MONTHLY WEATHER REVIEWVOLUME 137

A Posteriori Diagnostics in an Ensemble of Perturbed Analyses

GÉRALD DESROZIERS, LOÏK BERRE, VINCENT CHABOT, AND BERNARD CHAPNIK

GAME/CNRM, Météo-France, CNRS, Toulouse, France

ABSTRACT

Real-size assimilation systems basically rely on estimation theory in which two sources of information, background and observations, are combined in an optimal way. However, the optimality of such large problems is not guaranteed, since they rely on different approximations. The observed values of the subparts of the cost function measuring the distance between the analysis and the two sources of information can be compared to their theoretical statistical expectations. Such a posteriori diagnostics can be used to optimize the statistics of observation and also background errors. Moreover, the expectations of the subparts of the cost function associated with observations are related to the weights of the different sources of observations in the analysis. It is shown that those theoretical statistical expectations are direct by-products of an ensemble of perturbed assimilations.

1. Introduction

Most operational meteorological centers are now using a variational scheme to produce global or limited-area analyses (Lewis and Derber 1985; Talagrand and Courtier 1987; Rabier et al. 2000). Variational schemes allow the use of a large spectrum of observations in a very efficient way since today both the size of the observation set and of the model state vector are quite large. These analysis schemes rely on the theory of linear estimation (Talagrand 1997). In the linear estimation theory, the different pieces of information (i.e., the observations and the background estimate of the state vector provided by a short-range forecast) are given weights that are inversely proportional to their error covariances. However, those error statistics are not perfectly known. This is especially the case for background error statistics. The determination and representation of those error statistics remain a major challenge in assimilation schemes.

A potential way to determine background error statistics is to rely on an ensemble of perturbed analyses. Originally introduced by Evensen (1994), the idea of an ensemble of assimilations has been for example applied by Houtekamer et al. (1996) in the framework of en-

semble prediction and by Fisher (2003a) to improve the determination of background error statistics. In spite of their rather high numerical cost, large operational centers tend today to implement such Monte Carlo procedures to specify more realistic background error covariances (Berre et al. 2006; Belo Pereira and Berre 2006).

From a different point of view, Dee (1995) has defined a method based on the maximum likelihood to determine both observation and background error statistics.

Desroziers and Ivanov (2001) have proposed a method based on a consistency criterion originally proposed by Talagrand (1999). Chapnik et al. (2004) have further investigated the properties of the algorithm defined in Desroziers and Ivanov (2001) and have in particular shown that it was equivalent to a maximum likelihood approach. Chapnik et al. (2006) have applied this algorithm in an operational framework to determine observation error statistics. Such a tuning procedure has also been used by Buehner et al. (2005).

A specific feature of the previous algorithm is that it is based on a randomization approach that relies on a perturbation of observations. This is also the case in an ensemble of perturbed analyses, as it has been, for example, implemented by Fisher (2003a) or Berre et al. (2006). Such an ensemble of analyses is now operational at Météo-France in order to specify flow-dependent modulations of the background error variances in the operational four-dimensional variational data assimilation (4D-Var) assimilation scheme.

Corresponding author address: Gérald Desroziers, GAME/CNRM, Météo-France, CNRS, 42, av. G. Coriolis, Toulouse, 31057 CEDEX, France.
E-mail: gerald.desroziers@meteo.fr

DOI: 10.1175/2009MWR2778.1

© 2009 American Meteorological Society

Unauthenticated | Downloaded 03/25/23 09:18 PM UTC

An objective of this paper is to show that an ensemble of perturbed analyses also allows the computation of consistency diagnostics and then the determination of more exact observation and background error variances (Berre et al. 2007).

Another point is that such consistency diagnostics are related to quantities that appear to be the respective sensitivities of the analysis to the different subsets of observations. A second objective of the paper is then to show that an ensemble of perturbed analyses also provides, without any significant additional numerical cost, the weights of the different subsets of observations in an assimilation cycle.

In the next section, the ensemble approach is introduced. In section 3, the consistency diagnostics and the way they can be computed in an ensemble of assimilations are given. Section 4 explains how the analysis sensitivities to the observations can also be seen as byproducts of an ensemble of assimilations. The ensemble of perturbed analyses implemented at Météo-France in the global Action de Recherche Petite Echelle Grande Echelle (ARPEGE) model is described in section 5. The diagnostic of observation and background error statistics in the ARPEGE ensemble is shown in section 7. The analysis sensitivities to the different sets of observations are finally shown in the same operational framework. Conclusion and perspectives are given in section 9.

2. Ensemble formalism

In an assimilation cycle, the background $\mathbf{x}^b$ is given by the evolution of the previous analysis $\mathbf{x}^{a-}$ by the forecast model M. The subsequent analyzed state $\mathbf{x}^a$ is obtained as an optimal combination of the background and the observations $\mathbf{y}^o$. The two forecast and analysis steps are written as

$$\mathbf{x}^b = M(\mathbf{x}^{a-}) \quad \text{and} \tag{1}$$

$$\mathbf{x}^a = A(\mathbf{x}^b, \mathbf{y}^o), \tag{2}$$

where A stands for the possibly nonlinear analysis operator.

The estimate $\mathbf{x}^a$ can be classically obtained as the solution of the minimization of the following cost function:

$$\begin{aligned} J(\mathbf{x}) &= J^b(\mathbf{x}) + J^o(\mathbf{x}) \\ &= \left(\frac{1}{2}\right)\{(\mathbf{x}^b - \mathbf{x})^\mathrm{T}\mathbf{B}^{-1}(\mathbf{x}^b - \mathbf{x}) \\ &\quad + [\mathbf{y}^o - H(\mathbf{x})]^\mathrm{T}\mathbf{R}^{-1}[\mathbf{y}^o - H(\mathbf{x})]\}, \end{aligned} \tag{3}$$

where $J^b(\mathbf{x})$ and $J^o(\mathbf{x})$ are the background and observation terms, respectively. Matrices $\mathbf{B}$ and $\mathbf{R}$ stand for the background and observation error covariance matrices, respectively, and H is the possibly nonlinear observation operator including model integration in the 4D-Var formalism. Such a cost function can be solved by the minimization of a series of quadratic cost function with observation operators linearized around successive trajectories (Courtier et al. 1994).

It can be shown (see appendix A) that the following two equations stand for the evolution of the different errors involved in a forecast/analysis scheme, even in a slightly nonlinear analysis scheme such as 4D-Var:

$$\boldsymbol{\epsilon}^b = \mathbf{M}\boldsymbol{\epsilon}^{a-} + \boldsymbol{\epsilon}^m \quad \text{and} \tag{4}$$

$$\boldsymbol{\epsilon}^a = (\mathbf{I} - \mathbf{KH})\boldsymbol{\epsilon}^b + \mathbf{K}\boldsymbol{\epsilon}^o, \tag{5}$$

where $\mathbf{M}$ and $\mathbf{H}$ are linearized versions of the model M and the observation operator H, respectively, and

$$\mathbf{K} = \mathbf{BH}^\mathrm{T}(\mathbf{HBH}^\mathrm{T} + \mathbf{R})^{-1}. \tag{6}$$

Vectors $\boldsymbol{\epsilon}^b$, $\boldsymbol{\epsilon}^a$, $\boldsymbol{\epsilon}^m$, and $\boldsymbol{\epsilon}^o$ contain background, analysis, model, and observation errors, respectively, and $\boldsymbol{\epsilon}^{a-}$ is the analysis error at the previous analysis step [such a linearization of the analysis step can also be found in Fisher and Andersson (2001), Zagar et al. (2005), or more recently in Daescu (2008)].

As underlined in Berre et al. (2006), this derivation of background and analysis errors holds even if the analysis is not perfect. The aim of an ensemble of perturbed forecasts–analyses is to mimic the evolution of the estimation errors appearing in the previous equations. For a given member l of an ensemble of L forecasts–analyses, the evolution of the perturbations applied to the unperturbed states satisfy the following equations:

$$\boldsymbol{\delta}_l^b = \mathbf{M}\boldsymbol{\delta}_l^{a-} + \boldsymbol{\delta}_l^m \quad \text{and} \tag{7}$$

$$\boldsymbol{\delta}_l^a = (\mathbf{I} - \mathbf{KH})\boldsymbol{\delta}_l^b + \mathbf{K}\boldsymbol{\delta}_l^o, \tag{8}$$

where $\boldsymbol{\delta}_l^b$ and $\boldsymbol{\delta}_l^a$ are the perturbations applied to the unperturbed background $\mathbf{x}^b$ and analysis $\mathbf{x}^a$, respectively. The vector $\boldsymbol{\delta}_l^m$ is a vector of perturbations consistent with the covariance matrix $\mathbf{Q}$ of model errors:

$$\boldsymbol{\delta}_l^m = \mathbf{Q}^{1/2}\boldsymbol{\eta}_l^m, \tag{9}$$

where $\boldsymbol{\eta}_l^m$ is a vector of n random numbers (where n is the size of the state vector) with a standard Gaussian distribution (mean 0 and variance 1). Similarly, the vector $\boldsymbol{\delta}_l^o$ is a vector of perturbations consistent with the covariance matrix $\mathbf{R}$ of observation errors:

$$\boldsymbol{\delta}_l^o = \mathbf{R}^{1/2}\boldsymbol{\eta}_l^o, \tag{10}$$

where $\boldsymbol{\eta}_l^o$ is a vector of p random numbers (where p is the number of observations) with a standard Gaussian distribution.

Because the equations for the evolution of the perturbations are the same as those for the state estimation errors, the two equations for the perturbations, $\boldsymbol{\delta}_l^b$ and $\boldsymbol{\delta}_l^a$, will allow perturbations to evolve in a similar way as estimation errors. Then, the covariances of the perturbations will document the covariances of the estimation errors. In practice, this happens to be a powerful and relatively low-cost method when compared to the extended Kalman filter scheme, to obtain approximations of the background and analysis error covariance matrices.

Another point is that, although ensemble Kalman filter implementations most often rely on a relatively large ensemble (typically with 100 elements), some useful information on background error covariances can be obtained with small ensembles (5–10 elements). In that case a spatial filtering of the noisy raw variances has to be applied (Berre et al. 2007; Raynaud et al. 2008). Pannekoucke et al. (2007) also showed that the wavelet approach could be an adapted technique to filter the noisy correlations derived from a small ensemble. The wavelet approach to represent correlations in the $\mathbf{B}$ matrix was first proposed by Fisher and Andersson (2001).

3. A posteriori diagnostics

a. Analysis consistency diagnostics

As in Talagrand (1997) and Desroziers and Ivanov (2001), it is possible to introduce an extended vector of observations $\mathbf{z}$, combining the proper observations $\mathbf{y}^o$, with dimension p, and the background vector $\mathbf{x}^b$, with the same dimension n as the unknown true state $\mathbf{x}^t$.

This is written as

$$\mathbf{z} = [(\mathbf{x}^b)^{\mathrm{T}}(\mathbf{y}^o)^{\mathrm{T}}]^{\mathrm{T}} \quad \text{with}$$
$$\mathbf{x}^b = \mathbf{x}^t + \boldsymbol{\epsilon}^b, \tag{11}$$

and

$$\mathbf{y}^o = H(\mathbf{x}^t) + \boldsymbol{\epsilon}^o. \tag{12}$$

Then, it is possible to write $\mathbf{z}$ under the form:

$$\mathbf{z} = \Gamma(\mathbf{x}^t) + \boldsymbol{\epsilon}, \tag{13}$$

where Γ is an extended observation operator and $\boldsymbol{\epsilon}$ is the vector of background and observation errors with dimension $n + p$.

An important property pointed out by Talagrand (1999) is that if J_i stands for a subterm of J with p_i ele-

ments, then, the statistical expectation of $J_i(\mathbf{x}^a)$, in an optimal system, is

$$E\{J_i(\mathbf{x}^a)\} = p_i - \text{Tr}(\Gamma_i \mathbf{P}^a \Gamma_i^{\mathrm{T}} \mathbf{S}_i^{-1}), \tag{14}$$

where Γ_j and $\mathbf{S}_i$ stand for the observation operator and the error covariance matrix, respectively, associated with those p_i elements [a proof of Eq. (14) is given in appendix B for a slightly nonlinear assimilation scheme such as 4D-Var]. The matrix $\mathbf{P}^a$ is here defined by

$$\mathbf{P}^a = (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}. \tag{15}$$

In particular, for $\Gamma_i = \mathbf{I}_n$ and $\mathbf{S}_i = \mathbf{B}$, and knowing that $\mathbf{K} = \mathbf{P}^a\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}$, it follows that

$$E\{J^b(\mathbf{x}^a)\} = \text{Tr}(\mathbf{HK}), \tag{16}$$

and for $\Gamma_i = \mathbf{H}$ and $\mathbf{S}_i = \mathbf{R}$,

$$E\{J^o(\mathbf{x}^a)\} = p - \text{Tr}(\mathbf{HK}), \tag{17}$$

and then

$$E\{J(\mathbf{x}^a)\} = E\{J^b(\mathbf{x}^a)\} + E\{J^o(\mathbf{x}^a)\} = p. \tag{18}$$

This means that, if the background and error statistics are correctly specified, then the expectation of the global cost function at its minimum should be equal to p. As pointed out by Bennett et al. (1993) and Talagrand (1999), this is a simple a posteriori consistency criterion of the analysis scheme. A deviation of the expectation of $J(\mathbf{x}^a)$ from p indicates a misspecification in the statistics of the innovation vector.

b. Practical computation in an ensemble of assimilations

In practice, the statistical expectation of $J(\mathbf{x}^a)$ often differs from p in an operational system, which means that there may be a misspecification in either the background or observation error covariances, or both, but in any case this simple criterion is not sufficient to indicate where the misspecification lies. To go further, one would like to compute the expectations $E\{J_i(\mathbf{x}^a)\}$, for any subpart of J, to examine where the deviation of the statistics of $J_i(\mathbf{x}^a)$ from their theoretical values are the largest.

The direct use of Eq. (14) to derive this theoretical statistical expectation of $J_i(\mathbf{x}^a)$ cannot be use in practice, since, in particular, $\mathbf{P}^a$ is not explicitly computed in a large size variational analysis scheme.

However, as presented in Desroziers and Ivanov (2001), it can be shown that if the observation errors of a subset of observations i are uncorrelated with the remaining part of the observation errors, then

$$E\{J_i^o(\mathbf{x}^a)\} = p_i - \text{Tr}(\mathbf{\Pi}_i\mathbf{HK}\mathbf{\Pi}_i^{\text{T}}), \qquad (19)$$

where $\mathbf{\Pi}_i$ is the projector that enables one to pass from the complete set of observations $\mathbf{y}^o$ to the particular subset $\mathbf{y}_i^o$.

Similarly,

$$E\{J_i^b(\mathbf{x}^a)\} = \text{Tr}(\mathbf{\Pi}_i\mathbf{KH}\mathbf{\Pi}_i^{\text{T}}), \qquad (20)$$

where $\mathbf{\Pi}_i$ is the projector that allows one to pass from the complete set of the elements of $\mathbf{x}^b$ to a subset $\mathbf{x}_i^b$ of elements, the errors of which are uncorrelated with the other errors of $\mathbf{x}^b$. Note that Eqs. (19) and (20) are still valid in a slightly nonlinear assimilation scheme such as 4D-Var [see appendix B for the proof of the generic expression in Eq. (14)].

Still following Desroziers and Ivanov (2001), it can be shown (see appendix C) that an estimation of $E\{J_i^o(\mathbf{x}^a)\}$ is given by

$$E\{J_i^o(\mathbf{x}^a)\} = \frac{1}{L}\sum_l [\boldsymbol{\delta}_{l,i}^{o\text{T}}\mathbf{R}_i^{-1}\mathbf{\Pi}_i(\mathbf{I}_p - \mathbf{HK})\boldsymbol{\delta}_l^o] \qquad (21)$$

$$= \frac{1}{L}\sum_l [\boldsymbol{\delta}_{l,i}^{o\text{T}}\mathbf{R}_i^{-1}(\mathbf{\Pi}_i - \mathbf{H}_i\mathbf{K})\boldsymbol{\delta}_l^o], \qquad (22)$$

where $\boldsymbol{\delta}_l^o = \mathbf{R}^{1/2}\boldsymbol{\eta}_l^o$ is a vector of perturbations of the complete set of observations; $\mathbf{R}_i$ and $\mathbf{H}_i$ are the error covariance matrix and the observation operator, respectively, associated with the observation subset i; and $\boldsymbol{\delta}_{l,i}^o = \mathbf{\Pi}_i\boldsymbol{\delta}_l^o$ is the projection of $\boldsymbol{\delta}_l^o$ onto the subset of observations i.

Now, because $\boldsymbol{\delta}_l^a = (\mathbf{I} - \mathbf{KH})\boldsymbol{\delta}_l^b + \mathbf{K}\boldsymbol{\delta}_l^o$ in an ensemble of assimilations and also because the $\boldsymbol{\delta}_l^b$ and $\boldsymbol{\delta}_l^o$ perturbations are uncorrelated, a Monte Carlo estimation of $E\{J_i^o(\mathbf{x}^a)\}$ is given by

$$E\{J_i^o(\mathbf{x}^a)\} \simeq \frac{1}{L}\sum_l \boldsymbol{\delta}_{l,i}^{o\text{T}}\mathbf{R}_i^{-1}(\boldsymbol{\delta}_{l,i}^o - \mathbf{H}_i\boldsymbol{\delta}_l^a). \qquad (23)$$

Similarly, a Monte Carlo estimation of a subpart of $E\{J_i^b(\mathbf{x}^a)\}$ is given by

$$E\{J_i^b(\mathbf{x}^a)\} \simeq \frac{1}{L}\sum_l \boldsymbol{\delta}_{l,i}^{b\text{T}}\mathbf{B}_i^{-1}\mathbf{\Pi}_i\boldsymbol{\delta}_l^a, \qquad (24)$$

where $\boldsymbol{\delta}^b = \mathbf{B}^{1/2}\boldsymbol{\eta}^b$ is a vector of perturbations of the complete set of elements of $\mathbf{x}^b$, $\mathbf{B}_i$ is the background error covariance matrix associated with a subset i of elements of $\boldsymbol{\epsilon}^b$ uncorrelated with the other elements of $\boldsymbol{\epsilon}^b$, and $\boldsymbol{\delta}_{l,i}^b = \mathbf{\Pi}_i\boldsymbol{\delta}_l^b$ is the projection of $\boldsymbol{\delta}_l^b$ onto the subset i [see appendix C for a proof of Eqs. (23) and (24)].

Therefore, a key point of the present paper is that, since an ensemble of perturbed assimilations is based on perturbations of observations $\boldsymbol{\delta}_l^o$ and background $\boldsymbol{\delta}_l^b$, then, such an ensemble should provide a way to evaluate the previous theoretical statistical expectations of subparts of the cost function.

Given the ensemble observation perturbations $\boldsymbol{\delta}_l^o$ and the corresponding $\boldsymbol{\delta}_l^a$ on the analysis, it is easy to compute the $E\{J_i^o(\mathbf{x}^a)\}$ since the $\mathbf{R}_i$ matrix associated with the $\boldsymbol{\delta}_{l,i}^o$ is known (by construction $\boldsymbol{\delta}_l^o = \mathbf{R}^{1/2}\boldsymbol{\eta}_l^o$). Most often, $\mathbf{R}_i$ is a diagonal matrix in operational systems, which also facilitates the computation.

Because the density of satellite observations is steadily increasing and some of these observations might be correlated, implementations of nondiagonal $\mathbf{R}_i$ matrices are also under consideration (Bormann et al. 2003). In that case, $\mathbf{R}_i$ is however built under an operator form, which would make the computation of Eq. (23) still possible.

The diagnostic of the $E\{J_i^b(\mathbf{x}^a)\}$ through the use of Eq. (24) is more complicated, since in an ensemble of assimilations the perturbations on the background are not drawn from a given covariance matrix $\mathbf{B}$, but are rather the result of the evolution of the analysis perturbations by the forecast model. While in the future, an ensemble estimate of $\mathbf{B}_i^{-1}$ may be considered for use, no application of this expression is shown in the present paper. However, the computation of $E\{J_i^b(\mathbf{x}^a)\}$ for the whole J^b part, which is equal to $\text{Tr}(\mathbf{KH}) = \text{Tr}(\mathbf{HK})$, can be obtained indirectly by adding the $\text{Tr}(\mathbf{\Pi}_i\mathbf{HK}\mathbf{\Pi}_i^{\text{T}})$ associated with the different subsets of observations (see section 7b).

c. Diagnostic of error variances

As discussed in Desroziers and Ivanov (2001), the possibility to compute the theoretical statistical expectation of the subparts of $J(\mathbf{x}^a)$ also enables to optimize the error statistics associated to subsets of vector $\mathbf{z}$. This can be done by rewriting the original cost function under the following form:

$$J(\mathbf{x}^a) = \sum_i \frac{1}{s_i^{b^2}} J_i^b(\mathbf{x}^a) + \sum_i \frac{1}{s_i^{o^2}} J_i^o(\mathbf{x}^a), \qquad (25)$$

where $s_i^{b^2}$ and $s_i^{o^2}$ are weighting parameters of background and observation errors, respectively, supposed to be homogeneous over a subset i. The rationale of the tuning procedure is to find the values of those weighting coefficients such as $(1/s_i^{b^2})J_i^b(\mathbf{x}^a)$ and $(1/s_i^{o^2})J_i^o(\mathbf{x}^a)$ become close to the respective expected statistical estimations $E\{J_i^b(\mathbf{x}^a)\}$ and $E\{J_i^o(\mathbf{x}^a)\}$ they would have if the system were optimal. This produces

$$s_i^{b^2} = \frac{J_i^b(\mathbf{x}^a)}{E\{J_i^b(\mathbf{x}^a)\}} = \frac{J_i^b(\mathbf{x}^a)}{\text{Tr}(\mathbf{\Pi}_i\mathbf{KH}\mathbf{\Pi}_i^{\text{T}})} \qquad (26)$$

and

$$s_i^{o2} = \frac{J_i^o(\mathbf{x}^a)}{E\{J_i^o(\mathbf{x}^a)\}} = \frac{J_i^o(\mathbf{x}^a)}{p_i - \mathrm{Tr}(\mathbf{\Pi}_i \mathbf{KH} \mathbf{\Pi}_i^{\mathrm{T}})}. \qquad (27)$$

Those equations are nonlinear in s_i^{b2} and s_i^{o2}, but their particular forms suggest the use of an iterative fixed-point method to solve them.

Desroziers and Ivanov (2001) and Chapnik et al. (2004) have shown that, in practice, the convergence of the method is very fast and that the values of the s_i^{b2} and s_i^{o2} obtained at the first iteration are very close to those obtained after several iterations.

Chapnik et al. (2004) have also shown that the s_i^{b2} and s_i^{o2}—verifying the previous relations—are the solutions of the application of a maximum likelihood method to estimate these parameters (Dee and da Silva 1999). Dee and da Silva (1999) and Chapnik et al. (2004) have demonstrated that the ability of the maximum likelihood method—and then of the previous method—to provide correct estimations of the error variances relies on the fact that a part of the innovation vector is spatially correlated and that the other part is not. This hypothesis is also assumed in Hollingsworth and Lönnberg (1986) to evaluate the background and observation variances. In practice, this hypothesis is most often valid, since background errors are spatially correlated and observation errors are not.

Moreover, Chapnik et al. (2004) have shown that observation and background error variances provided by the algorithm are valid even if background error correlations are not correctly specified in matrix **B**.

However, in some cases, where observation errors are correlated, such tuning methods will be unable to provide valid estimations of observation error variances.

4. Degree of freedom for signal

An increased amount of research has been done recently in numerical weather prediction to assess the observation impact on analyses and subsequent short-range forecasts. Langland and Baker (2004) have for example proposed a procedure for estimating the impact of observations on a measure of short-range forecast error, using adjoint versions of the forecast model and the data assimilation procedure. A similar approach was followed by Zhu and Gelaro (2008). Desroziers et al. (2005b) have proposed a randomization procedure for evaluating the error variance reduction brought by observations on analyses and forecasts. Trémolet (2008) has also recently introduced a way of computing the adjoint of the assimilation scheme in an incremental variational formalism.

The analysis sensitivity to a particular subset i of observations can also be given by its degree of freedom for

signal (DFS) introduced by Rodgers (2000). This quantity is defined by

$$\mathrm{DFS}_i = \mathrm{Tr}\left[\frac{\partial H_i(\mathbf{x}^a)}{\partial \mathbf{y}_i^o}\right]. \qquad (28)$$

It measures the sensitivity of the analysis to a perturbation of a particular subset of observations and then the weight of these observations in the analysis. Cardinali et al. (2004) have used the DFS to measure the analysis sensitivity to observations in a real size assimilation system. Rabier et al. (2002) have also shown how to use such a diagnostic to select Infrared Atmospheric Sounding Interferometer (IASI) channels in order to extract the useful information from the very large amount of data provided by this instrument.

Fisher (2003b) has investigated the possibility to compute the total DFS brought by the complete set of observations in a real-size data assimilation. He has compared different methods to compute such a quantity. One of them is based on the estimation of the so-called influence matrix **HK** by a randomization procedure, as proposed by Girard (1987) and also by Hutchinson (1989). Such a randomization procedure has also been used by Wahba et al. (1995) to compute the generalized cross-validation criterion and inspired the randomized estimation of the $E\{J_i^o\}$ proposed by Desroziers and Ivanov (2001).

Again, since an ensemble of perturbed analyses is based on explicit perturbations of observations and implicit perturbations of the background, it should provide estimations of the DFS_i, with nearly no additional computational cost.

Hence, it can be shown (Chapnik et al. 2006) that the previous expression of the partial DFS can be rewritten as

$$\mathrm{DFS}_i = \mathrm{Tr}(\mathbf{\Gamma}_i \mathbf{P}^a \mathbf{\Gamma}_i^{\mathrm{T}} \mathbf{S}_i^{-1}), \qquad (29)$$

where i is a subset of observations.

One can recognize a part of Eq. (14) of the expectation of a subpart of the cost function. Talagrand (1999) has interpreted this expression as a measure of the contribution of the subset of observations i to the overall precision.

Relying on Eq. (8), the DFS_i can be also rewritten, even in a nonlinear analysis,

$$\mathrm{DFS}_i = \mathrm{Tr}(\mathbf{\Pi}_i \mathbf{HK} \mathbf{\Pi}_i^{\mathrm{T}}), \qquad (30)$$

that can be estimated by a Monte Carlo procedure (see appendix C):

$$\mathrm{DFS}_i \simeq \frac{1}{L} \sum_l \boldsymbol{\delta}_{l,i}^{o\mathrm{T}} \mathbf{R}_i^{-1} \mathbf{H}_i \boldsymbol{\delta}_l^a. \qquad (31)$$

Then, it appears that the analysis sensitivities to the different subsets of observations, through their DFS_i,

are available as direct by-products of an ensemble of perturbed analyses associated with a real-size assimilation scheme.

5. ARPEGE ensemble of perturbed assimilations

The French operational ARPEGE global model is a spectral model with 60 vertical levels and a stretched geometry that allows a higher resolution over France. With a T538 spectral resolution and a C2.4 stretching factor, the maximum spectral resolution is T1290, which corresponds to about 15 km in physical space.

The assimilation scheme associated with the ARPEGE model is based on an incremental 4D-Var formulation with two outer loops and with T107 and T149 spectral uniform resolution increments, respectively.

The variational formulation allows a wide range of observations to be used. The types of observations used in the ARPEGE assimilation are the following (listed by ARPEGE type number and code):

1) Synop (surface observations),
2) Aircraft,
3) Satwind (satellite-derived winds),
4) Driftbuoy,
5) Radiosounding,
6) Profiler,
7) Brightness temperature ("Radiance"),
9) Scatterometer, and
10) GPSRO (GPS radio occultation).

The representation of matrix **B** in the analysis scheme is based on a multivariate formulation (Derber and Bouttier 1999), that includes a normalization in physical space of the vorticity covariances by a field of standard deviations of vorticity background errors. This allows us to introduce some heterogeneity of the covariances, which especially induces a larger weight of observations in the areas where the vorticity error standard deviations are large, like in extratropical areas over oceans.

To make the fields of vorticity error standard deviations dependent on the meteorological situation and also on the data density, an ensemble of six perturbed analyses is now run operationally every 6 h, like for the operational 4D-Var. To reduce the cost of such an ensemble, the perturbed analyses are performed with a 3D-Var first guess at observation time (FGAT) scheme instead of 4D-Var and at a lower resolution than the unperturbed 4D-Var (T359 spectral uniform resolution for the nonlinear forecasts and T107 spectral resolution for the analysis increment). The 3D-Var FGAT scheme can be seen as a simplified implementation of 4D-Var, where the evolution of the analysis increment is neglected during the as-

simulation period (see e.g., Munro et al. 2004). Preliminary studies showed that the background error variances given by 4D-Var and 3D-Var FGAT were relatively close.

The set of observations used in the ensemble is the same as in the operational 4D-Var.

Since the perturbed backgrounds are different from each other and also from the unperturbed operational background, the quality control of observations can lead to different rejections of observations, especially for satellite data, since this quality control is partially based on the amplitudes of the differences between the observations and the background. This is part of the nonlinearity of the forecasting/assimilation system that must be kept in order to simulate the dynamics of the background and analysis errors in a real-size system more closely.

For the time being, no model error perturbations are added. As a consequence the background errors deduced from the ensemble of assimilations are known to show too low variances and have to be amplified.

6. Estimation of $\mathrm{Tr}(\mathbf{\Pi}_i \mathbf{HK\Pi}_i^{\mathrm{T}})$ in the ARPEGE ensemble

As shown in the previous sections, the $\mathrm{Tr}(\mathbf{\Pi}_i \mathbf{HK\Pi}_i^{T})$ are key quantities, since they appear in the $E\{J_i^o(\mathbf{x}^a)\}$ and then in the tuning procedure of the observation error variances, but also in the measure of the analysis sensitivities to the different sets of observations since they are directly equal to the DFS_i.

As discussed above, the $\mathrm{Tr}(\mathbf{\Pi}_i \mathbf{HK\Pi}_i^{T})$, associated with a subset of observations i, can be evaluated by using the following randomized estimation:

$$\mathrm{Tr}(\mathbf{\Pi}_i \mathbf{HK\Pi}_i^{\mathrm{T}}) = \frac{1}{L} \sum_l \boldsymbol{\delta}_{l,i}^{o\mathrm{T}} \mathbf{R}_i^{-1} \mathbf{H}_i \boldsymbol{\delta}_l^a. \quad (32)$$

Because there is no unperturbed member in the ARPEGE 3D-Var FGAT ensemble of assimilations, the perturbations $\boldsymbol{\delta}_l^a$ are calculated as the differences between two perturbed analyses $\mathbf{x}_l^a$ and $\mathbf{x}_{l+1}^a$, as in Fisher (2003a). Accordingly, $\boldsymbol{\delta}_{l,i}^o$ is the difference between two perturbed sets of observations $\mathbf{y}_{l,i}^o$ and $\mathbf{y}_{l+1,i}^o$, so that the previous expression becomes

$$\frac{1}{(2L)} \sum_{l,l'} (\mathbf{y}_{l',i}^o - \mathbf{y}_{l,i}^o)^{\mathrm{T}} \left(\frac{1}{\sigma_i^o}\right)^2 [H_i(\mathbf{x}_{l'}^a) - H_i(\mathbf{x}_l^a)], \quad (33)$$

with $l' = l + 1$, for $l = 1$ to $L - 1$ and $l' = 1$ for $l = L$.

Note that the $\boldsymbol{\delta}^a$ perturbations could have been computed as the differences between the perturbed 3D-Var FGAT analyses and the control 4D-Var analysis. However, since the perturbed 3D-Var FGAT analyses are also performed at a lower resolution than the control 4D-Var, such a computation of the $\boldsymbol{\delta}^a$ perturbations
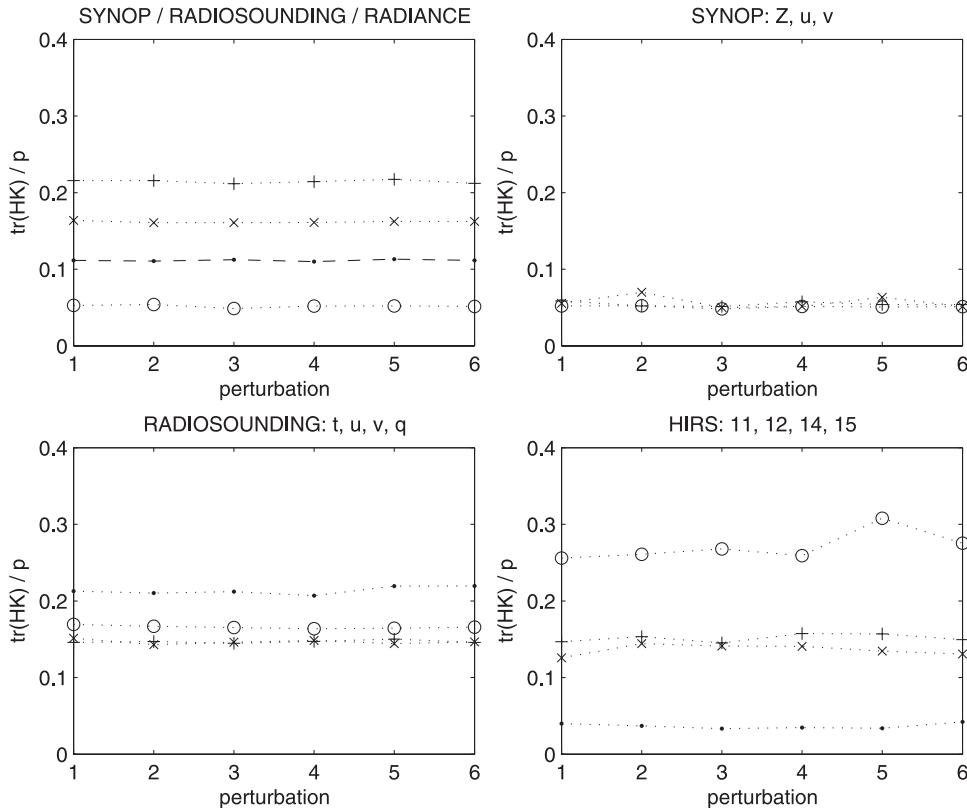
FIG. 1. Tr($\mathbf{\Pi}_i\mathbf{HK\Pi}_i^{\mathrm{T}}$)/$p_i$ for the different ARPEGE ensemble perturbations and a single date (0000 UTC 1 Jun 2008). (top left) Synop (dotted line with open circles), Radiosounding (dotted line with times signs), Radiance (dotted line with plus signs), and all observations (dashed line). (top right) Synop observations: geopotential Z (dotted line with open circles), zonal wind u (dotted line with times signs), and meridional wind v (dotted line with plus signs). (bottom left) Radiosounding observations: temperature t (dotted line with open circles), zonal wind u (dotted line with times signs), meridional wind v (dotted line with plus signs), and specific humidity (dotted line with small filled circles). (bottom right) HIRS Radiance channels: 11 (dotted line with open circles), 12 (dotted line with plus signs), 14 (dotted line with times signs), and 15 (dotted line with small filled circles).

would not be consistent (another possibility, which is envisaged, would be to perform an additional unperturbed analysis in the same low-resolution 3D-Var FGAT mode).

Equation (33) includes an additional normalization factor of 2, because of the fact that it involves perturbation differences (instead of single perturbation values). The operation $(1/\sigma_1^o)^2$ corresponds to a simple division by the variances of the observation errors used for the generation of the observation perturbations, since observation errors are, up to now, assumed to be uncorrelated in the ARPEGE assimilation. Here, $H_i(\mathbf{x}_l^a)$ stands for the evaluation of the perturbed analysis $\mathbf{x}_l^a$ for the subset of observations i by using the possibly nonlinear observation operator H_i associated with this particular subset.

In practice, the previous computation is rather easy to implement, because all the necessary elements, $\mathbf{y}_{l,i}^o$, $H_i(\mathbf{x}_l^a)$, σ_i^o, stand in the archived perturbed observation databases. Furthermore, the scalar product appearing in

the equation is simply the computation of the sum, over all the observations of subset i and all the perturbations l, of the product, observation by observation, of the difference between two observation perturbations and the difference between the two corresponding perturbed analyses evaluated at observation locations, normalized by the corresponding observation error variance.

Figure 1 shows the Tr($\mathbf{\Pi}_i\mathbf{HK\Pi}_i^{\mathrm{T}}$) quantities for a single date and three subsets of observations—Synop, Radiosounding, and satellite Radiance—and for the six perturbations l of the ARPEGE ensemble. To make those quantities comparable, they are divided by the corresponding numbers of observations p_i.

The stability of the Tr($\mathbf{\Pi}_i\mathbf{HK\Pi}_i^{\mathrm{T}}$)/$p_i$ is high along the perturbations. This is especially the case for the whole set of observations or for the whole set of parameters of a type of observations. Of course, the stability of the randomized quantities is slightly diminished when they
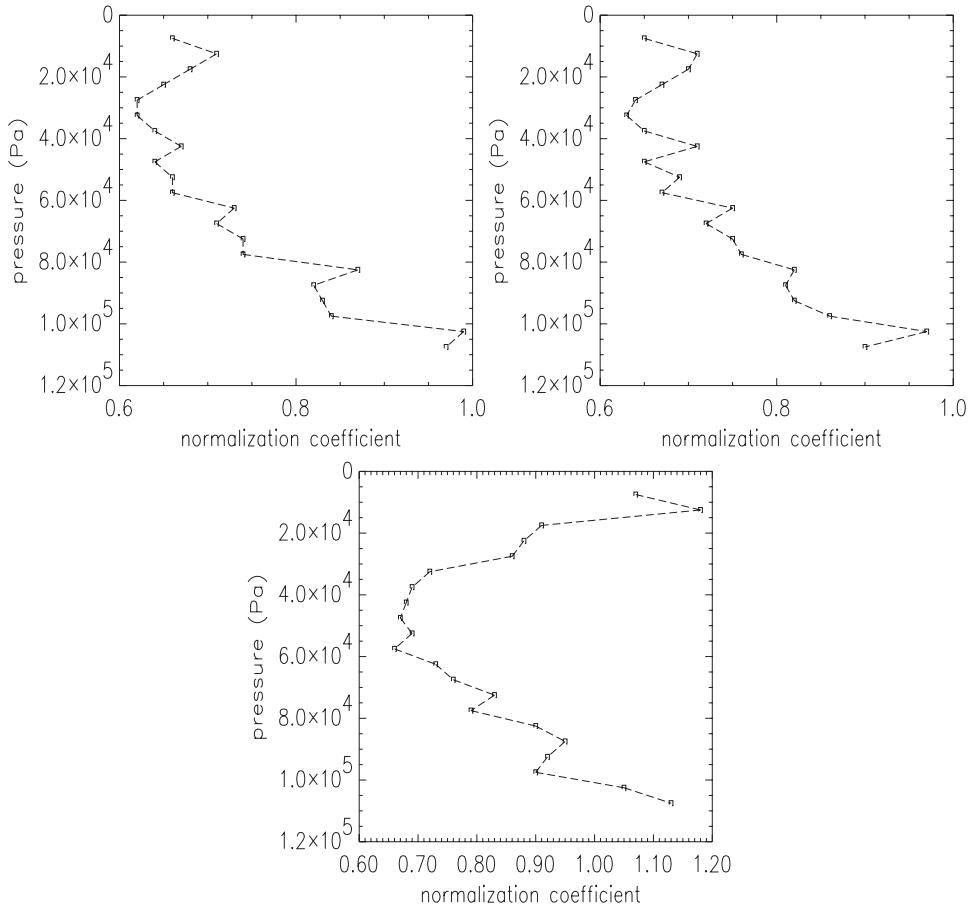
FIG. 2. Normalization coefficients s_i^o for the Radiosounding observation error standard deviations: (top left) zonal wind, (top right) meridional wind, and (bottom) temperature.

are associated with a more limited subset of observations, but it remains acceptable. This is likely due to the fact that the computations are done in a global system where the number of observations is quite large. Moreover, such quantities can be averaged over all the perturbations and also over different dates along a given period, that will still increase the robustness of their estimations.

Therefore, the possibility to compute the $\mathrm{Tr}(\mathbf{\Pi}_i\mathbf{H}\mathbf{K}\mathbf{\Pi}_i^{\mathrm{T}})/p_i$ quantities over more than one perturbed analyses can be useful to improve their evaluation. The dispersion of the values obtained for the different perturbations can also indicate the confidence to give to these evaluations.

7. Error variance diagnostic in the ARPEGE ensemble

a. Tuning of observation error variances

As indicated in the previous sections, a first application of the estimation of the $\mathrm{Tr}(\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{\Pi}_i\mathbf{H}_i\mathbf{K})$ quantities is to use them to diagnose and optimize the observation error variances of the different subsets of observations.

Such a diagnostic has been systematically performed for all the observations used in the ARPEGE global assimilation. For some observations, such as Aircraft, Satwind, Radiosounding, and Profiler, the variances specified in matrix $\mathbf{R}$ depend on the vertical.

Figure 2 shows the tuning coefficient s_i^o diagnosed by the randomization procedure for wind and temperature measurements given by the Radiosounding type. The coefficients have been determined over an 18-day period, from 3 to 20 June 2008. The coefficients indicate that the wind error standard deviations should be kept unchanged in the lowest levels but multiplied by a factor around 0.6 in the highest levels. Note that the diagnosed normalization coefficients are similar for the zonal and meridional components of wind, which is rather expected, but also gives confidence in the ensemble randomized diagnostic. For temperature, the standard deviations should be slightly increased in the lowest and highest levels, but decreased by a factor around 0.6 near the 500-hPa level.
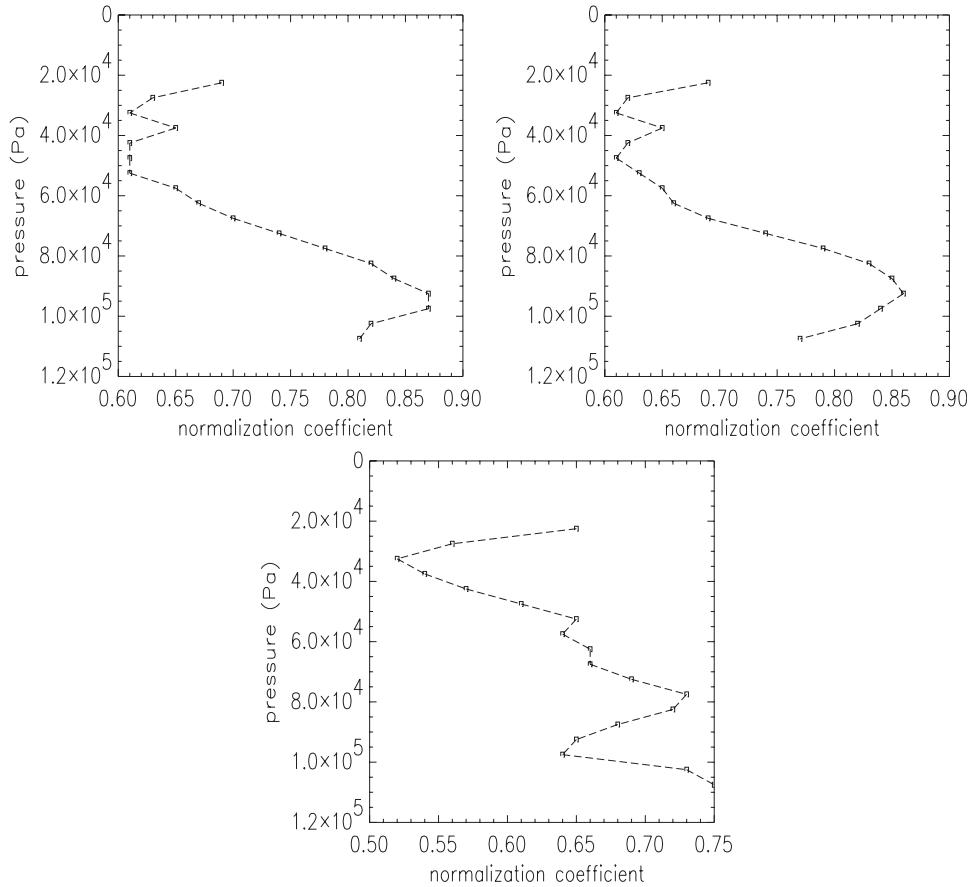
FIG. 3. As in Fig. 2, but for the Aircraft observation error standard deviations.

The normalization coefficients for the Aircraft observations (Fig. 3) indicate that the wind error standard deviations should be multiplied by values between 0.6 and 0.8, depending on the levels. Again, the normalization coefficients are very similar for the two components of wind. The values of the normalization coefficient are even lower for temperature than for wind, with values between 0.5 and 0.7.

The mean normalization coefficients by observation type (Fig. 4) show that there is a global overestimation of the observation error standard deviations in the ARPEGE assimilation. The needed normalization coefficients are around 0.6–0.8 for Synop, Aircraft, Driftbuoy, Radiosounding, Profiler, and GPSRO. They are close to 0.4–0.5 for Satwind, Radiance, and Scatterometer. For the three latter types of observations, the diagnosed large reductions of the observation error standard deviations may indicate a real overestimation of these standard deviations, but may be also due to the presence of error correlation for these types of observations. In that case, it is known that algorithms such as the one proposed by Desroziers and Ivanov (2001) will

fail to recover the right level of error variance and will, on the contrary, interpret observation error correlation as background error correlation. A way to check that the normalization coefficients are correct is to introduce
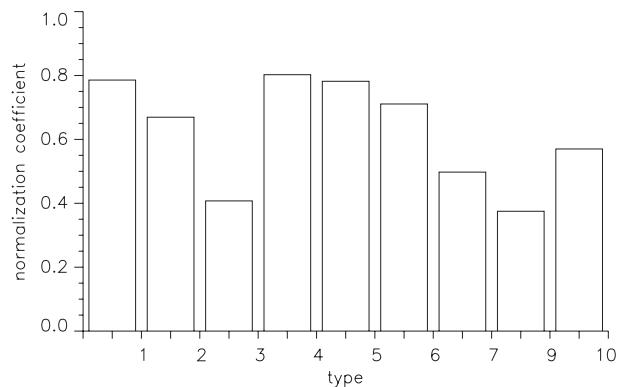

FIG. 4. Mean normalization coefficients s_i^o of observation error standard deviations by observation type: 1 = Synop, 2 = Aircraft, 3 = Satwind, 4 = Driftbuoy, 5 = Radiosounding, 6 = Profiler, 7 = Brightness temperature, 9 = Scatterometer, and 10 = GPSRO.

TABLE 1. The $J^b(\mathbf{x}^a)$ 4D-Var and 3D-Var FGAT, with unperturbed observations and backgrounds, compared with $E\{J^b(\mathbf{x}^a)\}$ at 0000, 0600, 1200, and 1800 UTC 1 Jun 2008.

	$J^b(\mathbf{x}^a)$ 4D-Var	$J^b(\mathbf{x}^a)$ 3D-Var FGAT	$E\{J^b(\mathbf{x}^a)\}$ $\simeq$ Tr(**HK**)
0000 UTC	21 053.59	24 215.48	49 799.64
0600 UTC	13 411.87	14 767.54	35 497.05
1200 UTC	18 120.32	20 589.83	35 899.43
1800 UTC	14 665.56	15 860.79	32 296.01

them in the analysis and to verify that the new diagnosed normalization coefficients are close to the previous ones.

b. Tuning of background error variances

As above-mentioned, the use of Eq. (24), to diagnose background variances, is difficult since the ensemble background perturbations are not drawn from a known covariance matrix, as it is, on the contrary, the case for observations.

However, if a single normalization factor of the **B** matrix is sought, only the expectation $E\{J^b(\mathbf{x}^a)\}$ is needed for the complete background vector. Moreover, as mentioned in section 3, $E\{J^b(\mathbf{x}^a)\}$ should be equal to Tr(**HK**) in a slightly nonlinear assimilation. The estimation of Tr(**HK**) can be in turn deduced from the sum over all the observations of the Tr($\mathbf{\Pi}_i\mathbf{HK}\mathbf{\Pi}_i^T$) associated with the different subsets of observations.

Table 1 shows the $J^b(\mathbf{x}^a)$ with unperturbed observations and backgrounds in comparison with the randomized estimation of Tr(**HK**). Values are given for a particular date for the four 6-h assimilations. The values of $J^b(\mathbf{x}^a)$ are also shown for 3D-Var FGAT analyses performed with the same unperturbed observations and backgrounds used in the 4D-Var assimilations, as the ensemble of ARPEGE assimilations is also, up to now, performed in 3D-Var FGAT mode.

It can be seen that the values of $J^b(\mathbf{x}^a)$ are systematically slightly larger for 3D-Var FGAT than for 4D-Var. This is expected since observations are known to have more weight in 3D-Var FGAT than in 4D-Var. Of course, this is not evidence that 3D-Var FGAT performs better than 4D-Var. On the contrary, this may be interpreted as a better filtering of the observation errors in 4D-Var.

The analysis sensitivity to observations is larger at 0000 and 1200 UTC than at 0600 and 1800 UTC, which is simply because radiosondes are mainly taken at 0000 UTC and that the required normalization coefficient s^b can be computed with

$$s^b = \left[\frac{J^b(\mathbf{x}^a)}{E\{J^b(\mathbf{x}^a)\}}\right]^{1/2} = \left[\frac{J^b(\mathbf{x}^a)}{\text{Tr}(\mathbf{HK})}\right]^{1/2}. \quad (34)$$
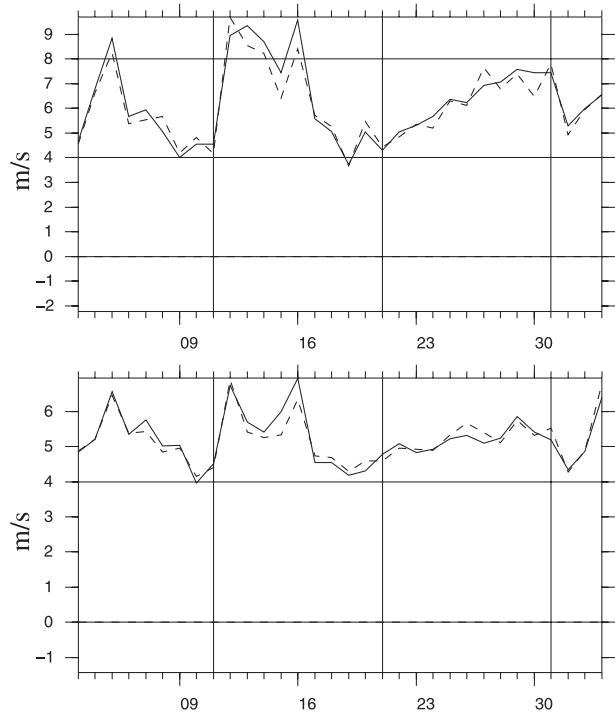

FIG. 5. RMS scores of 48-h wind forecasts vs radiosonde observations at 250 hPa: operational scores (solid line); scores with a tuning of observation error variances and a global tuning of background error variances (dashed lines). Scores have been computed from 1 Nov to 4 Dec 2008. They are shown for two domains: (top) North America and (bottom) 20°–90°N area.

It varies from 0.61 to 0.71 with 4D-Var and from 0.64 to 0.76 with 3D-Var FGAT (Table 1), which means that the average background error standard deviations are too large in the ARPEGE 4D-Var. This says, in turn, that the background is given too less confidence in the assimilation.

c. Impact of the error variance tuning

The tuned variances for observation errors and the diagnosed normalization coefficient s^b have been introduced in the 4D-Var scheme to measure the impact of such a tuning on the subsequent short-range forecasts.

Figure 5 shows an example of the root-mean-square (RMS) scores of the forecasts issued from the tuned analyses, compared to the operational scores. The scores are shown here for the 48-h wind forecasts against radiosondes at 250 hPa, for North America and 20°–90°N domains. The scores are relatively neutral globally, although the impact is slightly positive during the 12–17 November period. This relative neutrality is related to the fact that observation and background error variances tend to be reduced here, by chance, by the same amount of tuning.

An observed side effect of the reduction of both observation and background errors is that more observations are rejected, since a first part of the ARPEGE quality
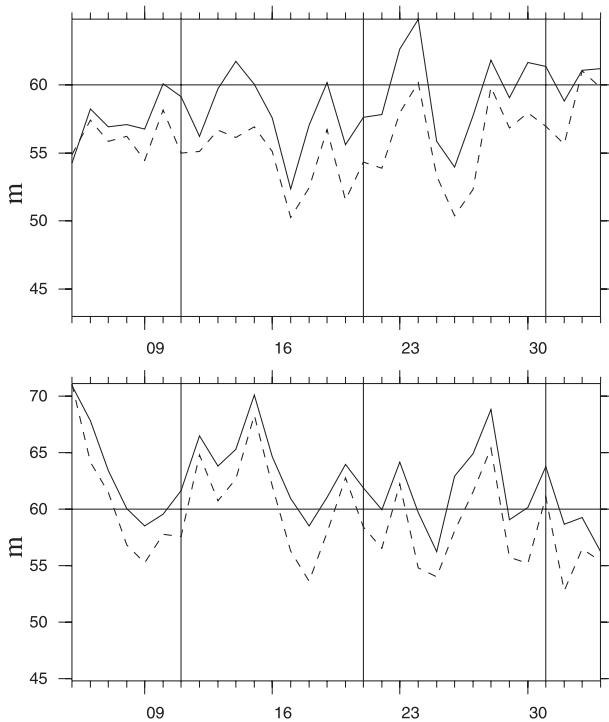
FIG. 6. As in Fig. 5, but for RMS scores of 96-h geopotential forecasts against ECMWF analyses at 20 hPa. Scores are shown for two domains: (top) 20°–90°N area and (bottom) 20°–90°S area.

control is done by comparing the squared innovations with the corresponding sums of the background and observation error variances. Such a larger rejection of erroneous observations can explain a part of the improvement on the forecast scores seen in Fig. 5.

Satellite observations are among those for which the error variance has been proportionally the most reduced. This seems to be related to large positive impact observed at high levels (Fig. 6).

8. DFS in the ARPEGE ensemble

Another application of the estimation of the $\mathrm{Tr}(\boldsymbol{\Pi}_i^{\mathrm{T}}\boldsymbol{\Pi}_i\mathbf{H}\mathbf{K})$ is the evaluation of the DFS_i associated with the different subsets of observations and variables.

Figure 7 (top panel) shows the DFS by observation type. Not surprisingly, the largest DFS are observed for Radiosounding and Brightness temperature types, followed by the DFS of Aircraft observations. Cardinali et al. (2004) have obtained similar values for the DFS per observation type, but with a larger value for satellite observations. This is likely due to the fact that the European Centre for Medium-Range Weather Forecasts (ECMWF) analysis relies on a larger number of satellite data.

The comparison of the DFS with the corresponding numbers of observations shows that the DFS is not simply a question of number, since the weight of Radiosounding

observations is proportionally larger than the weight of Brightness temperature observations. This can be interpreted as a lower confidence accorded to the last type of measurements, but it also reflects the fact that they are more integrated measures. Relative to their numbers, the DFS of Satwind and GPSRO measurements are high, which can be explained by the fact that they are given a large confidence and also because they are present in data-sparse areas like over the oceans.

In the Northern Hemisphere, the largest DFS are clearly for Radiosounding data, where they are quite numerous, followed by Aircraft and Brightness temperature data (Fig. 7, middle panel).

Despite the overall lower DFS per observation of Brightness temperature observations in the ARPEGE assimilation, they appear to be the main source of information in the Southern Hemisphere (Fig. 7, bottom panel), since there are few Radiosounding observations in this area of the globe.

Figure 8 (top panel) shows the DFS of the different observed variables given by all types of observations, except the Brightness temperature type. It can be seen that the largest DFS is for the wind components given by the upper-air observations (Aircraft, Satwind, Radiosounding, and Profiler) followed by the temperature measurements given by the same type of observations, except Profiler. For these variables, the DFS is nearly directly proportional to the number of measurements. The large DFS per observation of GPSRO observations again has to be noted. On the contrary, the DFS per observation is small for surface observations (Synop, Driftbuoy, and Scatterometer), except for the geopotential Z measurements given by Synop.

The DFS is a measure of the sum of the variations of the analysis brought by the variations of a given set of observations. Those variations of the analysis are clearly directly linked to the confidence given to the observations, through their observation error variances. The variations of the analysis are also constrained by the length scale of the background error correlations. If the length scale is relatively large in comparison with the density of observations, then the individual enabled variation of the analysis at an observation location will be small. On the contrary, if the length scale of the background error correlations is shorter, as for specific humidity, then the impact of observations will be larger. This can explain why the DFS per observation is high for specific humidity observations (Fig. 8, top panel). Of course, this does not say that humidity observations will have a main contribution on the reduction of forecast errors. The DFS has simply to be interpreted as a measure of the degrees of freedom for the analysis (the signal) to deviate from the background.

The DFS associated with different Brightness temperature sounders (Fig. 8, middle panel) are particularly
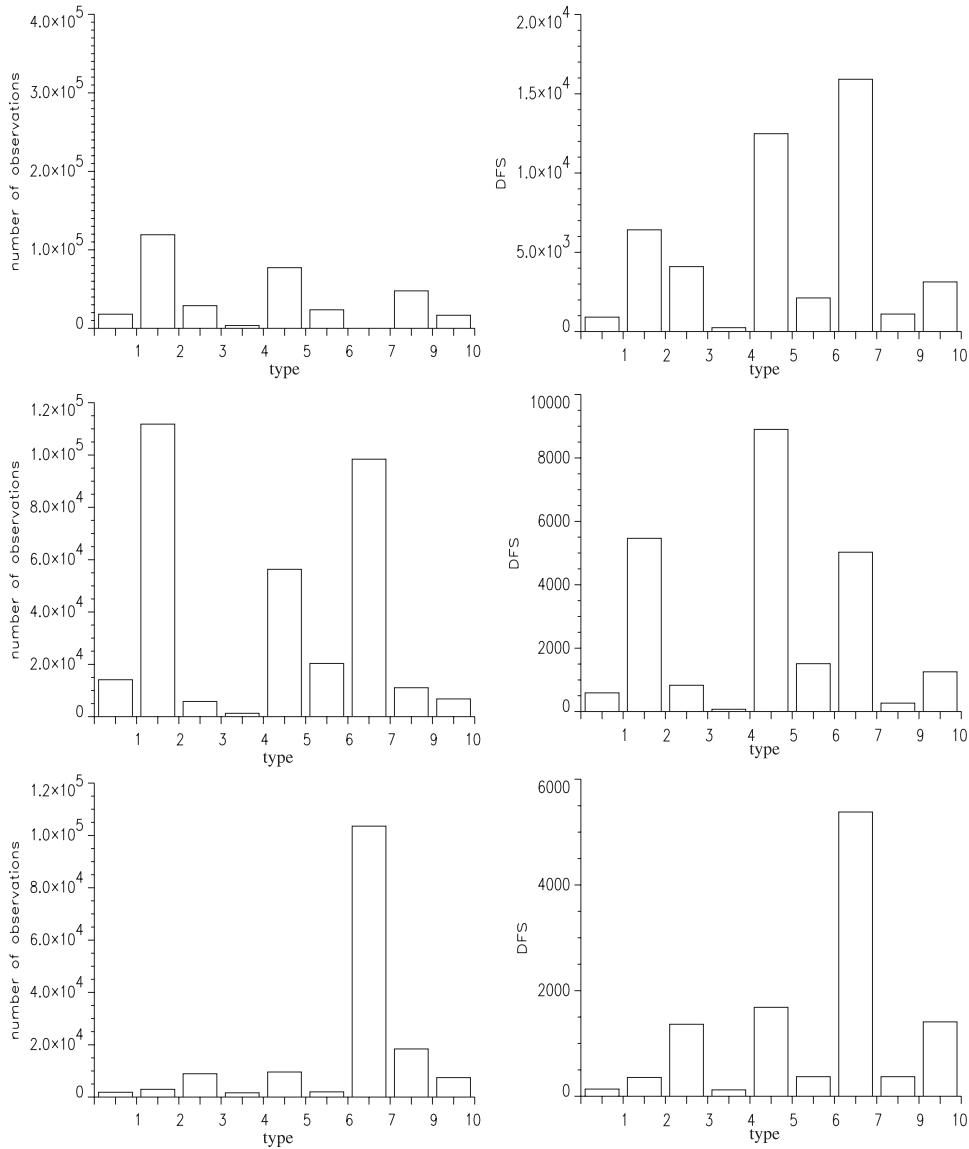
FIG. 7. Number and partial DFS of observations by observation type: 1 = Synop, 2 = Aircraft, 3 = Satwind, 4 = Driftbuoy, 5 = Radiosounding, 6 = Profiler, 7 = Brightness temperature, 9 = Scatterometer, and 10 = GPSRO. (top) Globe, (middle) 20°–90°N area, and (bottom) 20°–90°S area.

interesting to examine, since they show that the DFS are not again simply functions of the numbers of observations. Note, in particular, the large values of the DFS per observation for High Resolution Infrared Radiation Sounder (HIRS), Advanced Microwave Sounding Unit-B (AMSU-B), Meteorological Satellite (Meteosat) clear-sky radiances (CSR), and, on the contrary, the low DFS/p for IASI observations.

Figure 8 (bottom panel) shows that the DFS per observation associated with HIRS channels 4–7 and 15 are quite low, when compared with those associated with channels 11–12, although they have comparable observation error standard deviation. A likely explanation of this

behavior is that channels 4–7 and 15 are associated with much larger weighting functions than those of channels 11–12. As a consequence, the background error variances (not shown here), as diagnosed by a randomization procedure proposed by Andersson et al. (2000), show quite low values for channels 4–7 and 15 and larger values for channels 11–12. It means that the former channels will not produce a large deviation from the background. Another way to interpret this is that there is not a lot of information to get from a very integrated observation. A complementary explanation is that channels 11–12 are more sensitive to humidity than channels 4–7 and 15. Since background error correlations are shorter for humidity
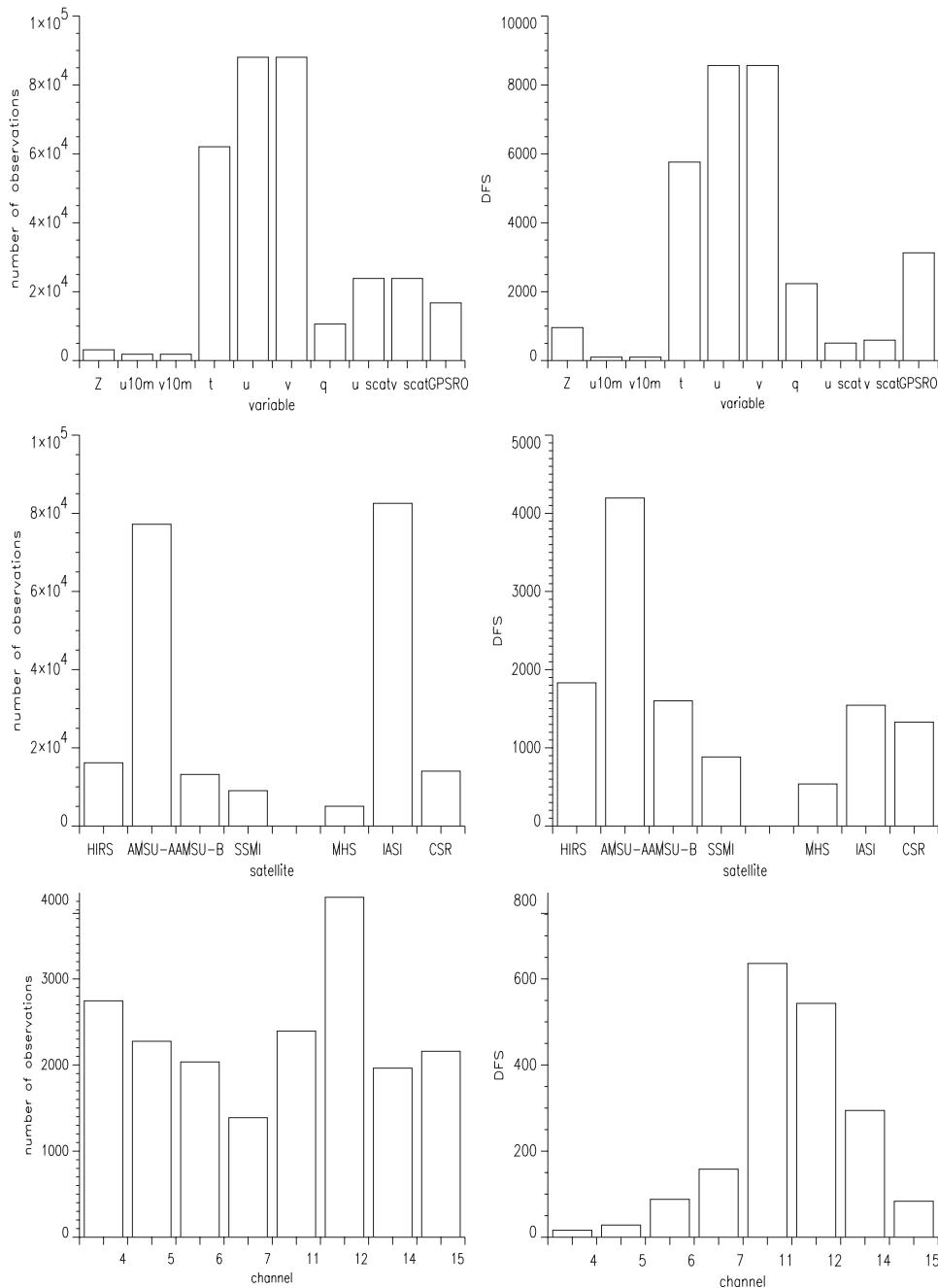
FIG. 8. (left) Number and (right) partial DFS: (top) by variable, except brightness temperatures; (middle) by sounders; and (bottom) by channels (HIRS channels).

than for temperature, this will also allow channels 11–12 to produce larger local modifications to the background and then to be characterized by larger DFS.

9. Conclusions

Real-size assimilation systems rely on approximations that can make them suboptimal with respect to the combination of background and observation information. A posteriori diagnostics give a way to check the optimality of such systems and to improve them.

The use of an ensemble of assimilations is a way to optimize the description of background error covariance matrix **B**. Ensembles of assimilations can be based on an explicit perturbation of observations in order to provide implicit background perturbations, representative

of background errors, from which useful statistics can be deduced. They are implemented, or planned to be implemented, in operational centers such as the ECMWF (Isaksen et al. 2007) or Météo-France (Berre et al. 2007).

It has been shown that such an ensemble can be used to evaluate the theoretical statistical estimations of subparts of the cost function measuring the distance between the analysis and the two sources of information, the background and the observations. The estimations appear to be robust in a real-size analysis system where the number of observations is large. Moreover, the robustness of the estimations can be increased by averaging them upon several dates.

The statistical estimations of subparts of the cost function can be used to tune the complete set of observation error standard deviations that appear to be overestimated in the ARPEGE 4D-Var assimilation. They can also allow the determination of a single normalization coefficient for the overall background error standard deviations; this error also happens to be (by chance) overestimated in the ARPEGE system.

The impact on the short-range forecasts of the reduction of both observation and background errors has been shown to be neutral to slightly positive at the low levels. The slight improvement on scores could be partly due to an improved rejection of erroneous observations against background, due to the better scaling of both observation and background error variances. A stronger positive impact has been observed in the stratosphere, where satellite radiance errors were overestimated to a larger extent than other error variances.

Since the ensemble of assimilations is based on perturbations of observations consistent with observation error variances, the right amplitude of those variances is also expected to have an influence on the correct dispersion of the ensemble. Although the ARPEGE analysis is based on the 4D-Var formalism, the ensemble of assimilations is run, for the time being, in a 3D-Var FGAT mode at Météo-France. 4D-Var ensembles have been occasionally run and actually show that the DFS given by the two types of ensembles are quite close. However, it is planned to more carefully investigate the differences between the background errors and DFS produced by 3D-Var FGAT and 4D-Var ensembles.

Desroziers et al. (2005a) have proposed other diagnostics, in observation space, to diagnose observation, background, and analysis errors. A first comparison, not shown here, between this method and the procedure described in the present paper shows a nice agreement between the two approaches. It is planned to systematically compare the error variances given by the two methods and to try to establish why the two methods

compare so well although they rely on apparently different theoretical rationale.

Up to now, only a single normalization coefficient of matrix **B** has been determined. There is a scope for investigating whether the a posteriori diagnostics described in this paper but also the observation space diagnostics proposed in Desroziers et al. (2005a) can be used to more finely optimize matrix **B**.

As shown in the paper, a direct by-product of an ensemble of perturbed analyses is the evaluation of the weights of observations through their DFS. Since the associated computations are costless, they can be implemented easily to provide an additional routine monitoring tool in a real-size system.

APPENDIX A

Proof of Eq. (5)

The incremental implementation of 4D-Var, with the linearization of observation operators around successive updated nonlinear trajectories, attempts to solve the original problem of minimizing the following cost function, with nonlinear observation operators, including model integration:

$$J(\mathbf{x}) = \left(\frac{1}{2}\right)\{(\mathbf{x}^b - \mathbf{x})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}^b - \mathbf{x})$$
$$+ [\mathbf{y}^o - H(\mathbf{x})]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{y}^o - H(\mathbf{x})]\}. \quad (\text{A1})$$

The analysis $\mathbf{x}^a$ is the vector that makes the gradient of J vanish. That can be written as

$$0 = \nabla J(\mathbf{x}^a), \quad (\text{A2})$$

$$= \mathbf{B}^{-1}(\mathbf{x}^b - \mathbf{x}^a) + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{y}^o - H(\mathbf{x}^a)], \quad (\text{A3})$$

$$= \mathbf{B}^{-1}[\mathbf{x}^t + \boldsymbol{\epsilon}^b - (\mathbf{x}^t + \boldsymbol{\epsilon}^a)]$$
$$+ \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}[H(\mathbf{x}^t) + \boldsymbol{\epsilon}^o - H(\mathbf{x}^t + \boldsymbol{\epsilon}^a)], \quad (\text{A4})$$

$$\simeq \mathbf{B}^{-1}(\boldsymbol{\epsilon}^b - \boldsymbol{\epsilon}^a) + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\boldsymbol{\epsilon}^o - \mathbf{H}\boldsymbol{\epsilon}^a), \quad (\text{A5})$$

where **H** is the linearized observation operator around $\mathbf{x}^a$. It follows that

$$\boldsymbol{\epsilon}^a = (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}(\mathbf{B}^{-1}\boldsymbol{\epsilon}^b + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\boldsymbol{\epsilon}^o) \quad (\text{A6})$$

$$= (\mathbf{I} - \mathbf{KH})\boldsymbol{\epsilon}^b + \mathbf{K}\boldsymbol{\epsilon}^o, \quad (\text{A7})$$

with

$$\mathbf{K} = (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1} \qquad (A8)$$

$$= \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}. \qquad (A9)$$

APPENDIX B

Proof of Eq. (14)

Chapnik (2005) has given a proof of Eq. (14) with a linear observation operator. This proof is generalized here for the nonlinear case.

The difference between the analysis $\mathbf{x}^a$ and the extended observation vector $\mathbf{z}$ (merging the background $\mathbf{x}^b$ and the proper observation vector $\mathbf{y}^o$) can be written as

$$\Gamma(\mathbf{x}^a) - \mathbf{z} = \Gamma(\mathbf{x}^a) - [\Gamma(\mathbf{x}^t) + \boldsymbol{\epsilon}] \qquad (B1)$$

$$\simeq \Gamma(\mathbf{x}^a) - \Gamma(\mathbf{x}^a) + \Gamma\boldsymbol{\epsilon}^a - \boldsymbol{\epsilon} \quad (B2)$$

$$= \Gamma\boldsymbol{\epsilon}^a - \boldsymbol{\epsilon}, \qquad (B3)$$

where Γ is the linearized generalized observation operator around $\mathbf{x}^a$.

Following Eq. (5), the analysis error $\boldsymbol{\epsilon}^a$ can be written as

$$\boldsymbol{\epsilon}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\boldsymbol{\epsilon}^b + \mathbf{K}\boldsymbol{\epsilon}^o \qquad (B4)$$

$$= (\Gamma^{\mathrm{T}}\mathbf{S}^{-1}\Gamma)^{-1}\Gamma^{\mathrm{T}}\mathbf{S}^{-1}\boldsymbol{\epsilon} \qquad (B5)$$

$$= \Gamma^{+}\boldsymbol{\epsilon}, \qquad (B6)$$

where Γ^{+} is the left inverse of Γ such as $\Gamma^{+}\Gamma = \mathbf{I}$.

Using Eqs. (B3) and (B6) and the linearity property of the expectation operator, it follows that

$$E\{J_i(\mathbf{x}^a)\} = [\Gamma_i(\mathbf{x}^a) - \boldsymbol{\epsilon}_i]^{\mathrm{T}}\mathbf{S}_i^{-1}[\Gamma_i(\mathbf{x}^a) - \boldsymbol{\epsilon}_i] \qquad (B7)$$

$$= \{[\mathbf{\Pi}_i(\Gamma(\mathbf{x}^a) - \boldsymbol{\epsilon})]^{\mathrm{T}}\mathbf{S}_i^{-1}[\mathbf{\Pi}_i(\Gamma(\mathbf{x}^a) - \boldsymbol{\epsilon})]\} \qquad (B8)$$

$$= \mathrm{Tr}\{\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{\Pi}_i(\Gamma\Gamma^{+} - \mathbf{I})E\{\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}}\}(\Gamma\Gamma^{+} - \mathbf{I})^{\mathrm{T}}\} \qquad (B9)$$

$$= \mathrm{Tr}[\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{\Pi}_i(\Gamma\Gamma^{+} - \mathbf{I})\mathbf{S}(\Gamma\Gamma^{+} - \mathbf{I})^{\mathrm{T}}]. \qquad (B10)$$

It is easy to check that $(\Gamma\Gamma^{+} - \mathbf{I})\mathbf{S}(\Gamma\Gamma^{+} - \mathbf{I})^{\mathrm{T}} = \mathbf{S} - \Gamma\mathbf{P}^a\Gamma^{\mathrm{T}}$ and thus

$$E\{J_i(\mathbf{x}^a)\} = \mathrm{Tr}[\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{\Pi}_i(\mathbf{S} - \Gamma\mathbf{P}^a\Gamma^{\mathrm{T}})] \quad (B11)$$

$$= p_i - \mathrm{Tr}(\Gamma_i\mathbf{P}^a\Gamma_i^{\mathrm{T}}\mathbf{S}_i^{-1}). \qquad (B12)$$

APPENDIX C

Proof of Eqs. (23) and (24)

Using Eq. (14), the expectation of a subpart of $J^o(\mathbf{x}^a)$ can be written as

$$E\{J_i^o(\mathbf{x}^a)\} = p_i - \mathrm{Tr}(\mathbf{H}_i\mathbf{P}^a\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}) \qquad (C1)$$

$$= p_i - \mathrm{Tr}(\mathbf{\Pi}_i\mathbf{H}\mathbf{P}^a\mathbf{H}^{\mathrm{T}}\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{R}_i^{-1}). \quad (C2)$$

If the observation errors associated with subset i are independent from the other observation errors, then $\mathbf{R}_i^{-1} = \mathbf{\Pi}_i\mathbf{R}^{-1}\mathbf{\Pi}_i^{\mathrm{T}}$ and $\mathbf{R}^{-1}$ commutes with $\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{\Pi}_i$ so that

$$E\{J_i^o(\mathbf{x}^a)\} = p_i - \mathrm{Tr}(\mathbf{\Pi}_i\mathbf{H}\mathbf{P}^a\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{\Pi}_i\mathbf{\Pi}_i^{\mathrm{T}}) \qquad (C3)$$

$$= \mathrm{Tr}(\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{\Pi}_i) - \mathrm{Tr}(\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{\Pi}_i\mathbf{H}\mathbf{P}^a\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}) \qquad (C4)$$

$$= \mathrm{Tr}[\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{\Pi}_i(\mathbf{I} - \mathbf{H}\mathbf{K})]. \qquad (C5)$$

On another hand, the trace of a matrix $\mathbf{A}$ can be approximated with

$$\mathrm{Tr}(\mathbf{A}) = \mathrm{Tr}E\{\boldsymbol{\eta}\mathbf{A}\boldsymbol{\eta}^{\mathrm{T}}\}, \qquad (C6)$$

$$\simeq \frac{1}{L}\sum_l \boldsymbol{\eta}_l^{\mathrm{T}}\mathbf{A}\boldsymbol{\eta}_l, \qquad (C7)$$

where the $\boldsymbol{\eta}_l$ are a set of L vectors of random numbers with standard Gaussian distribution (mean 0 and variance 1; Girard 1987).

Equation (C5) can be rewritten as

$$E\{J_i^o(\mathbf{x}^a)\} = \mathrm{Tr}[\mathbf{R}^{1/2}\mathbf{R}^{-1}\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{\Pi}_i(\mathbf{I} - \mathbf{H}\mathbf{K})\mathbf{R}^{1/2}] \quad (C8)$$

$$= \mathrm{Tr}[\mathbf{R}^{1/2}\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{\Pi}_i(\mathbf{I} - \mathbf{H}\mathbf{K})\mathbf{R}^{1/2}]. \quad (C9)$$

Applying the randomized estimation of the trace of a matrix to this last equation gives

$$E\{J_i^o(\mathbf{x}^a)\} = \frac{1}{L}\sum_l \boldsymbol{\eta}_l^{\mathrm{T}}\mathbf{R}^{1/2}\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{\Pi}_i(\mathbf{I} - \mathbf{H}\mathbf{K})\mathbf{R}^{1/2}\boldsymbol{\eta}_l \qquad (C10)$$

$$= \frac{1}{L}\sum_l \boldsymbol{\delta}_l^{o\mathrm{T}}\mathbf{\Pi}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{\Pi}_i(\mathbf{I} - \mathbf{H}\mathbf{K})\boldsymbol{\delta}_l^o \qquad (C11)$$

$$= \frac{1}{L}\sum_l \boldsymbol{\delta}_{l,i}^{o\mathrm{T}}\mathbf{R}_i^{-1}(\boldsymbol{\delta}_{l,i}^o - \mathbf{H}_i\mathbf{K}\boldsymbol{\delta}_l^o), \qquad (C12)$$

where $\boldsymbol{\delta}_l^o = \mathbf{R}^{1/2}\boldsymbol{\eta}_l$ is a vector of perturbations on the complete set of observations and $\boldsymbol{\delta}_{l,i}^o$ is the restriction of $\boldsymbol{\delta}_l^o$ to the subset of observations i. In an ensemble of

assimilations, the perturbations on the background, the observations, and the analysis are linked by the relation $\boldsymbol{\delta}^a = (\mathbf{I} - \mathbf{KH})\boldsymbol{\delta}^b + \mathbf{K}\boldsymbol{\delta}^o$. Furthermore, the perturbations $\boldsymbol{\delta}^o$ and $\boldsymbol{\delta}^b$ are uncorrelated, so that $1/L \sum_l \boldsymbol{\delta}_{l,i}^{o\mathrm{T}} \mathbf{R}_i^{-1} \mathbf{H}_i \boldsymbol{\delta}_l^a \simeq 1/L \sum_l \boldsymbol{\delta}_{l,i}^{o\mathrm{T}} \mathbf{R}_i^{-1} \mathbf{H}_i \mathbf{K} \boldsymbol{\delta}_l^o$, which proves Eq. (23).

The proof of Eq. (24) can be obtained with the same kind of derivations.

REFERENCES

Andersson, E., M. Fisher, R. Munro, and A. McNally, 2000: Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence. *Quart. J. Roy. Meteor. Soc.,* **126,** 1455–1472.

Belo Pereira, M., and L. Berre, 2006: The use of an ensemble approach to study the background error covariances in a global NWP model. *Mon. Wea. Rev.,* **134,** 2466–2489.

Bennett, A., L. Leslie, C. Hagelberg, and P. Powers, 1993: Tropical cyclone prediction using a barotropic model initialized by a generalized inverse method. *Mon. Wea. Rev.,* **121,** 1714–1729.

Berre, L., S. Stefanescu, and M. Belo Pereira, 2006: The representation of analysis effect in three error simulation techniques. *Tellus,* **58A,** 196–209.

——, O. Pannekoucke, G. Desroziers, S. E. Stefanescu, B. Chapnik, and L. Raynaud, 2007: A variational assimilation ensemble and the spatial filtering of its error covariances: Increase of sample size by local spatial averaging. *Proc. ECMWF Workshop on Recent Developments in Data Assimilation for Atmosphere and Ocean,* Reading, United Kingdom, ECMWF, 151–168.

Bormann, N., S. Saarinen, G. Kelly, and J.-N. Thépaut, 2003: The spatial structure of observation errors in atmospheric motion vector from geostationary satellite data. *Mon. Wea. Rev.,* **131,** 706–718.

Buehner, M., P. Gauthier, and Z. Liu, 2005: Evaluation of new estimates of background- and observation-error covariances for variational assimilation. *Quart. J. Roy. Meteor. Soc.,* **131,** 3373–3383.

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

Chapnik, B., 2005: Réglage des statistiques d'erreur en assimilation variationnelle (Tuning of error statistics in variational assimilation). Thèse de doctorat de l'Université Toulouse III, 196 pp.

——, G. Desroziers, F. Rabier, and O. Talagrand, 2004: Properties and first application of an error statistics tuning method in variational assimilation. *Quart. J. Roy. Meteor. Soc.,* **130,** 2253–2275.

——, ——, ——, and ——, 2006: Diagnosis and tuning of observational error statistics in a quasi operational data assimilation setting. *Quart. J. Roy. Meteor. Soc.,* **132,** 543–565.

Courtier, P., 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.

Daescu, D., 2008: On the sensitivity equations of four-dimensional variational (4DVAR) data assimilation. *Mon. Wea. Rev.,* **136,** 3050–3065.

Dee, D. P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.,* **123,** 1128–1145.

——, and A. 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 F. Bouttier, 1999: A reformulation of the background error covariances in the ECMWF global data assimilation system. *Tellus,* **51A,** 195–221.

Desroziers, G., and S. Ivanov, 2001: Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. *Quart. J. Roy. Meteor. Soc.,* **127,** 1433–1452.

——, 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.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99** (C5), 3104–3127.

Fisher, M., 2003a: Background error covariance modelling. *Proc. ECMWF Workshop on Recent Developments in Data Assimilation for Atmosphere and Ocean,* Reading, United Kingdom, ECMWF, 97–123.

——, 2003b: Estimation of entropy reduction and degrees of freedom for signal for large variational analysis systems. Tech. Rep. 397, ECMWF, 20 pp.

——, and E. Andersson, 2001: Developments in 4D-Var and Kalman filtering. Tech. Rep. 347, ECMWF, 38 pp.

Girard, D., 1987: A fast Monte Carlo cross-validation procedure for large least squares problems with noisy data. Tech. Rep. 687-M, IMAG, Grenoble, France, 22 pp.

Hollingsworth, A., and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus,* **38A,** 111–136.

Houtekamer, P., L. Lefaivre, J. Derome, H. Ritchie, and H. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.,* **124,** 1225–1242.

Hutchinson, M., 1989: A stochastic estimator for the trace of the influence matrix for Laplacian smoothing splines. *Commun. Stat. Simul.,* **18,** 1059–1076.

Isaksen, L., M. Fisher, and J. Berner, 2007: Use of analysis ensembles in estimating flow-dependent background error variance. *Proc. ECMWF Workshop on Flow-Dependent Aspects of Data Assimilation,* Reading, United Kingdom, ECMWF, 65–86.

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

Lewis, J. M., and J. C. Derber, 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus,* **37A,** 309–322.

Munro, R., C. Köpken, G. Kelly, J.-N. Thépaut, and R. Saunders, 2004: Assimilation of Meteosat radiance data within the 4D-Var system at ECMWF: Data quality monitoring, bias correction and single-cycle experiments. *Quart. J. Roy. Meteor. Soc.,* **130,** 2293–2313.

Pannekoucke, O., L. Berre, and G. Desroziers, 2007: Filtering properties of wavelets for the local background error correlations. *Quart. J. Roy. Meteor. Soc.,* **133,** 363–379.

Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.,* **126,** 1143–1170.

——, N. Fourri, D. Chafa, and P. Prunet, 2002: Channel selection method for infrared atmospheric sounding interferometer radiances. *Quart. J. Roy. Meteor. Soc.,* **128,** 1011–1032.

Raynaud, L., L. Berre, and G. Desroziers, 2008: Spatial averaging of ensemble-based background-error variances. *Quart. J. Roy. Meteor. Soc.,* **134,** 1003–1014.

Rodgers, C., 2000: *Inverse Methods for Atmospheric Sounding Theory and Practice*. World Scientific Publishing, 256 pp.

Talagrand, O., 1997: Assimilation of observations, an introduction. *J. Meteor. Soc. Japan,* **75,** 191–209.

——, 1999: A posteriori verification of analysis and assimilation algorithms. *Proc. ECMWF Workshop on Diagnosis of Data Assimilation Systems,* Reading, United Kingdom, ECMWF, 17–28.

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

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

Wahba, G., D. Johnson, F. Gao, and J. Gong, 1995: Adaptive tuning of numerical weather prediction models: Randomized GCV in three- and four-dimensional data assimilation. *Mon. Wea. Rev.,* **123,** 3358–3369.

Zagar, N., E. Andersson, and M. Fisher, 2005: Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors. *Quart. J. Roy. Meteor. Soc.,* **131,** 987–1011.

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.