction of the tropical error norm used previously to diagnose the control experiment. For example, in the control experiment for July 2005, the total reduction in the tropical error norm from the global set of observations

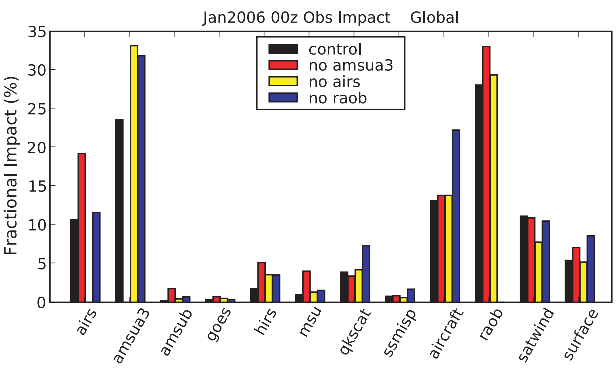

Fig. 12. Adjoint-based fractional impacts of various observing systems on the reduction in 24-h global forecast error during January 2006 for different OSEs.

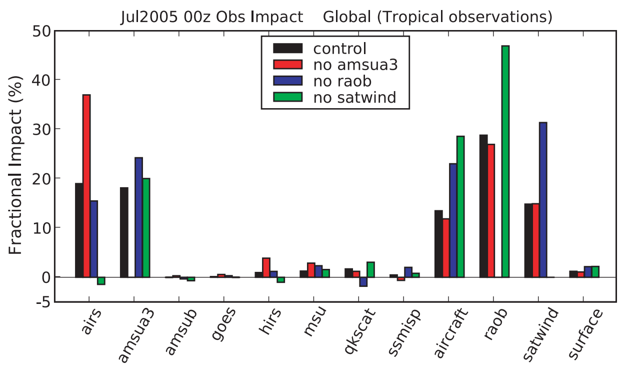

Fig. 13. As in Fig. 12, except for July 2005 and including only contributions from observations in the tropics to the reduction in global forecast error.

is just over 16.1 $\mathrm{J\,kg^{-1}}$, with the contribution from AMSU-A being 3.4 $\mathrm{J\,kg^{-1}}$, or 21%. This is comparable to the reduction of 15.3 $\mathrm{J\,kg^{-1}}$ in the global error norm from observations in the tropics alone, with the contribution from AMSU-A to this total being 2.8 $\mathrm{J\,kg^{-1}}$, or 18%.

There are large variations in the impacts of several observing systems in Fig. 13. The overall importance of wind observations in the tropics is evident from the variations in the impacts of rawinsondes and satellite winds, as well as aircraft observations. Removal of the satellite winds increases the impact of rawinsondes by more than two thirds compared with the control, from 28% to 47%. There is a reciprocal response in the impact of satellite winds to the removal of rawinsonde observations, which more than doubles with respect to the control experiment, increasing from 15% to more than 30%. It has been verified that the vast majority of the rawinsonde impacts seen here are due to the wind components of these observations: 81% of the total impact of the rawinsondes comes from the winds (with the zonal and meridional components contributing nearly equally, or approximately 40% each), while 18% comes from temperature. This is more than double the ratio in the Northern Hemisphere where the winds account for 63% of the rawinsonde impact while the temperature accounts for 35%; that is, the contributions from each wind component and the temperature are approximately equal in the Northern Hemisphere. In the case of the aircraft observations, their impact approximately doubles with respect to the control when either the rawinsondes or satellite winds are removed.

The response of AIRS in Fig. 13 reveals yet another type of data dependence with possible ramifications for the way balance relationships are imposed in the current data assimilation system. As observed previously, the removal of AMSU-A radiances results in a near doubling of the fractional impact of AIRS from 19% to 37% with respect to the control. In sharp contrast to this, however, the removal of the satellite winds erases the benefit of AIRS entirely such that the impact in fact becomes slightly detrimental to the forecast in terms of this metric. The impact of HIRS is similarly affected by the removal of the satellite winds, although the magnitude of the response is many times smaller than for AIRS. A similar result is obtained in the tropics during January 2006 (not shown), but in that case the removal of rawinsondes, rather than satellite winds, changes the impact of AIRS from moderately positive to substantially negative. The fact that the AIRS impact is similarly affected by different observing systems in January and July is another interesting aspect of these results. This difference may be related to seasonal shifts in the tropical circulation and, in the case of the satellite winds, corresponding changes in the spatial distribution of available observations.

The AIRS impact results point to a possible deficiency in the wind–mass relationship imposed through the background error covariances used in the analysis system. The current formulation imposes a linear balance relationship between the stream function and geopotential height fields such that observations of either the mass or wind field produce an increment in both fields in accordance with this relationship. The relationship is most appropriate for the extratropics and is therefore relaxed, but only partially, at tropical latitudes other than the equator. While AIRS provides substantial benefit to the forecast in the tropics when assimilated in conjunction with sufficient observations of the winds, the results here suggest that changes to the wind field induced by AIRS through the balance relationship alone may be detrimental at these latitudes. It is unclear at this time why the impact of AMSU-A is not similarly affected by the removal of the satellite winds. Among other things, differences in the spatial distribution and density of observations from these observing systems may play a role (twice as many AIRS radiances are assimilated in this region as AMSU-A radiances), but these issues require further investigation.

The results shown here are consistent with those of Žagar et al. (2008), who demonstrate in an idealized context the problems with current (three-dimensional) multivariate data assimilation schemes in the tropics, and especially the inability of the schemes to reconstruct the tropical wind field from mass observations and applied balances in the background error covariance model. These authors also provide requirements for additional wind observations in the tropics in assessing the potential impact of the upcoming Doppler wind lidar Atmospheric Dynamics Mission (ADM)-Aeolus, currently scheduled for launch in 2010.

7. Conclusions

Despite fundamental differences in their underlying methodologies and assumptions, OSEs and adjoint-derived measures of observation impact appear to provide consistent estimates of the overall ‘importance’ of most of the major observing systems assimilated in the GEOS-5 atmospheric data assimilation system, at least in terms of their contribution to the change in a mean squared (energy-based) metric of 24-h forecast error in several verification regions. The most important underlying difference between these methods is that, in the adjoint, the impacts of any or all observations are estimated simultaneously and in the context of all other observations present in a single execution of the assimilation system, while in an OSE, selected subsets of observations are removed from the assimilation system in separate experiments and the results compared to a control experiment including all observations. Although OSEs typically are used to measure the impacts of observations on forecasts out to several days or more, quantitative comparisons between the two methods are restricted by the forecast range and metric for which the adjoint results are valid on the one hand (24-h energy in this study) and by the selected observing systems removed in the OSEs on the other. More generally, the adjoint measures the response of a single forecast metric to all perturbations of the observing system, while an OSE measures the effect of a single perturbation on all forecast metrics.

Outside the tropics, the fractional impacts of most observing systems, defined uniquely in the OSE and adjoint contexts to measure the contributions of individual observing systems to the net change in the 24-h forecast error metric due to all observations, were found to be in reasonable to good quantitative agreement with relatively few exceptions. Most notably among the latter, the impact of satellite winds was found to be considerably smaller in the adjoint results than in the OSEs. Both the different treatments of the background information and lack of moist physical processes in the current adjoint model (the majority of the satellite winds are in the tropics) may contribute to this difference, although the current results are inconclusive in this regard. The removal of all AMSU-A radiances (from three satellites) in the OSEs also produced significantly larger impacts on forecast skill during winter in the Southern Hemisphere than the combined impact of these same observations as measured by the adjoint. This difference was attributed to the severe degradation of the analysis during southern winter when all AMSU-A data were removed. The two methods yielded similar results when measuring the impacts of up to two AMSU-A satellites during southern winter, and up to three AMSU-A satellites during northern winter. Within the tropics, the OSEs yielded fractional impacts that, while larger in absolute magnitude than those measured by the adjoint, were similar overall in terms of the relative contributions of the various observing systems to the quality of the forecast in that region.

Differences should be expected in the results produced by the two methods, and do not necessarily point to shortcomings in one or the other. When interpreting results, it is important to keep in mind that the adjoint measures the impact of observations in each analysis cycle separately and against the control background containing all previous information, while the OSEs measure the impact of removing observational information from both the background and analysis in a cumulative manner. This distinction can be significant, especially if an observing system contributes disproportionately to the quality of the analysis and subsequent background state. At the same time, the weights of the remaining observations can change substantially when large subsets of observations are removed from the assimilation system. In cases where these observations have a large impact on the quality of the analysis, as in the case of AMSU-A cited above, then other aspects of the data assimilation system including the specification of background and observation errors and accounting for biases may need to be reconsidered to obtain an optimal analysis with the reduced observation set; that is, the OSE might otherwise overestimate the true impact of these observations. No such adjustments were made in the OSEs conducted for this study.

Information gleaned from OSEs and adjoints should be viewed as complementary since both address relevant questions about how observations influence the quality of weather forecasts. While, ultimately, the impact of any one observing system occurs in the context of all other observations present in the assimilation system (as measured by the adjoint), it also important to understand how the absence or loss of an observing system (as measured by an OSE) affects the assimilation system, especially as it pertains to the quality of the background state. Not discounting the effects of deficiencies in the current GEOS-5 adjoint model, including the lack of moist physics, and without evidence to the contrary, it seems reasonable to speculate that the different treatment of the background information in the OSEs and adjoint method explains some of the observed differences in results. However, experiments designed to elucidate these effects in the present study suggest that it may not be straightforward to assess the extent to which this affects the comparison of results.

The combined use of OSEs and adjoints provides insights into how changing the mix of observations in a data assimilation system affects their impacts. By applying the adjoint to both the control and perturbed members of an OSE, it is possible to measure such changes as observations are added or removed from the assimilation system, thereby revealing dependencies and redundancies between observing systems not revealed by either method separately. Examples of these calculations in the present study showed significant compensatory impacts between, for example, AIRS and AMSU-A, as well as between rawinsondes and aircraft observations among other data types, and also revealed various aspects of the criticality of wind observations in the tropics. Given the large volume of data used in modern data assimilation systems, the net impact of any one observing system may reflect as much about the quality and quantity of other nearby observations as it does about that one observing system itself. Information about these dependencies may therefore be useful for making intelligent data selection decisions. Moreover, to the extent that redundancies between observing systems are considered desirable (or not) in designing an optimal global observing system for numerical weather prediction, such information also may be useful in identifying needs for future observations.

Having said this, we stress that the primary use of OSEs and the adjoint method described here is to assess the impact of existing observations. Requirements for the design of future observing systems pertaining to, for example, the specific type, accuracy and coverage of new observations are best determined by other methods including traditional observing system simulation experiments (OSSEs) and recently introduced sensitivity observing system experiments (SOSEs, Marseille et al., 2008). The latter is a variation of the OSSE in which adjoint sensitivity information is used to define a so-called pseudo true atmospheric state for simulating proposed observations in actual case studies.

Because of its computational efficiency and flexibility in allowing arbitrary aggregation of results by data type, location, channel, etc., the adjoint method used in this study is especially well suited for routine monitoring of observation impacts in an operational setting. The OSEs, on the other hand, provide critical, if less regularly available, information, especially for forecast ranges beyond a few days. The latter are also important since observation impacts can vary with forecast length, making it difficult to draw conclusions about the overall importance of observing systems based on a specific forecast range or verification metric. Used together, these methods should increase our capability to assess and understand the complex, complementary nature of how observations are used by a data assimilation system.

8. Acknowledgments

The authors thank Ricardo Todling for his development of key components of the GEOS-5 system and related tools used in this study, and Ron Errico for his guidance and insightful discussions of this work. We also acknowledge the efforts of two anonymous reviewers whose comments helped improve the paper. This work was supported by the Atmospheric Data Assimilation Development component of the NASA Modeling, Analysis and Prediction program (MAP/04-0000-0080).

References

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

English, S., Saunders, R., Candy, B., Forsythe, M. and Collard, A. 2004. Met Office satellite data OSEs. In: *Proceedings of Third WMO Workshop on the Impact of Various Observing Systems on Numerical Weather Prediction, Alpbach, Austria, WMO/TD No. 1228*, 146–156.

Errico, R. M. 2000. Interpretations of the total energy and rotational energy norms applied to determination of singular vectors. *Quart. J. R. Meteorol. Soc.* **126**, 1581–1599.

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

Gelaro, R., Rosmond, T. E. and Daley, R. 2002. Singular vector calculations with an analysis error variance metric. *Mon. Wea. Rev.* **130**, 1166–1186.

Gelaro, R., Zhu, Y. and Errico, R. M. 2007. Examination of various-order adjoint-based approximations of observation impact. *Meteorologische Zeitschrift* **16**, 685–692.

Kelly, G., McNally, T., Thépaut, J.-N. and Szyndel, M. 2004. OSEs of all main data types in the ECMWF operation system. In: *Proceedings of Third WMO Workshop on the Impact of Various Observing Systems on Numerical Weather Prediction, Alpbach, Austria, WMO/TD No. 1228*, 63–94.

Langland, R. H. and Baker, N. 2004. Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus* **56A**, 189–201.

Le Marshall, J., Jung, J., Derber, J., Chahine, M., Treadon, R. and co-authors. 2006. Improving global analysis and forecasting with AIRS. *Bull. Am. Meteorol. Soc.* **87**, 891–894.

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

Lord, S., Zapotocny, T. and Jung, J. 2004. Observing system experiments with NCEP's global forecast system. In: *Proceedings of Third WMO Workshop on the Impact of Various Observing Systems on Numerical Weather Prediction, Alpbach, Austria, WMO/TD No. 1228*, 56–62.

Marseille, G.-J., Stoffelen, A. and Barkmeijer, J. 2008. Sensitivity Observing System Experiment (SOSE)—a new effective NWP-based tool in designing the global observing system. *Tellus* **60A**, 216–233.

Orrell, D., Smith, L., Barkmeijer, J. and Palmer, T. N. 2001. Model error in weather forecasting. *Nonlinear Processes Geophys.* **8**, 357–371.

Rabier, F., Klinker, E., Courtier, P. and Hollingsworth, A. 1996. Sensitivity of forecast errors to initial conditions. *Quart. J. R. Meteorol. Soc.* **122**, 121–150.

Reynolds, C. A. and Gelaro, R. 2001. Remarks on Northern Hemisphere forecast error sensitivity from 1996–2000. *Mon. Wea. Rev.* **129**, 2145–2153.

Rienecker, M. M., Suarez, M. J., Todling, R., Bacmeister, J., Takacs, L. and co-authors. 2007. The GEOS-5 Data Assimilation System—documentation of versions 5.0.1 and 5.1.0. *NASA TM 104606, Technical Report Series on Global Modeling and Data Assimilation* Volume 27.

Talagrand, O. 1981. A study of the dynamics of four-dimensional data assimilation. *Tellus* **33**, 43–60.

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

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

Žagar, N., Stoffelen, A., Marseille, G.-J., Accadia, C. and Schlüssel, P. 2008. Impact assessment of simulated Doppler wind lidars with a multivariate variational assimilation in the tropics. *Mon. Wea. Rev.* **136**, 2443–2460.

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