Q. J. R. Meteorol. Soc. (2001), **127**, pp. 1775–1794

Estimation of the representativeness error caused by the incremental formulation of variational data assimilation

By GERALD DESROZIERS[1]*, OMAR BRACHEMI[1,2] and BACHIR HAMADACHE[1,2]

[1]*Centre National de Recherches Météorologiques, France*
[2]*Office National de la Météorologie, Algeria*

(Received 23 May 2000; revised 18 December 2000)

SUMMARY

A formalization of the representativeness error caused by the use of a simplified increment in certain implementations of the variational formalism is introduced. The incremental representativeness-error can be determined by summing up the part of the background-error power-spectrum that is beyond the truncation of the simplified increment. Thus, this error will be more important for fields, such as humidity, whose forecast errors have large components at small scales. The length scale and the oscillating structure of this incremental representativeness-error are explained as the consequences of the application of a square window in spectral space to the corresponding background-error covariances. A computation of the covariances of this error has been performed in a realistic framework given by the French numerical weather prediction system. In particular, it shows that this error is not negligible, and that its size is of the same order as that of observational instrument error which occurs when using a T63 spectral resolution of the analysis increment.

KEYWORDS: Representativeness error Variational assimilation

1. INTRODUCTION

Since the first applications of the variational formalism in data assimilation with simplified models (Lewis and Derber 1985; Le Dimet and Talagrand 1986; Courtier and Talagrand 1987), the potential of this approach for numerical weather prediction (NWP) has been amply demonstrated. In particular, variational methods allow the use of a wider range of observations that are only indirectly linked with model variables: as an example, the benefit of using a variational scheme for assimilating satellite radiances has been established by a number of authors (Eyre *et al.* 1993; Andersson *et al.* 1994; Derber and Wu 1998). This gain is already possible in a three-dimensional variational scheme (3D-Var), that also permits one to avoid the data selection necessary in optimal interpolation schemes. The use of four-dimensional variational assimilation (4D-Var) adds the advantage of producing an analysis that is dynamically consistent with the prediction model. The implicit use of more realistic flow-dependent structure-functions in 4D-Var and the relation between this algorithm and the extended Kalman filter have also been shown by Thépaut *et al.* (1996).

Variational assimilation schemes are now being progressively implemented in large operational schemes (Parrish and Derber 1992; Rabier *et al.* 1998; Thépaut *et al.* (1998) in either their 3D-Var or 4D-Var formulation. This implementation has been made possible by reducing the cost of such methods: a pre-conditioning of the problem (Lorenc 1988), based on a change of the control variables, is commonly used that allows one to limit the number of iterations, but it is especially the use of the so-called incremental formulation that has made the application of variational methods affordable.

The principle of the incremental formulation of 3D/4D-Var algorithms (Courtier *et al.* 1994) is to reduce the dimension of the minimization space and to simplify the assimilating model through, for example, the use of a simplified physical package (Janisková *et al.* 1999). Various implementations of the incremental formulation have shown that it provides an excellent approximation to the original full 4D-Var formulation

* Corresponding author: Météo-France, CNRM/GMME/RECYF, 42 av. G. Coriolis, 31057 Toulouse Cedex, France. email: gerald.desroziers@meteo.fr

and is now in operational use at ECMWF (Rabier *et al.* 2000). A theoretical justification of this approach has also been investigated by Laroche and Gauthier (1998).

In order to obtain a still larger reduction of the 4D-Var computational cost, a multiple-truncation incremental approach was proposed by Veersé and Thépaut (1998) and is implemented in the operational 4D-Var scheme used by Météo-France. In such a formulation, which is a generalization of the incremental approach and which also has connections with multigrid methods, the resolution of the analysis increment is progressively increased, starting from an increment whose resolution can be much coarser than that of the prediction model. In this case, with such a large reduction of the resolution of the control variable, it appears that some scales will obviously not be described by the tangent linear model used in the assimilation procedure and that this will lead to a particular kind of representativeness error. The aim of this paper is to formalize this error (section 2), to illustrate it in a simple one dimensional (1-D) case (section 3) and also to evaluate it in a realistic framework (section 4).

2. Incremental representativeness-error

(*a*) *Incremental formulation*

The principle of the incremental formulation of 3D/4D-Var algorithms (Courtier *et al.* 1994) is to seek the increment $\delta\mathbf{x}$ to add to the background $\mathbf{x}^b$—so that the analysis is given by $\mathbf{x}^a = \mathbf{x}^b + \delta\mathbf{x}$—as the vector minimizing the cost function

$$J(\delta\mathbf{x}) = \tfrac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \tfrac{1}{2}(\mathbf{d} - \mathbf{H}\delta\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{d} - \mathbf{H}\delta\mathbf{x}), \quad (1)$$

with $\mathbf{d} = \mathbf{y}^o - H(\mathbf{x}^b)$ the vector of departures between observations $\mathbf{y}^o$ and the background $\mathbf{x}^b$; $\mathbf{H}$ is the linearized version of the observation operator H, that allows the computation of the model equivalents at observation locations. In Eq. (1) the superscript T denotes matrix transpose.

The background term measures the distance between the analysis $\mathbf{x}^a$ and the short-range forecast $\mathbf{x}^b$, with $\mathbf{B}$ the forecast-error covariance matrix. In the observation term, $\mathbf{R}$ denotes the observation-error covariance matrix, including the representativeness error (Lorenc 1986).

This formalism rather corresponds to the 3D-Var case. The extension of the temporal dimension (4D-Var) can, however, be introduced. The observation operator H then includes the integration of the forecast model, $\mathbf{x}^a$ stands for the initial conditions at the beginning of the assimilation period and an outer loop in the minimization process is introduced in order to take into account some parts of the nonlinearities contained in H. Although the following developments stand in a 4D-Var framework, this paper focuses on the 3D-Var case for the sake of simplicity.

The practical interest of the incremental formulation is to allow a further simplification by reducing the resolution of the analysis increment $\delta\mathbf{x}$. This is formalized by introducing a simplification operator S (possibly nonlinear) and performing the change of variable $\delta\mathbf{w} = \mathbf{S}\delta\mathbf{x}$ where $\mathbf{S}$ is the linearized version of S. Then the previous cost-function J can be rewritten as a function of $\delta\mathbf{w}$:

$$J(\delta\mathbf{w}) = \tfrac{1}{2}\delta\mathbf{w}^T(\mathbf{B}^w)^{-1}\delta\mathbf{w} + \tfrac{1}{2}(\mathbf{d} - \mathbf{G}\delta\mathbf{w})^T(\mathbf{R}^S)^{-1}(\mathbf{d} - \mathbf{G}\delta\mathbf{w}), \quad (2)$$

with $\mathbf{G}$ an approximation of $\mathbf{HS}^{-I}$, where $\mathbf{S}^{-I}$ is the generalized inverse of $\mathbf{S}$ such that $\mathbf{x}^a = \mathbf{x}^b + \mathbf{S}^{-I}\delta\mathbf{w}$ and $\mathbf{B}^w$ is the appropriate forecast-error covariance matrix in simplified space. Matrix $\mathbf{R}^S$ is the new observation-error covariance matrix, associated with the simplified space, that should be different from $\mathbf{R}$.

(*b*) *Observation-error expression*

Following Lorenc (1986), Daley (1993) and Ide *et al.* (1997), observation errors can be split in two parts in the first formulation (1) of J (with no simplification by **S**); if $\mathbf{x}^t$ stands for the unknown true state and $\delta\mathbf{x}^t = \mathbf{x}^t - \mathbf{x}^b$ the difference between the truth and the background, it follows that

$$\begin{aligned}\mathbf{d} - \mathbf{H}\delta\mathbf{x}^t &= \mathbf{y}^o - \mathbf{H}(\mathbf{x}^b) - \mathbf{H}\delta\mathbf{x}^t \\ &\approx \mathbf{y}^o - \mathrm{H}(\mathbf{x}^t) \\ &\approx \{\mathbf{y}^o - \mathbf{y}^t\} + \{\mathbf{y}^t - \mathrm{H}(\mathbf{x}^t)\} \\ &\approx \boldsymbol{\epsilon}^o + \boldsymbol{\epsilon}^H.\end{aligned}$$

The term $\boldsymbol{\epsilon}^o$ denotes the observational instrument error which depends on the engineering specifications only and is henceforward called merely 'observation error'. The second term $\boldsymbol{\epsilon}^H$ corresponds to the forward interpolation error directly related to the form of H. This last error depends on the type of relation between $\mathbf{y}^o$ and $\mathbf{x}^t$, but also on the resolution of the analysis grid.

Introducing the simplification by **S**, and $\delta\mathbf{w}^t$ such as $\delta\mathbf{w}^t = \mathbf{S}\delta\mathbf{x}^t$, the observation error takes a new form,

$$\begin{aligned}\mathbf{d} - \mathbf{G}\delta\mathbf{w}^t &= \mathbf{y}^o - \mathbf{H}(\mathbf{x}^b) - \mathbf{G}\delta\mathbf{w}^t \\ &\approx \{\mathbf{y}^o - \mathbf{y}^t\} + \{\mathbf{y}^t - \mathrm{H}(\mathbf{x}^b) - \mathbf{H}\delta\mathbf{x}^t\} + \{\mathbf{H}\delta\mathbf{x}^t - \mathbf{G}\delta\mathbf{w}^t\} \\ &\approx \{\mathbf{y}^o - \mathbf{y}^t\} + \{\mathbf{y}^t - \mathrm{H}(\mathbf{x}^t)\} + \mathbf{H}\delta\mathbf{x}^t - \mathbf{G}\delta\mathbf{w}^t \\ &\approx \boldsymbol{\epsilon}^o + \boldsymbol{\epsilon}^H + \boldsymbol{\epsilon}^{o,S}.\end{aligned}$$

Thus, the introduction of the simplification operator S has a consequence in terms of observation error: it adds a new error $\boldsymbol{\epsilon}^{o,S}$ that could be defined as 'incremental representativeness-error', caused by the lack of resolution in the simplified increment. This error can be rewritten

$$\begin{aligned}\boldsymbol{\epsilon}^{o,S} &= \mathbf{H}\delta\mathbf{x}^t - \mathbf{G}\delta\mathbf{w}^t \\ &= \mathbf{H}\delta\mathbf{x}^t - \mathbf{H}\mathbf{S}^{-I}\mathbf{S}\delta\mathbf{x}^t \\ &= \mathbf{H}\boldsymbol{\mathcal{R}}\delta\mathbf{x}^t \\ &= \mathbf{H}\boldsymbol{\epsilon}^S\end{aligned}$$

with $\boldsymbol{\mathcal{R}} = \mathbf{I} - \mathbf{S}^{-I}\mathbf{S}$ (where **I** is the identity matrix) and $\boldsymbol{\epsilon}^S = \boldsymbol{\mathcal{R}}\delta\mathbf{x}^t$ is the incremental representativeness-error in model space. It means that $\boldsymbol{\epsilon}^S$ is the difference between a given increment $\delta\mathbf{x}^t$ at the full model-resolution and this same increment projected on the simplified space by **S** and then brought back to the full resolution space by $\mathbf{S}^{-I}$. Since $\mathbf{S}^{-I}$ is not the true inverse of **S**, there is an obvious loss of the smallest structures in this operation.

(*c*) *Observation-error covariance*

The covariance of the incremental representativeness-error in observation space, denoted $\mathbf{F}^S$, is

$$\begin{aligned}\mathbf{F}^S &= \mathrm{E}(\boldsymbol{\epsilon}^{o,S}(\boldsymbol{\epsilon}^{o,S})^T) \\ &= \mathbf{H}\boldsymbol{\mathcal{R}}\mathrm{E}(\delta\mathbf{x}^t(\delta\mathbf{x}^t)^T)\boldsymbol{\mathcal{R}}^T\mathbf{H}^T \\ &= \mathbf{H}\boldsymbol{\mathcal{R}}\mathbf{B}\boldsymbol{\mathcal{R}}^T\mathbf{H}^T.\end{aligned}$$

If the cross-correlations between $\boldsymbol{\epsilon}^{o}$, $\boldsymbol{\epsilon}^{H}$ and $\boldsymbol{\epsilon}^{o,S}$ are neglected, then the observation-error covariance is written $\mathbf{R}^{S} = \mathbf{E} + \mathbf{F} + \mathbf{F}^{S}$.

This expression is what is needed for the analysis algorithm with the incremental formulation. At this stage, the difference between the two kinds of representativeness error must be pointed out.

The original representativeness-error covariance $\mathbf{F}$ contains all the energy in the atmospheric scales that are not resolved by the model. On the other hand, the above expression of $\mathbf{F}^{S}$ shows that the incremental-error covariances are related to the forecast errors that are not described by the increment. Since the level of the forecast error is normally lower than the energy of the corresponding atmospheric field, then $\mathbf{F}^{S}$ will not be as large as the increase in F that would occur if the resolution of the model was reduced to that of the increment.

The present study was motivated by suspicious results obtained with a dense network of dropsondes (Desroziers *et al.* 1997) and the 3D-Var assimilation scheme of the Action de Recherche Petite Echelle Grande Echelle (ARPEGE) model. Spurious increments in the analysis of the humidity field were observed around the edges of the dense network. Furthermore, the amplitude of these increments increased as the resolution of the increments became coarser. The ARPEGE prediction model is a spectral global model (Courtier *et al.* 1991) based on a stretched geometry (Schmidt 1977; Courtier and Geleyn 1988), that allows a very high resolution locally. In that case, the classic representativeness-error covariance $\mathbf{F}$ should be very low in the high-resolution area of interest and non-negligible in the low-resolution area at the antipodes of that region. On the other hand, the ARPEGE 3D/4D-Var scheme is based on the use of a non-stretched increment with a uniform resolution over the globe: thus, the incremental representativeness covariance $\mathbf{F}^{S}$ could be large in the high-resolution area (see section 4 for a quantitative evaluation of this error). Then, in this particular case, the covariance of $\boldsymbol{\epsilon}^{H} + \boldsymbol{\epsilon}^{o,S}$ is more homogeneous than either term separately.

(*d*) *Innovation covariance and optimal solution*

The use of the incremental formulation also has a consequence on the way the matrix $\mathbf{B}^{w}$ has to be specified in the J^{b} term of the cost function. Since this matrix should reflect the covariance of the simplified forecast error $\delta\mathbf{w}^{t} = \mathbf{S}\delta\mathbf{x}^{t}$, $\mathbf{B}^{w}$ is given by $\mathbf{B}^{w} = \mathbf{S}\mathrm{E}\{\delta\mathbf{x}^{t}(\delta\mathbf{x}^{t})^{T}\}\mathbf{S}^{T} = \mathbf{S}\mathbf{B}\mathbf{S}^{T}$ (this is what is done, for example, in the operational implementation of 4D-Var at Météo-France). This specification of $\mathbf{B}^{w}$ corresponds to a decrease of the variances of the forecast errors and an increase of the length-scales of the associated structure-functions, as a result of the filtering of the small-scale part of these errors. (See section 3 for an illustration of this in a simple case.)

Since the innovation vector $\mathbf{d} = \mathbf{y}^{o} - \mathrm{H}(\mathbf{x}^{b})$ is unchanged in the incremental formulation, its error covariance remains equal to $\mathbf{H}\mathbf{B}\mathbf{H}^{T} + \mathbf{E} + \mathbf{F}$. If $\mathbf{D}$ denotes this covariance of the innovation vector, using the classical expression for the solution of the analysis problem, gives $\delta\mathbf{x}^{a} = \mathbf{B}\mathbf{H}^{T}(\mathbf{H}\mathbf{B}\mathbf{H}^{T} + \mathbf{E} + \mathbf{F})^{-1}\mathbf{d} = \mathbf{B}\mathbf{H}^{T}\mathbf{D}^{-1}\mathbf{d}$, in the non-incremental formulation and $\delta\mathbf{x}^{a} = \mathbf{S}^{-1}\mathbf{B}^{w}\mathbf{G}^{T}\mathbf{D}^{-1}\mathbf{d}$, in the incremental formulation. The vector $\delta\mathbf{x}^{o} = \mathbf{D}^{-1}\mathbf{d}$ has the same dimension as the vector of observations $\mathbf{y}^{o}$ and is known to be the solution of the dual formulation of the analysis problem in observation space (Courtier 1997). Thus, it will be unchanged with a low-resolution increment since $\mathbf{D}$ is not modified. The change in the final analysis will only come from the term $\mathbf{S}^{-1}\mathbf{B}^{w}\mathbf{G}^{T}$ instead of $\mathbf{B}\mathbf{H}^{T}$ that will map the modification to the background coming from each observation differently. Omission of $\mathbf{F}^{S}$ from the observation term results

in the observations being given too much weight. The impact of this is discussed in subsection 3(b).

3. ILLUSTRATION WITH A SIMPLE 1-D CASE

(a) *Documentation of the incremental representativeness-error*

The incremental representativeness-error is first illustrated in a simple toy problem given by an analysis on a circle. A spectral Fourier decomposition of the signal, assumed here to be a temperature increment (K), is used. The analysis formulation is as in (2) with $\delta\mathbf{w}$ the vector of corresponding spectral coefficients in the simplified space.

To define the background error covariance matrix $\mathbf{B}$, a Gaussian shape structure function is first specified in physical space. The representation of $\mathbf{B}$ in spectral space is given by $\mathbf{B} = \mathcal{T}\widetilde{\mathbf{B}}\mathcal{T}^{\mathrm{T}}$, where $\widetilde{\mathbf{B}}$ is the forecast-error covariance-matrix in physical space and $\mathcal{T}$ is the Fourier transform. Assuming a homogeneity hypothesis on the domain, it can be shown that $\mathbf{B}$ is diagonal and that its diagonal is obtained by applying the Fourier transform to the Gaussian correlation in physical space. Note that some of the following relations, derived in this simple case, are valid only under this homogeneity hypothesis introduced here, for the sake of simplicity, to give an insight into the structure of the representativeness-error covariance.

A background-error vector $\boldsymbol{\epsilon}^{\mathrm{b}}$ is simulated with the following procedure: if $\boldsymbol{\eta}^{\mathrm{b}}$ is a vector of random numbers with Gaussian distribution, then writing $\boldsymbol{\epsilon}^{\mathrm{b}} = \mathbf{B}^{1/2}\boldsymbol{\eta}^{\mathrm{b}}$ is a way to simulate background errors consistent with the matrix $\mathbf{B}$ since $\mathrm{E}\{\boldsymbol{\eta}^{\mathrm{b}}(\boldsymbol{\eta}^{\mathrm{b}})^{\mathrm{T}}\} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix, and then $\mathrm{E}(\boldsymbol{\epsilon}^{\mathrm{b}}(\boldsymbol{\epsilon}^{\mathrm{b}})^{\mathrm{T}}) = \mathbf{B}^{1/2}\mathrm{E}\{\boldsymbol{\eta}^{\mathrm{b}}(\boldsymbol{\eta}^{\mathrm{b}})^{\mathrm{T}}\}\mathbf{B}^{1/2} = \mathbf{B}$. It will be further assumed that $\mathbf{x}^{\mathrm{b}} = 0$, so that $\mathbf{x}^{\mathrm{t}} = -\boldsymbol{\epsilon}^{\mathrm{b}}$.

Figure 1 shows the signal obtained with such a procedure (solid line), using a 100-km length-scale for the structure function. Here the standard deviation of the background error is assumed to be uniform on the domain and equal to 1 K. The length of the domain is arbitrarily set to 40,000 km and a high $K = 400$ spectral truncation is chosen to specify this reference signal. The simplification operator S introduced in the previous section is defined as the operator projecting the signal onto its K^{S} first spectral components only (S is thus linear in this simple case). The generalized inverse $\mathbf{S}^{-\mathrm{I}}$ is also defined by completing with zero padding a vector of spectral coefficients of the simplified space. The composition of operators $\mathcal{F} = \mathbf{S}^{-\mathrm{I}}\mathbf{S}$ then acts like a low-pass filter in spectral space. The difference between the original signal $\boldsymbol{\epsilon}^{\mathrm{b}}$ and the filtered signal $\mathcal{F}\boldsymbol{\epsilon}^{\mathrm{b}}$ is the error of representativeness (dashed line in Fig. 1) resulting from the loss of resolution caused by the projection of $\boldsymbol{\epsilon}^{\mathrm{b}}$ on the simplified space (this error is also given by $\mathcal{R}\boldsymbol{\epsilon}^{\mathrm{b}}$, with the definition of $\mathcal{R}$ introduced in the previous section).

Panels (a) and (b) of Fig. 2 show the spectrum of the spectral variances (the diagonal of $\mathbf{B}$) associated with two structure functions with different length-scales (100 km and 200 km respectively) in physical space. The covariance of the incremental representativeness-error is written $\mathcal{R}\mathbf{B}\mathcal{R}^{\mathrm{T}}$ with $\mathcal{R} = \mathbf{I} - \mathbf{S}^{-\mathrm{I}}\mathbf{S} = \mathbf{I} - \mathcal{F}$. This means that the spectrum of the spectral variances for the incremental representativeness-error is simply the smaller-scale part of the background-error spectrum beyond the truncation K^{S} of the simplified space. The total variance of the representativeness error in physical space is, in particular, the sum of the spectral variances between K^{S} and K and the covariance of this error is obtained by applying the spectral inverse transform to this part of the spectrum (in this simple case where the original forecast errors are assumed to be homogeneous).

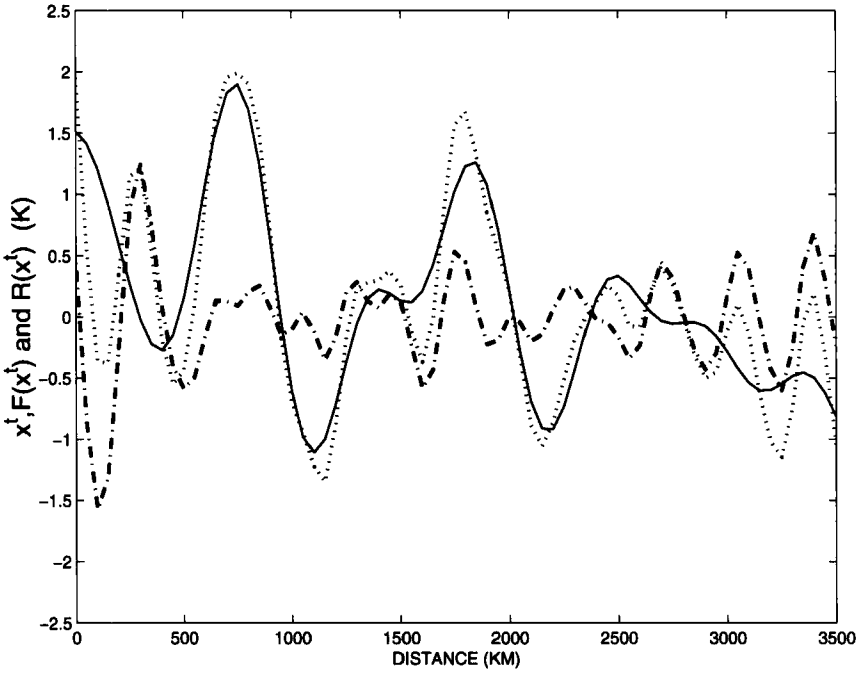

Figure 1. Simulated signal resulting from the specification of a Gaussian structure-function in physical space and a 100 km length-scale (solid line); the dotted line shows the filtered signal with $K^S = 79$ and the dash-dotted curve the resulting incremental representativeness-error.

If $\mathbf{b}$, $\mathbf{b}^{\mathcal{F}}$ and $\mathbf{b}^{\mathcal{R}}$ respectively stand for the diagonals of matrices $\mathbf{B}$, $\mathbf{B}^{\mathcal{F}} = \mathcal{F}\mathbf{B}\mathcal{F}^{\mathrm{T}}$, $\mathbf{B}^{\mathcal{R}} = \mathcal{R}\mathbf{B}\mathcal{R}^{\mathrm{T}}$, then the corresponding structure-functions for the background error ρ, the filtered background-error $\mathcal{F}\epsilon^{\mathrm{b}} - \rho^{\mathcal{F}}$ and the incremental representativeness-error $\rho^{\mathcal{R}}$, are obtained with $\rho = \mathcal{T}^{\mathrm{i}}(\mathbf{b})$, $\rho^{\mathcal{F}} = \mathcal{T}^{\mathrm{i}}(\mathbf{b}^{\mathcal{F}})$ and $\rho^{\mathcal{R}} = \mathcal{T}^{\mathrm{i}}(\mathbf{b}^{\mathcal{R}})$, where $\mathcal{T}^{\mathrm{i}}$ is the inverse spectral transform.

Panels c, d, e and f in Fig. 2 show the corresponding covariances in physical space for two length-scales of the original structure function (100 km for the left-hand panels and 200 km for the right-hand panels) and two resolutions of the increment ($K^S = 95$ and $K^S = 63$ respectively, corresponding to a 200 km and 300 km length for the shortest half-wave described). It can be seen that the forecast error correlations appear broader in the simplified space (dotted lines) than in the full resolution space (solid lines). This is especially true when the original length-scale becomes significantly shorter than the resolution of the increment (Figs. 2(c), (e)). In this case, the variance of the incremental representativeness error also becomes large (dashed lines). Of course, this simply says that, if the cost of computation is not to be too high, the increments need to have a resolution which resolves the background-error covariance properly.

The filtered spectrum $\mathbf{b}^{\mathcal{F}}$ can also be written $\mathbf{b}^{\mathcal{F}} = f\mathbf{b}$, with f a square window in spectral space, zeroing the values of the spectrum beyond truncation K^S. Then, the convolution theorem leads to the relation $\rho^{\mathcal{F}} = \mathcal{T}^{\mathrm{i}}(\mathbf{b}^{\mathcal{F}}) = \mathcal{T}^{\mathrm{i}}(f\mathbf{b}) = \mathcal{T}^{\mathrm{i}}(f) * \rho$, where $*$ is the convolution operator and $\mathcal{T}^{\mathrm{i}}(f)$ is the image in physical space of the square window in spectral space f. This last function is known to turn on and off rapidly and thus $\rho^{\mathcal{F}}$ will also show oscillations.

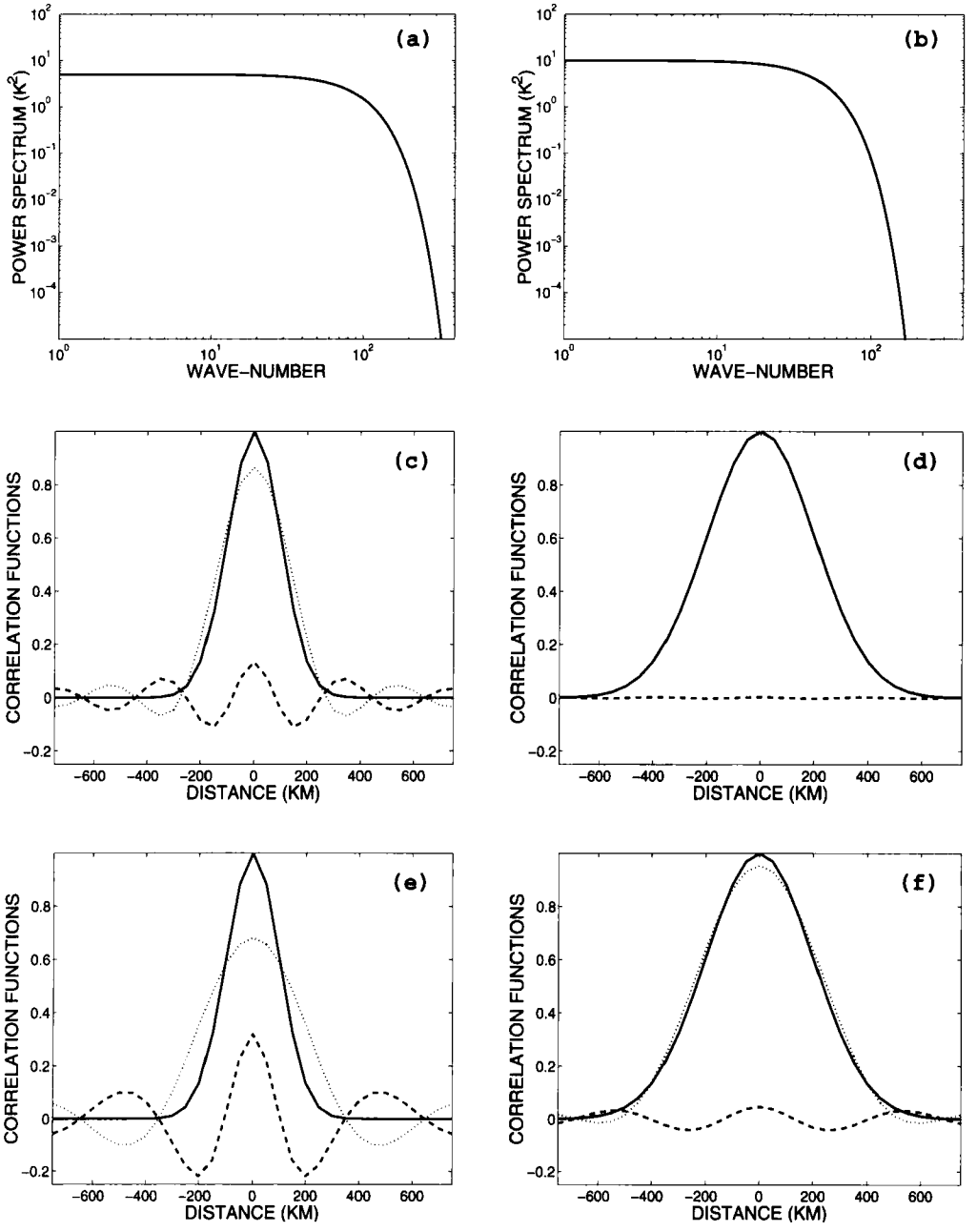

Figure 2. (a, b) Power spectra resulting from the specification of a Gaussian structure-function in physical space with length-scales of (a) 100 km and (b) 200 km. (c, d, e, f) Structure function in full-resolution physical space (solid line) and in simplified space (dotted line), and incremental representativeness-error covariance (pecked line): (c) 100 km length-scale, $K^S = 95$; (d) 200 km length-scale, $K^S = 95$; (e) 100 km length-scale, $K^S = 63$; (f) 200 km length-scale, $K^S = 63$.

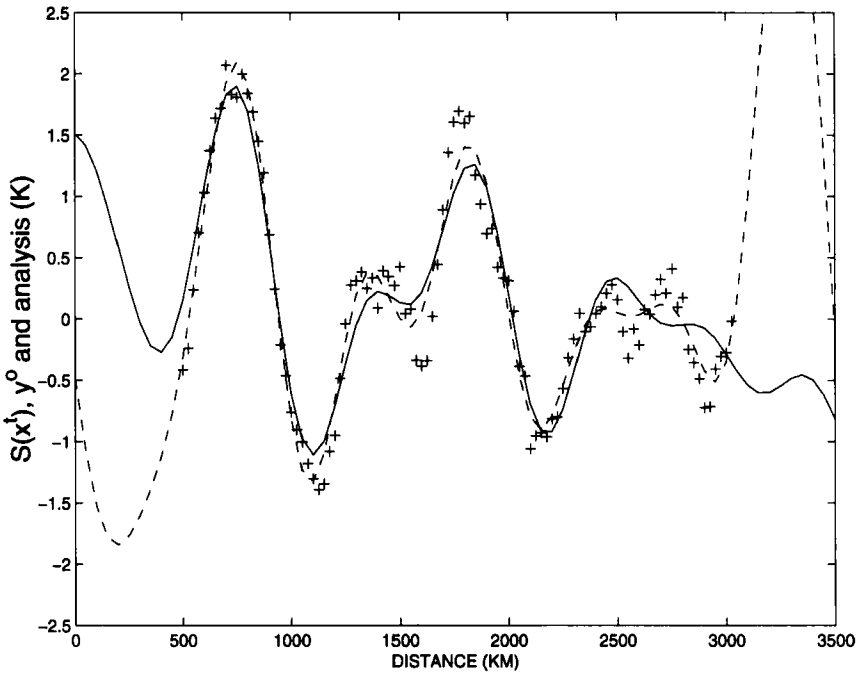

Figure 3. True signal $\mathbf{S}(\mathbf{x}^t)$, corresponding to a 100 km length-scale with a uniform standard deviation σ^b after projection on simplified space (solid line), associated observations with $\mathcal{R} = (\sigma^o)^2\mathbf{I}$ and $\sigma^o = \sigma^b/10$ (crosses), and resulting sub-optimal analysis with no introduction of the incremental representativeness-error (pecked line).

This explains the oscillating shape of the representativeness-error structure-function $\rho^{\mathcal{R}}$ and that the length-scale of the representativeness error correlation is directly related to the smallest scale that can be described in the simplified space (the length-scale of $\rho^{\mathcal{R}}$ increases when K^S decreases).

(*b*) *Impact of the representativeness error in the analysis*

The impact of neglecting the incremental representativeness-error is shown here in the same simple 1-D analysis scheme. It is again assumed that $\mathbf{x}^b = 0$ and then that $\mathbf{x}^t = -\boldsymbol{\epsilon}^b$. A realization of $\boldsymbol{\epsilon}^b$ corresponding to a 100 km length-scale is produced with the same procedure as above with a high $K = 400$ spectral truncation. Precise observations $\mathbf{y}^o$ are also simulated with $\mathbf{y}^o = \mathbf{H}(\mathbf{x}^t) + \mathbf{R}\boldsymbol{\eta}^o$, where $\boldsymbol{\eta}^0$ is a vector of random numbers with normal distribution, $\mathbf{R} = (\sigma^o)^2\mathbf{I}$ and σ^o is ten times less than σ^b. In order to evaluate the performance of the analysis scheme, the true signal $\mathbf{x}^t$ is first projected on the simplified space (this projection $\mathbf{S}(\mathbf{x}^t)$ is represented by the solid line in Fig. 3). (Here $K^S = 79$ corresponds to a roughly 250-km length for the shortest half-wave described (and thus much broader than the length-scale of the forecast errors)). Dense observations (one observation every 50 km) are set in a part of the domain only (crosses in Fig. 3).

The resulting analysis with no introduction of the incremental representativeness-error is shown in Fig. 3 (dashed line): locally, it is far from the filtered truth (solid line) and especially presents unrealistic structures on both sides of the observation network. (These spurious increments are similar to those observed with the ARPEGE 3D-Var and dense observations as mentioned above.) On the other hand, the analysis with the correct

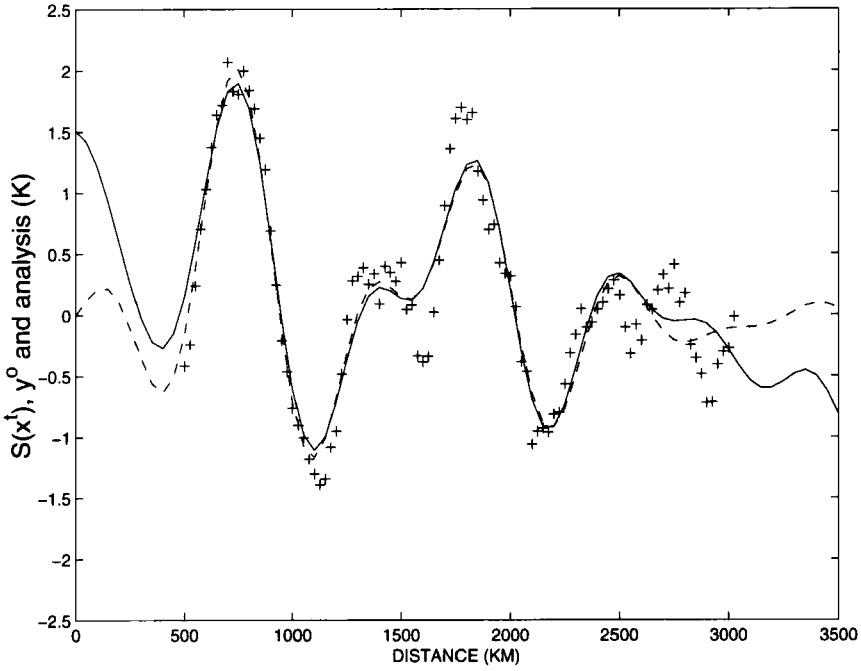

Figure 4. True signal $\mathbf{S}(\mathbf{x}^t)$, corresponding to a 100 km length-scale with a uniform standard deviation σ^b after projection on simplified space (solid line), associated observations with $\mathcal{R} = (\sigma^o)^2\mathbf{I}$ and $\sigma^o = \sigma^b/10$ (crosses), and resulting optimal analysis taking into account the incremental representativeness-error (pecked line).

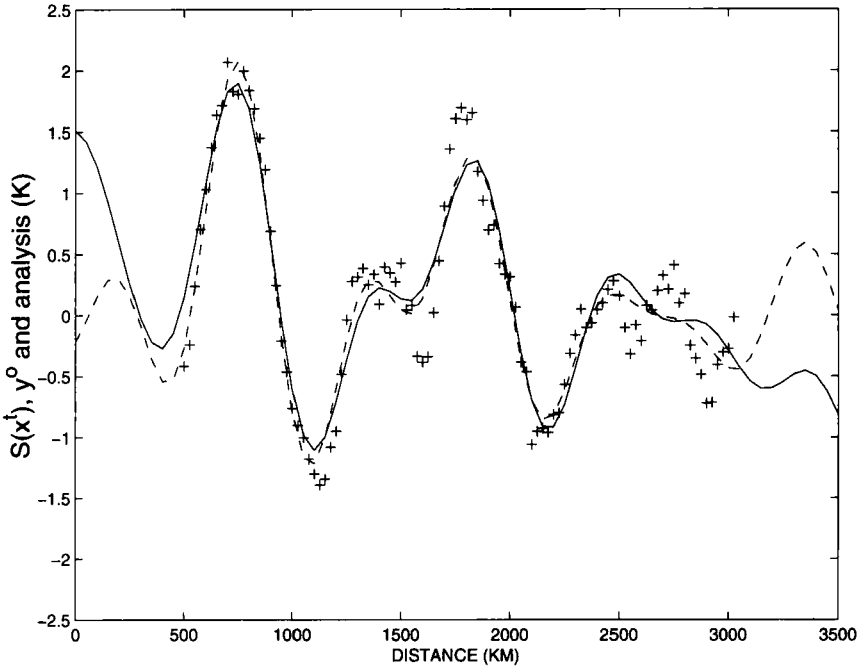

Figure 5. True signal $\mathbf{S}(\mathbf{x}^t)$, corresponding to a 100 km length-scale with a uniform standard deviation σ^b after projection on simplified space (solid line), associated observations with $\mathcal{R} = (\sigma^o)^2\mathbf{I}$ and $\sigma^o = \sigma^b/10$ (crosses), and resulting analysis taking into account the correct variance of the representativeness error (pecked line), but assuming it to be diagonal in observation space.

specification of the representativeness error (Fig. 4) shows a better fit to $S(\mathbf{x}^t)$ and a clear damping of the oscillations on the edges of the observation network.

Finally, the analysis resulting from the correct specification of the variance of the representativeness error but assuming it to be diagonal in observation space, appears to be between both previous ones (Fig. 5). In particular, it allows for a damping of the spurious increments on the edges of the observation network. This could be a simple but rather efficient way to take the incremental representativeness-error into account in practical implementations that do not usually allow for the specification of correlations between observation errors.

4. Evaluation in a realistic framework

(*a*) *Introduction*

The aim of this section is to present a description of the incremental representativeness-error in a realistic framework, given by the French ARPEGE NWP system. This operational system currently uses a 4D-Var (previously 3D-Var) analysis scheme based on an incremental formulation.

In the ARPEGE 3D/4D-Var formulation, the simplification operator S corresponds to both a reduction of the resolution of the analysis increment and a change of geometry (the full-resolution model has a stretched geometry whereas the resolution of the analysis increment is uniform over the whole globe).

(*b*) *Computation of forecast-error statistics*

The spectral background-error covariances are determined using the method proposed by Parrish and Derber (1992), where statistics are computed for the differences between forecasts at different ranges (here respectively 24 h and 48 h) but valid at the same time, assuming that these differences represent forecast errors. Here, this method has been applied to a winter period (three months from 1 December 1996 to 28 February 1997). At that time, the model had a T149 spectral resolution on the stretched grid and a stretching factor 3.5 corresponding to a T149 $\times$ 3.5 $\approx$ T520 resolution on the true geographical sphere near the pole of interest centred on France.

Since the resolution of the ARPEGE model varies over the geographical sphere, a spatial window was applied in order to obtain forecast-error covariances valid for the high-resolution area (Desroziers *et al.* (1995) had already applied such a method to determine the geographical dependency of the ARPEGE model forecast-errors). Since the ARPEGE stretching varies only slowly in the high-resolution area, the resolution of the model in this region is assumed to be homogeneous and close to T520.

If Υ is a field of background errors over the sphere, then $\Upsilon(\lambda, \mu)$ is written

$$\Upsilon(\lambda, \mu) = \sum_{n} \sum_{m=-n}^{n} \epsilon_n^m Y_n^m(\lambda, \mu),$$

where λ is the longitude, μ the sine of the latitude φ, ϵ_n^m the spectral coefficients associated with the total wave-number n and the zonal wave-number m. Here Y_n^m is a spherical harmonic defined by $Y_n^m(\lambda, \mu) = P_n^m(\mu)\, e^{im\lambda}$, where $P_n^m(\mu)$ is the Legendre polynomial with degree n and order m.

Although this is certainly a crude approximation, especially in the high-resolution area, the spectral covariances are assumed to be homogeneous. The two spherical harmonics Y_n^m and $Y_{n'}^{m'}$ are then correlated only if they are equal, and the modal variance

b_n is independent of the zonal wave-number m. The total variance P_n for a given wave-number n is given by

$$P_n = \sum_{m=-n}^{n} E(\epsilon_n^{m^2}) = (2n+1)b_n.$$

The procedure used to get local error statistics valid for the high-resolution area can be summarized as follows. First, a spatial window is applied, on the ARPEGE transformed sphere, to the physical space representation of the fields:

$$\begin{cases} \Upsilon^w(\lambda, \varphi) = \Upsilon(\lambda, \varphi)\left(1 - \left(\dfrac{\pi/2 - \varphi}{\pi/2}\right)^2\right), & 0 \leqslant \varphi \leqslant \pi/2 \\ \Upsilon^w(\lambda, \varphi) = 0, & -\pi/2 \leqslant \varphi < 0. \end{cases}$$

Since the northern hemisphere of the transformed sphere corresponds to the high-resolution area on the true geographical sphere, this parabolic window gradually changes the fields to zero outside the high-resolution area. The fields Υ^w are then transformed to the spectral space corresponding to the ARPEGE computation sphere, and to the spectral space corresponding to the geographical space, by a spectral transform implemented in the ARPEGE system. The tests performed by Desroziers *et al.* (1995) showed that this procedure allows one to recover a good description of the forecast-error covariances in the high-resolution area.

These statistics have been computed over the three-month period for each model variable and level (viz. temperature, specific humidity, vorticity and divergence at all 27 model levels and at surface pressure). As an example, the resulting power-spectra for temperature are shown in Fig. 6(a). They present characteristic shapes with, in particular, increasing energy in the small scales at the lower levels.

(*c*) *Computation of incremental representativeness-error statistics*

The same statistics have also been computed for the incremental representativeness-errors defined by $\epsilon^{\mathcal{R}} = \epsilon - \mathbf{S}^{-\mathrm{I}}\mathbf{S}\epsilon$, as in previous section. Because there is no linearized version $\mathbf{S}$ of the simplification operator S, the term $\mathbf{S}^{-\mathrm{I}}\mathbf{S}\epsilon$ has been computed as $\mathrm{S}^{-\mathrm{I}}\mathrm{S}(\mathbf{x}^{24}) - \mathrm{S}^{-\mathrm{I}}\mathrm{S}(\mathbf{x}^{48})$, where $\mathbf{x}^{24}$ and $\mathbf{x}^{48}$ respectively correspond to ARPEGE forecasts at 24 h and 48 h ranges and valid at the same time.

Figure 6(b) shows the representativeness-error power-spectra for temperature. There is a clear jump in the spectra at the increment resolution (here T95). Some energy remains below this resolution, since S is not a perfect filter in spectral space. However, it appears that the parts of these representativeness-error power-spectra located between the resolution of the simplified increment (T95) and the maximal resolution (T520) are very similar to their corresponding parts for the background-error power-spectra (Fig. 6(a)). The same result is found for specific humidity, vorticity, divergence and surface pressure (not shown). This is in agreement with the idea developed in the previous section that the incremental representativeness-error variances can be deduced simply from the background-error spectra.

The structure function ρ of the background error can be obtained by applying the inverse spectral transform. Because of the hypothesis of homogeneity and isotropy, ρ can be projected on zonal spherical harmonics P_n^0 only (Courtier *et al.* 1998):

$$\rho(r) = \sum_n b_n \sqrt{(2n+1)} P_n^0(\mu),$$

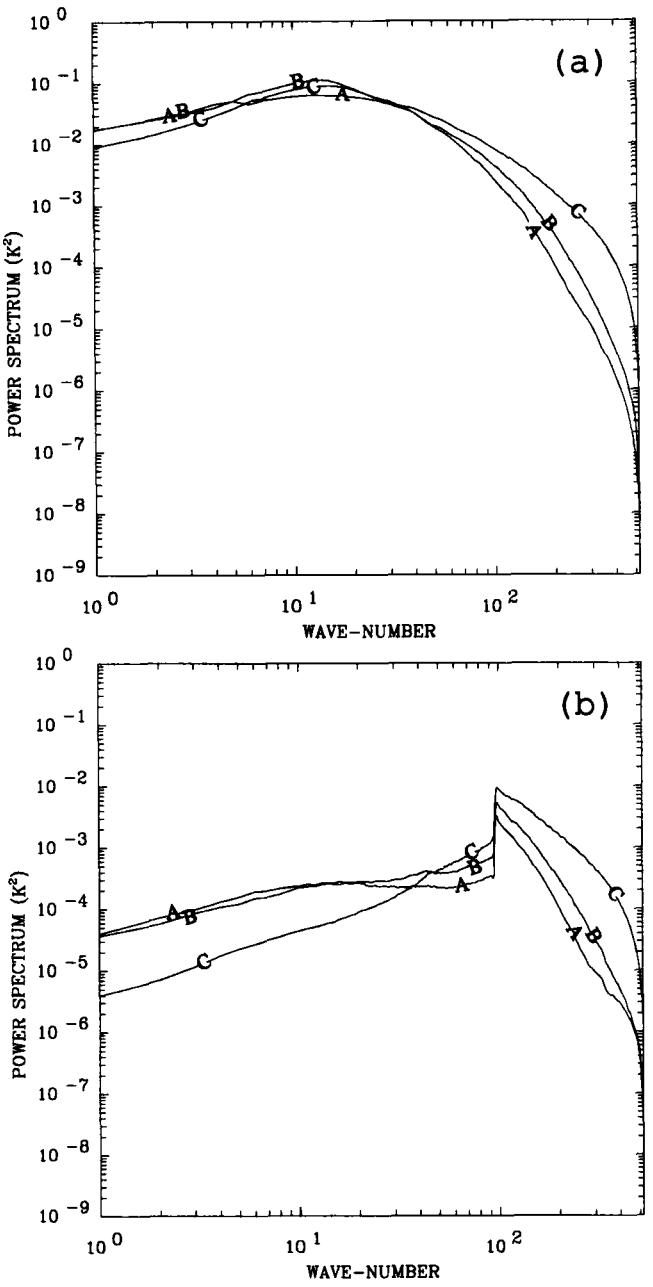

Figure 6. ARPEGE spectra for temperature for three model-levels, 8 (curve A ≈ 300 hPa), 12 (curve B ≈ 500 hPa) and 20 (curve C ≈ 850 hPa): (a) background error; (b) incremental representativeness-error.

with $r = a\varphi/2\pi$, a is the earth radius and $\mu = \sin(\varphi)$.

The structure function $\rho^{\mathcal{R}}$ of the incremental representativeness-error can be determined in the same way.

The temperature background-error correlations drop to zero monotonically at a distance around 1500 km, but with length-scales increasing with altitude (Fig. 7(a)). On the other hand, the corresponding correlations for the incremental error (Fig. 7(b)) show oscillations and shorter length-scales in good agreement with the 1-D case.

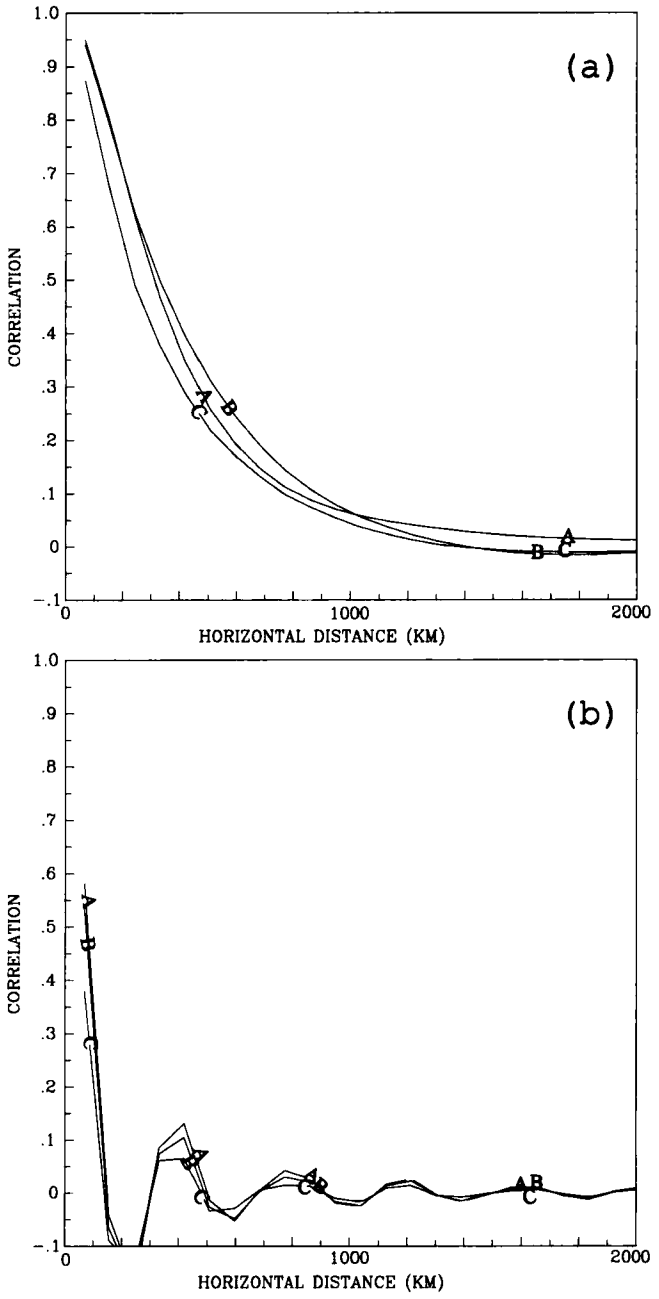

Figure 7. Structure functions for temperature corresponding with the spectra shown in Fig. 6 for three model-levels, 8 (curve A ≈ 300 hPa), 12 (curve B ≈ 500 hPa) and 20 (curve C ≈ 850 hPa): (a) background error; (b) incremental representativeness-error.

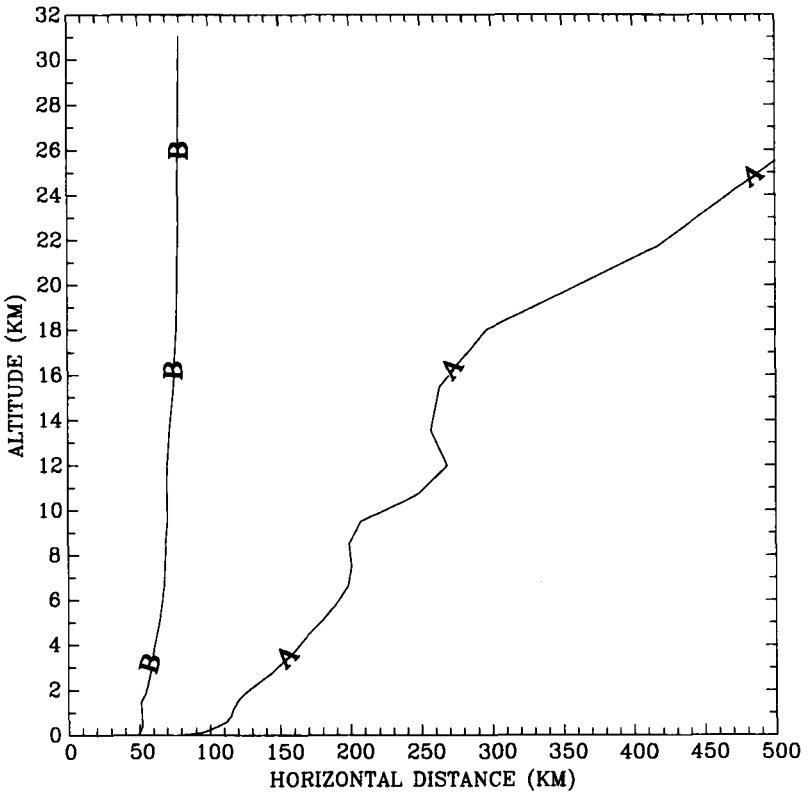

Figure 8. Length scales of the correlation functions for temperature: curve A, background errors; curve B, incremental representativeness-errors.

The length-scales $\mathcal{L}$ at different levels for both errors can be computed explicitly using an expression used by Daley (1991):

$$\mathcal{L}^2 = 2\frac{\sum_n b_n}{\sum_n (n(n+1)b_n)/a^2}.$$

Figure 8 shows that the horizontal length-scales of the background-error correlations for temperature increase with altitude over horizontal distances of 100 km to 500 km. These values are less than those obtained by some other authors, such as Rabier *et al.* (1998). This may be because the ARPEGE model has a very high resolution in this area and because Rabier *et al.* (1998) computed the error statistics at T106 resolution. As we showed in subsection 3(a), this may have increased the derived length-scales artificially. On the other hand, Fig. 8 shows that the length-scales of the incremental error are rather constant and much shorter (around 50–80 km), which is also consistent with the 1-D case.

The homogeneity hypothesis can be extended to the covariance between two distinct model levels. Then the vertical covariance $b_n^{l,l'}$ between two levels l and l', for the total horizontal wave-number n is written $b_n^{l,l'} = \sum_{m=-n}^{n} E(\epsilon_{l_n}^m (\epsilon_{l',n}^m)^*)$, where $(\epsilon_{l',n}^m)^*$ is the complex conjugate of $\epsilon_{l',n}^m$. Such a computation allows the specification of fully non-separable correlation structure-functions with a different vertical correlation for each wave-number (Courtier *et al.* 1998; Rabier *et al.* 1998). The total vertical correlation

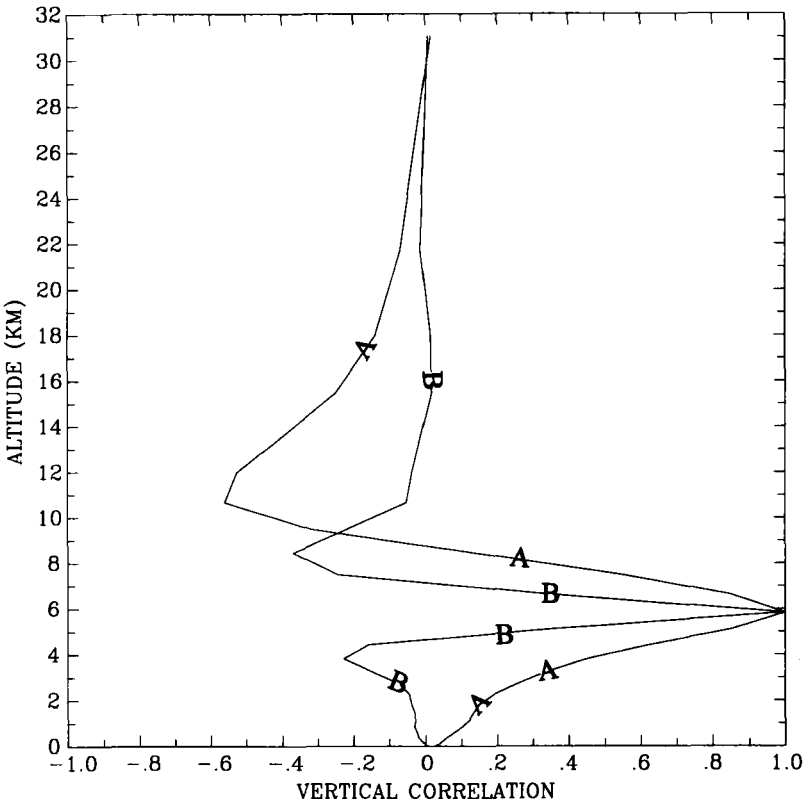

Figure 9. Vertical correlation for temperature between model level 12($\approx$ 6 km $\approx$ 500 hPa) and levels above and below it: curve A, background error; curve B, incremental representativeness-error.

$V^{l,l'}$ is given by $V^{l,l'} = \sum_n b_n^{l,l'}/(\sigma^l \sigma^{l'})$, where σ^l and $\sigma^{l'}$ are the standard deviations of the error at level l and l' respectively. Figure 9 shows that the simplification made in the horizontal on the vertical correlation of the incremental representativeness-error has an impact. Indeed, this vertical correlation becomes very similar to the one obtained by summing up the individual vertical correlations for each wave number n between K^S (here T95) and the full model resolution (here T520). Since the length-scales of the vertical correlations of the forecast errors are shorter at highest wave-numbers, then the length-scales of the total vertical correlations of the incremental representativeness-error (Fig. 9, curve B) will be shorter than the length-scales of the vertical correlations of the forecast errors (Fig. 9, curve A).

Finally, since the covariances of the incremental error can be deduced from the covariances of the background error, the variance of this error has been determined for the model variables and for different truncations K^S. For each level these variances $(\sigma^S)^2$ are simply computed with $(\sigma^S)^2 = \sum_{n=K^S}^{K} P_n$, where $K = 520$ is the full resolution of the ARPEGE model.

The background error for temperature (Fig. 10(a)) shows decreasing values with altitude: the profiles for representativeness error resemble the background-error profile, but with diminishing amplitudes with the augmentation of K^S. The ratio between representativeness error and background error also decreases with altitude because background-error spectra have more energy in the large scales at the highest levels. Note

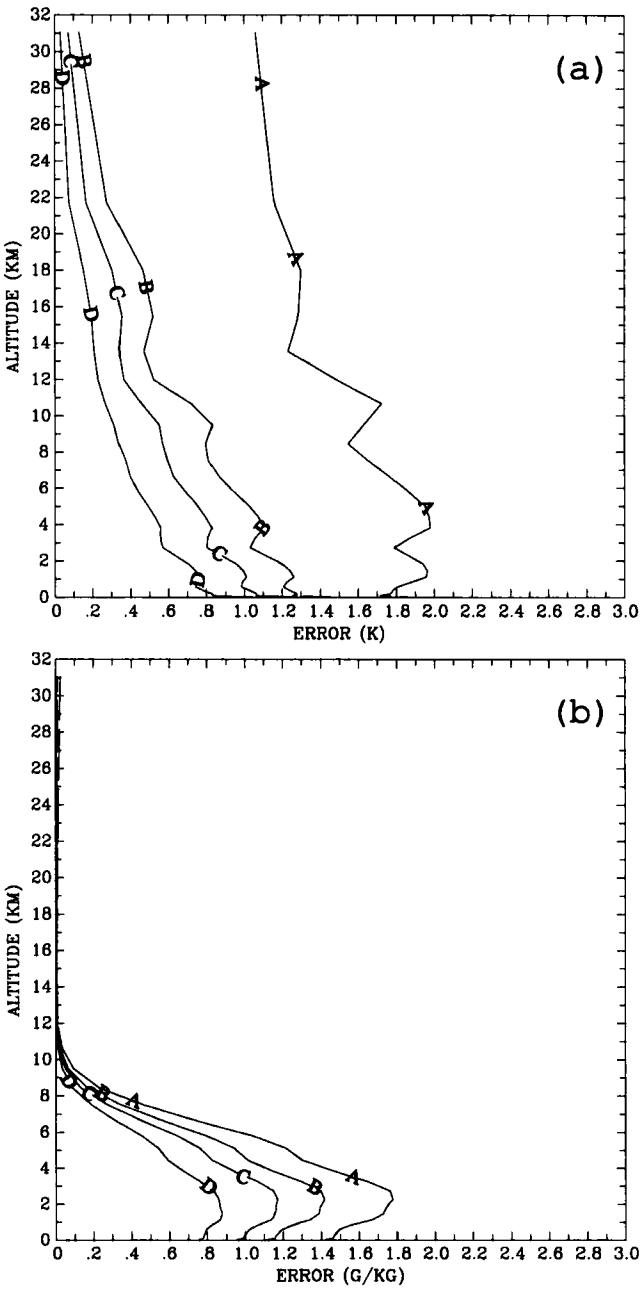

Figure 10. Variation with height of standard deviations of errors: (a) temperature; (b) specific humidity. In each panel, curve A shows the background errors and curves B, C and D, the (progressively smaller) associated representativeness-errors for resolutions of T42, T63 and T95 respectively.

that for $K^S = 42$ the mean value of this representativeness error is close to the typical value of observation error given to radiosounding measurements (about 1 K).

Figure 10(b) shows background and incremental error standard deviations for specific humidity: all curves show a maximum around 700 hPa and a ratio between representativeness error and background error larger than for temperature, since humidity-error spectra have more energy in the small scales than temperature-error spectra. The

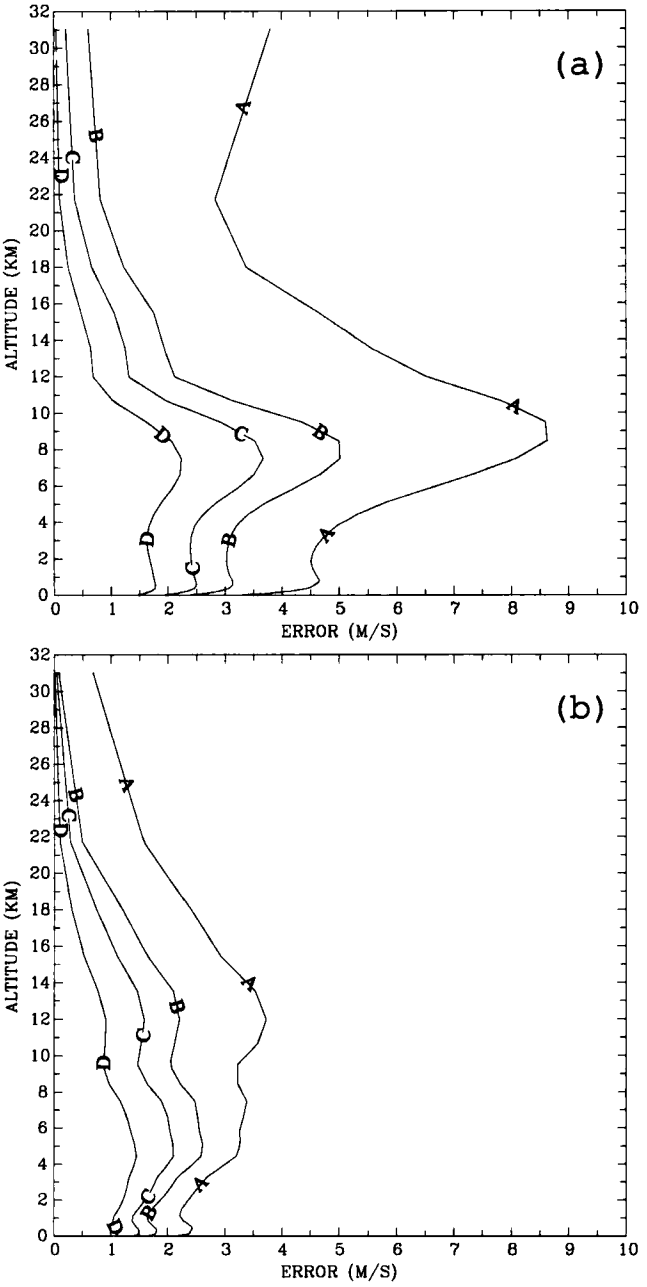

Figure 11. Variation with height of standard deviations of errors: (a) rotational component of the wind; (b) divergent component of the wind. In each panel, curve A shows the background errors and curves B, C and D, the (progressively smaller) associated representativeness-errors for resolutions of T42, T63 and T95 respectively.

representativeness error at 700 hPa represents about 70% of the background error and is greater than the observation errors given to radiosounding measurements (about $0.8\ g\ kg^{-1}$).

The standard deviation of background error for the rotational part of the wind shows a typical peak at 9 km (Fig. 11(a)). The representativeness error also has large values around this level, that can again be more important than typical values of observation error (about $3.5\ m\ s^{-1}$) for a truncation $K^S = 42$, and still very important for $K^S = 63$. The representativeness error for the divergent part of the wind is lower but increases close to the surface (Fig. 11(b)): the ratio of representativeness error to background error is also fairly large, because, just as do those for specific humidity, divergent-wind error spectra have a good deal of energy in the smaller scales.

5. Conclusion

The incremental formulation of the variational assimilation schemes has contributed greatly to the operational implementation of the variational approach in a number of NWP centres. However, this formulation leads to an inevitable representativeness-error, that is not the result of the potential high resolution of the forecast model, but of the resolution of the analysis increment (which can be much coarser).

The aim of this paper has been to formalize this particular representativeness-error but also to try to evaluate it in a realistic framework. Such results can lead to the choice of a sufficient resolution in order to cancel this error, or to provide a way to take it into account if the required computation cost with this resolution is too high.

In particular, it has been shown that the variance of the incremental representativeness-error can be determined by summing up the part of the background-error power-spectrum that is beyond the truncation of the simplified increment: this error will thus be larger for fields such as humidity that show more energy in the forecast errors for the smallest scales (and so are associated with shorter length-scales in physical space). Since the incremental representativeness-error can be seen as the result of the convolution of the background-error spectrum with a square window in spectral space, the correlation for this error presents a typical oscillating shape. Furthermore, the length-scale of the incremental representativeness-error correlation is directly related to the truncation K^S of the simplified space: this length-scale is longer for a coarser resolution K^S.

The computation of the statistics of both the ARPEGE background and incremental representativeness-errors confirms the fact that the latter can be roughly deduced from the application of a square window to the former. The characteristics of the representativeness-error correlations shown in the simple 1-D case are retrieved in this realistic framework: they show the same oscillating behaviour and their length-scales are more constant in the vertical than the corresponding background length-scales (which typically increase with altitude).

So, the variance of the representativeness error can be parametrized by summing up the part of the background-error power-spectra that is truncated by using a simplified increment. It is shown that the level of this error is not negligible at truncations T63 or T42, for instance, when values may be close to the observation error, or even greater for parameters such as humidity.

The computation of these variances was here assumed to be made for *in situ* instruments, measuring parameters close to model variables, such as radiosounding or aircraft measurements. Of course, there is a need to extend this computation to observations, such as satellite radiances, that give measurements indirectly linked to

model variables and that also sample a given volume of the atmospheric flow. This corresponds to a case when the observation operator H becomes more complex.

In a similar way, the present study focused on the 3D-Var case: however, it is clear that the assimilating model, used in a 4D-Var formulation (and that can be formally included in H), will also spread the representativeness error in a complex way. In particular, Fischer *et al.* (1998) showed, using a Kalman filter with a low-dimensional semi-geostrophic uniform potential-vorticity model, that maximal forecast-errors are concentrated along fronts and lead to some flow-dependency and anisotropy in the covariances that will also appear for the representativeness error. Some authors have proposed changes of coordinates in order to recover isotropy that may also help for the representativeness-error problem (Benjamin *et al.* 1991; Desroziers 1997).

ACKNOWLEDGEMENTS

The authors are grateful to Zhiquan Liu, Philippe Lopez and Florence Rabier for helpful discussion and comments on this article. We also wish to thank Andrew Lorenc, Associate Editor for the *Q. J. R. Meteorol. Soc.*, and the two referees for their thoughtful and constructive comments on the first version of this paper.

REFERENCES

Andersson, E., Pailleux, J., Thépaut, J.-N., Eyre, J. R., McNally, A. P., Kelly, G. A. and Courtier, P. 1994 Use of cloud-cleared radiances in three/four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.*, **120,** 627–653

Benjamin, S. G., Brewster, K. A., Brummer, R., Jewett, B. F. and Schlatter, T. W. 1991 An isentropic three-hourly data assimilation system using ACARS aircraft observations. *Mon. Weather Rev.*, **119,** 888–906

Courtier, P. 1997 Dual formulation of four-dimensional variational assimilation. *Q. J. R. Meteorol. Soc.*, **123,** 2449–2461

Courtier, P. and Geylen, J.-F. 1988 A global numerical weather model with variable resolution: Application to the shallow-water equations. *Q. J. R. Meteorol. Soc.*, **114,** 1321–1346

Courtier, P. and Talagrand, O. 1987 Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Q. J. R. Meteorol. Soc.*, **113,** 1329–1347

Courtier, P., Freydier, C., Geleyn, J.-F., Rabier, F. and Rochas, M. 1991 'The Arpège project at Météo-France'. Pp. 193–231 in *Proceedings of the ECMWF workshop on numerical methods in atmospheric models*, Reading, 9–13 September 1991, available from ECMWF, Reading, UK

Courtier, P., Thépaut, J.-N. and Hollingsworth, A. 1994 A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120,** 1367–1387

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F. and Fisher, M. 1998 The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Q. J. R. Meteorol. Soc.*, **124,** 1783–1807

Daley, R. 1991 *Atmospheric data analysis.* Cambridge University Press, Cambridge, UK

1993 Estimating observation error statistics for atmospheric data assimilation. *Ann. Geophysicae*, **11,** 634–647

Derber, J. C. and Wu, W.-S. 1998 The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Weather Rev.*, **126,** 2287–2297

Desroziers, G. 1997 A coordinate change for data assimilation in spherical geometry of frontal structures. *Mon. Weather Rev.*, **125,** 3030–3038

Desroziers, G., Mathiot, V., Orain, F. and Bernadet, P. 1995 'Estimation locale des covariances des erreurs de prévision d'un modèle spectral sur la sphère: Application au modèle à résolution variable ARPEGE'. [Local estimation of covariances of forecast errors in a spectral model on the sphere: Application to the ARPEGE variable resolution model.] Technical report 34, Centre National de Recherches Météorologiques, Toulouse, France

Desroziers, G., Nechad, B., Sadiki, W. and Thépaut, J.-N. 1997 'Analyse variationnelle du réseau de dropsondes de l'expérience FRONTS 92: Application du 3D-Var Arpège et discussion de l'erreur due à la formulation incrémentale de l'analyse à partir d'une maquette 1D-Var sur le cercle'. [Variational analysis of the network of dropsondes used during the FRONTS 92 experiment: Application of the Arpège 3D-Var system and discussion of the error caused by the incremental formulation of the analysis from a 1D-Var model on the circle.] Technical report 53, Centre National de Recherches Météorologiques, Toulouse, France

Eyre, J. R., Kelly, G., McNally, A. P., Andersson, E. and Persson, A. 1993 Assimilation of TOVS radiance information through one-dimensional variational analysis. *Q. J. R. Meteorol. Soc.*, **119,** 1427–1463

Fischer, C., Joly, A. and Lalaurette, F. 1998 Error growth and Kalman filtering within an idealized baroclinic flow. *Tellus*, **50A,** 596–615

Ide, K., Courtier, P., Ghil, M. and Lorenc, A. C. 1997 Unified notation for data assimilation: operational, sequential and variational. *J. Meteorol. Soc. Japan*, **75,** 181–189

Janisková, M., Veersé, F., Thépaut, J.-N., Desroziers, G. and Pouponneau, B. 1999 Impact of a simplified physical package in 4D-Var analyses of FASTEX situations. *Q. J. R. Meteorol. Soc.*, **125,** 2465–2485

Laroche, S. and Gautier, P. 1998 A validation of the incremental formulation of 4-D variational data assimilation in a nonlinear barotropic flow. *Tellus*, **50A,** 97–110

Le Dimet, F.-X. and Talagrand, O. 1986 Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A,** 97–110

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

Lorenc, A. C. 1986 Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **112,** 1177–1194

1988 Optimal nonlinear objective analysis. *Q. J. R. Meteorol. Soc.*, **114,** 205–240

Parrish, D. F. and Derber, J. C. 1992 The National Meteorological Center's Spectral Statistical Interpolation Analysis System. *Mon. Weather Rev.*, **120,** 1747–1763

Rabier, F., McNally, A., Andersson, E., Courtier, P., Undén, P., Eyre, J., Hollingsworth, A. and Bouttier, F. 1998 The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions. *Q. J. R. Meteorol. Soc.*, **124,** 1809–1829

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

Schmidt, F. 1977 Variable fine mesh in spectral global model. *Beitr. Phys. Atmos.*, **50,** 211–217

Thépaut, J.-N., Courtier, P., Belaud, G. and Lemaître, G. 1996 Dynamical structure functions in a four-dimensional variational assimilation: A case study. *Q. J. R. Meteorol. Soc.*, **122,** 535–561

Thépaut, J.-N., Alary, P., Caille, P., Cassé, V., Geleyn, J.-F., Moll, P., Pailleux, J., Piriou, J.-M., Puech, D. and Taillefer, F. 1998 'The operational global data assimilation system at Météo-France'. Pp. 25–31 in *Proceedings of HIRLAM 4 Workshop on Variational Analysis in Limited Area Models*, Toulouse, France

Veersé, F. and Thépaut, J.-N. 1998 Multiple-truncation incremental approach for four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.*, **124,** 1889–1908