Q. J. R. Meteorol. Soc. (2002), **128**, pp. 1367–1386

The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study

By Z.-Q. LIU[1,2]* and F. RABIER[1]
[1]*Centre National de Recherches Météorologiques, France*
[2]*National Satellite Meteorological Centre, China*

(Received 17 October 2000; revised 17 October 2001)

SUMMARY

In this paper, the optimal configurations of model resolution, observation resolution and observation density are investigated in a simple one-dimensional framework. In this context, the representativeness error is formalized and estimated before being used in the analysis-error formulation. Some optimal and suboptimal assimilation-schemes, differing from different approximations of observation-error covariance and observation operator, are compared. The optimal observation-extent is determined as a function of model resolution. Increasing the observation density is usually beneficial, except for suboptimal schemes similar to the ones used in operational practice. The impact of thinning the observations with correlated error is also studied from a suboptimal viewpoint.

KEYWORDS: Correlation Suboptimal Thinning

1. INTRODUCTION

In the past decade, a substantial part of the improvement of operational global numerical weather prediction (NWP) has come from use of variational assimilation techniques (Le Dimet and Talagrand 1986; Parrish and Derber 1992; Andersson *et al.* 1998; Lorenc *et al.* 2000; Rabier *et al.* 2000) and of satellite data. An essential parameter for remote-sensing instruments is the resolution of the observations (i.e., 'footprint' of a nadir measurement). The horizontal resolution of the observations can vary from several kilometres for a nadir-viewing instrument to hundreds of kilometres for a limb sounder. Another important parameter is the observation density, which can usually be characterized by the spacing between two adjacent observations. Finding optimal observation-resolution and density for various applications should be part of the assessment of current instruments and the design of future ones. In particular, for NWP, one would like to know how this optimal observation-resolution and density are related to model resolution. In general, there will always be an incentive to increase the model resolution to better resolve phenomena of smaller scale. Miller (1999) has shown that increasing the model resolutions (both of the analysis and of the forecast) at the European Centre for Medium-Range Weather Forecasts (ECMWF) is beneficial in providing better analyses and forecasts, particularly for severe weather events. Consequently, it is particularly interesting to examine how observations should be provided to best adapt to the model resolution.

According to the Bayesian probabilistic theorem (see Lorenc (1986) for a derivation of data assimilation), the variational formulation requires correct error-statistics of the background and the observations to be used in the cost function to obtain optimal-state estimation. However, these error statistics are still far from perfect. For example, the representativeness error (Lorenc 1986; Daley 1993), which is thought to introduce spatial correlations in the observational error, is difficult to estimate and expensive to specify in the minimization of the cost function. Thinning of observations (Järvinen and Undén 1997) is often used in operational practice, and it is an efficient way to reduce the effective error-correlation. Two open questions are (a) under which conditions the best balance between observation correlation and thinning can be reached and

---

* Corresponding author: Météo-France, CNRM/GMAP, 42 av. Coriolis, 31057 Toulouse cédex, France.
e-mail: zhiquan.liu@cnrm.meteo.fr

(b) whether the consideration of observation-error correlation in the cost function can further improve the analysis quality.

In this paper, these general issues will be dealt with in a systematic way in a simple one-dimensional (1-D) framework. In section 2, the 1-D experimental framework is presented. In section 3, experiments are performed to find some optimal configurations of model resolution, observation resolution and observation density for different observation-types and different analysis-schemes. Conclusions are given in section 4.

2. EXPERIMENTAL FRAMEWORK

In this section, we shall introduce the experimental 1-D framework used to simulate the results of various analyses. The characteristics of the observations, true background errors and true signal will be defined and the analysis equations will be recalled.

(*a*) *True signal and its covariance*

Let us consider a 1-D periodic domain $-\pi a \leqslant r \leqslant \pi a$. The length of the domain is taken as $L = 2\pi a \approx 8000$ km ($a = 1250$ km). We assume that the true signal $x^{\mathrm{t}}(r)$ in this 1-D domain is a continuous real function and can be expressed as a Fourier series, truncated at wave number K_{t}. The function values $x^{\mathrm{t}}(r_j)$ at any N discrete positions $r_j (j = 1, \ldots, N)$ can be written as $\mathbf{x}^{\mathrm{t}} = (x^{\mathrm{t}}(r_j))$,

$$\mathbf{x}^{\mathrm{t}} = \mathbf{F}^{\mathrm{t}}\hat{\mathbf{x}}^{\mathrm{t}}, \tag{1}$$

where $\hat{\mathbf{x}}^{\mathrm{t}}$ stands for the complex Fourier coefficients vector of length $M_{\mathrm{t}} = 2K_{\mathrm{t}} + 1$ with complex conjugate elements $\hat{x}_k^{\mathrm{t}} = (\hat{x}_{-k}^{\mathrm{t}})^*$. $\mathbf{F}^{\mathrm{t}}$ is the Fourier-transform matrix of dimension $N \times M_{\mathrm{t}}$ with elements $F_{jk} = \exp(ikr_j/a)$. Here the superscript t signifies 'true' and is used to distinguish the true values for the analysis variables described in the next subsection. In this study, there is no real atmospheric variable and therefore the unit of $x^{\mathrm{t}}(r)$ is arbitrary. However, the characteristics of the signal will be specified to approximate some realistic situation.

We assume that the covariance of the truth in Fourier space $\hat{\mathbf{S}}$ is diagonal, with elements $\sigma_{\mathrm{t}}^2\hat{\rho}_k^{\mathrm{t}}$. Here σ_{t}^2 is the true-signal variance. Throughout this study, one takes $\sigma_{\mathrm{t}} = 10$ — ten times the magnitude of the background and measurement errors which are specified to be 1. $\hat{\rho}_k^{\mathrm{t}}$ is the power spectrum of the second-order autoregressive correlation-function (Thiébaux 1976) given by

$$\rho_{\mathrm{t}}(r) = \left\{\cos(br) + \frac{\sin(br)}{L_{\mathrm{t}}b}\right\}\exp(-r/L_{\mathrm{t}}), \tag{2}$$

where b and L_{t} are constants and r $(0 \leqslant r \leqslant \pi a)$ is the absolute distance between two points. We chose $b = 4/a$ and $L_{\mathrm{t}} = a/3 = 417$ km as did Daley (1993). For these specifications, the value $K_{\mathrm{t}} = 200$ will be used, which is large enough to allow the scales beyond it to be ignored.

(*b*) *Analysed signal and background-error covariance*

The variables to be analysed are the Fourier spectral coefficients ranging from wave number $-K_{\mathrm{m}}$ to K_{m} ($K_{\mathrm{m}} \leqslant K_{\mathrm{t}}$). The associated analysis-mesh spacing is $\Delta r = 2\pi a/M$ with the dimension of the model $M = 2K_{\mathrm{m}} + 1$. Consequently, the model representation of the truth is a wave-number-limited filtered truth, which can be expressed as

$$\mathbf{x} = \mathbf{F}\hat{\mathbf{x}}. \tag{3}$$

Equation (3) is similar to (1), except that in $\hat{\mathbf{x}}$ and $\mathbf{F}$ there are no terms with wave number higher than K_m. Note that the superscript t has been removed to distinguish (3) from (1). In fact, the wave number-limited filter (rectangular window) approximately corresponds to a spatial smoothing given by

$$x(r_j) = \int_{-\pi a}^{\pi a} x^\mathrm{t}(r) w(r - r_j)\,\mathrm{d}r, \tag{4}$$

with the so-called cardinal weighting-function (Petersen and Middleton (1963); see appendix of the present paper for a derivation)

$$w(r - r_j) = \frac{\sin\{K_\mathrm{m}(r - r_j)/a\}}{\pi(r - r_j)}. \tag{5}$$

The background spectral vector is denoted by $\hat{\mathbf{x}}_\mathrm{b}$. Its error covariance $\hat{\mathbf{B}}$ is assumed to be diagonal with elements $\sigma_\mathrm{b}^2 \hat{\rho}_k^\mathrm{b}$. Here $\sigma_\mathrm{b}^2 = 1$ is the background-error variance and $\hat{\rho}_k^b$ is the power spectrum of a degenerate second-order autoregressive correlation-function (Thiébaux *et al.* 1986) given by

$$\rho_\mathrm{b}(r) = \left(1 + \frac{r}{L_\mathrm{b}}\right) \exp(-r/L_\mathrm{b}), \tag{6}$$

where $L_\mathrm{b} = a/6 = 208$ km is the background-error correlation-length.

(*c*) *Observations*

Most instruments can be roughly grouped into two categories: *in situ* instruments that measure the signal value at a point and remote-sensing instruments that measure some sort of spatial integral of the signal. Here we suppose that the perfect measure $y^\mathrm{t}(r_\mathrm{o})$ (without noise) at a point r_o from a remote-sensing instrument is a linear weighted integral of the underlying signal centred in r_o, namely

$$y^\mathrm{t}(r_\mathrm{o}) = \int_{-\pi a}^{\pi a} x^\mathrm{t}(r) w(r - r_\mathrm{o})\,\mathrm{d}r. \tag{7}$$

The weighting function $w(r - r_\mathrm{o})$ plays a spatial smoothing role. Moreover, (7) can be expressed as an equivalent filtering form in Fourier space given by

$$y^\mathrm{t}(r_\mathrm{o}) = \sum_{k=-K_\mathrm{t}}^{K_\mathrm{t}} \hat{w}_k \hat{x}_k \exp(ikr_\mathrm{o}/a), \tag{8}$$

where $\hat{w}_k$ is the spectral weighting coefficient at wave number k. For an observation vector $\mathbf{y}^\mathrm{t}$, (8) can be written as

$$\mathbf{y}^\mathrm{t} = \mathbf{G}^\mathrm{t}\hat{\mathbf{x}}^\mathrm{t} = \mathbf{F}^\mathrm{t}\mathbf{W}^\mathrm{t}\hat{\mathbf{x}}^\mathrm{t}, \tag{9}$$

where $\mathbf{W}^\mathrm{t}$ is a diagonal matrix with elements $\hat{w}_k$. $\mathbf{F}^\mathrm{t}$ and $\hat{\mathbf{x}}^\mathrm{t}$ are respectively the Fourier-transform matrix and the spectral-coefficients vector. We define $\mathbf{G}^\mathrm{t} = \mathbf{F}^\mathrm{t}\mathbf{W}^\mathrm{t}$ as the weighted Fourier-transform matrix.

In the present study, we consider weighting functions associated with two remote-sensing instruments:

(1) the uniform weighting-function with its corresponding spectral form

$$w(r)=\begin{cases}0 & \left(|r|>\dfrac{L_{\mathrm{o}}}{2}\right)\\ \dfrac{1}{L_{\mathrm{o}}} & \left(|r|\leqslant\dfrac{L_{\mathrm{o}}}{2}\right)\end{cases}\qquad \hat{w}_k=\begin{cases}1 & (k=0)\\ \dfrac{\sin(kL_{\mathrm{o}}/2a)}{kL_{\mathrm{o}}/2a} & (k\neq 0)\end{cases}, \tag{10}$$

(this is practical for horizontal sampling of a nadir-viewing instrument with a limited field of view L_{o})

(2) and the Gaussian weighting-function with its corresponding spectral form

$$\begin{aligned} w(r)&=\exp(-4r^2/L_{\mathrm{o}}^2)\Big/\int_{-\pi a}^{\pi a}\exp(-4r^2/L_{\mathrm{o}}^2)\,\mathrm{d}r\\ \hat{w}_k&=\exp(-k^2L_{\mathrm{o}}^2/16a^2).\end{aligned} \tag{11}$$

Here the constant 4 allows L_{o} to be roughly related to the realistic extent of the observation (e.g., in the sense that the weight at $r=L_{\mathrm{o}}/2$ is reduced to $1/e$ of the maximal weight). This Gaussian-type weighting-function has been used by Daley (1993) to study the representativeness error due to the misspecification of the observation operator in a similar 1-D context.

In situ measurements can be considered as a special case of the above two remote-sensing measurements with $L_{\mathrm{o}}=0$ in which $w(r)$ becomes a Dirac Delta function and with $\hat{w}_k$ being equal to one for all values of k. In Fig. 1, three weighting functions are plotted respectively in physical space (Fig. 1(a)) and Fourier space (Fig. 1(b)). For the sake of clarity, the weighting functions in physical space are plotted only in the interval [−200 km, 200 km]. In addition to the uniform and Gaussian weighting-functions, the cardinal weighting-function is also shown, as it corresponds to the representation of the model. The model spectral-truncation $K_{\mathrm{m}}=39$ corresponds to an analysis mesh of 100 km, which is the same as the observation resolution L_{o} of the uniform and Gaussian instruments. One can see that the model representation and the observation representation of the signal are significantly different even though they have the same resolution. In the next subsection, we will show how these differences are transformed into the term of the representativeness error. Here we have referred to L_{o} as the observation resolution. In some literature, the notion of the observation resolution is often referred to as the observation network spacing, whereas we will use the phrase ‘observation resolution’ to indicate the property of a remote-sensing instrument itself which should be carefully distinguished from the station interval or density of an observation network.

(*d*) *Representativeness error and the observation operator*

In this study, we consider only an observation network consisting of N equally-spaced stations. At each station, the instrument measures the signal with a weighting function described as above so that the observation vector $\mathbf{y}$ of length N can be expressed as

$$\mathbf{y}=\mathbf{G}^{\mathrm{t}}\hat{\mathbf{x}}^{\mathrm{t}}+\epsilon^{\mathrm{i}}, \tag{12}$$

where ϵ^{i} is the instrument-error vector. Now we define the observation operator $\mathbf{H}$ ($N\times M$ matrix) which computes the model representation of the observations $\mathbf{y}^{\mathrm{m}}$, starting from the analysed spectral vector $\hat{\mathbf{x}}$ of length M given by

$$\mathbf{y}^{\mathrm{m}}=\mathbf{H}\hat{\mathbf{x}}=\mathbf{H}\mathbf{T}\hat{\mathbf{x}}^{\mathrm{t}}, \tag{13}$$

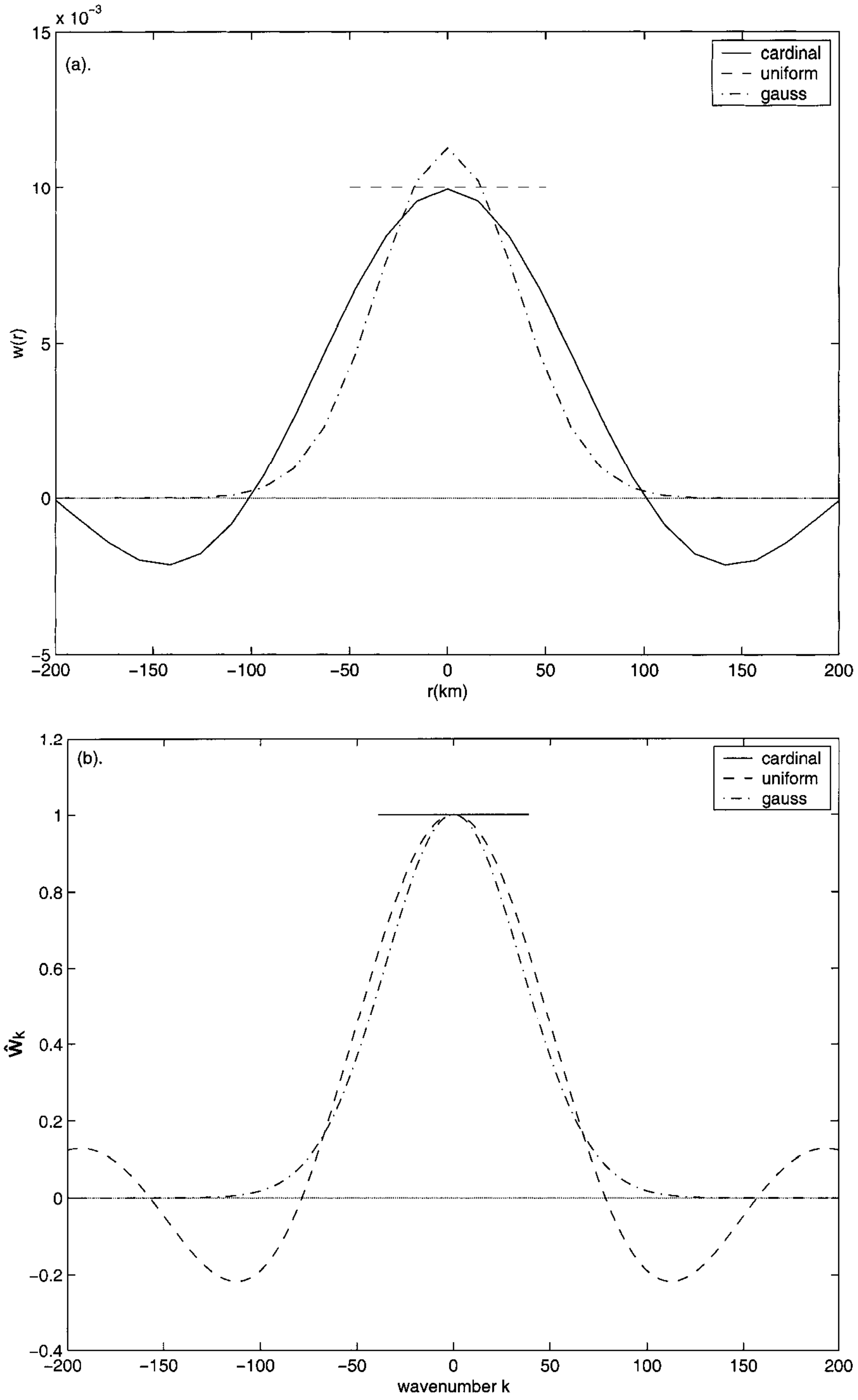

Figure 1. Observation weighting-functions: (a) spatial weighting-functions $w(r)$ at point $r = 0$ as a function of the ordinate r; (b) spectral weighting $\hat{w}_k$ as a function of wave number k. Corresponding parameters are the following: cardinal weighting-function: $K_{\rm m} = 39$ (analysis mesh $\Delta r = 100$ km); uniform and Gauss weighting-functions: $L_{\rm o} = 100$ km.

where the truncation operator $\mathbf{T}$ cuts off the complete spectral vector $\hat{\mathbf{x}}^{\mathrm{t}}$ to the analysed spectral vector $\hat{\mathbf{x}}$. Subtracting (12) from (13) and then performing the expectation operation, we obtain the observation-error covariance

$$\mathbf{R} = \mathbf{R}^{\mathrm{i}} + \mathbf{R}^{\mathbf{H}}, \tag{14}$$

where $\mathbf{R}^{\mathrm{i}}$ is the instrument-error covariance which is supposed to be diagonal with elements σ_{o}^2 equal to 1 (variance) and

$$\mathbf{R}^{\mathbf{H}} = (\mathbf{HT} - \mathbf{G}^{\mathrm{t}})\hat{\mathbf{S}}(\mathbf{HT} - \mathbf{G}^{\mathrm{t}})^* \tag{15}$$

is the representativeness-error covariance (Daley 1993; Mitchell and Daley 1997). Here, $\hat{\mathbf{S}}$ is the true signal spectral covariance defined in subsection 2(a) and the star stands for the complex conjugate transpose of the matrix. We have assumed that there is no correlation either between the instrument error and the representativeness error or between the error and the signal. The representativeness errors are correlated ($\mathbf{R}^{\mathbf{H}}$ is not diagonal) as a result of the correlated signal.

In practice, the weighting function of the remote-sensing instrument could either be unknown or be known but difficult to use in the computation, so that the measurement has to be considered simply as a point value in model space. To examine the effect of this, two observation operators are considered:

- $\mathbf{H}_{\mathbf{W}} = \mathbf{G} = \mathbf{FW}$, a version of the weighted Fourier transform matrix $\mathbf{G}^{\mathrm{t}}$ truncated at wave number K_{m}, which assumes that the observation weighting-function $w(r)$ and hence $\hat{w}_k$ are perfectly known and modelled correctly.
- $\mathbf{H}_{\mathbf{I}} = \mathbf{F}$, the Fourier-transform matrix truncated at wave number K_{m}, which ignores the weighting matrix $\mathbf{W}$ and means that any observation is always regarded as if it had the same cardinal weighting-function as the model. Here the subscript '**I**' indicates that the weighted matrix $\mathbf{W}$ is replaced by an identity matrix.

Note that the matrices $\mathbf{G}$, $\mathbf{F}$ and $\mathbf{W}$ only contain wave-number components with $|k| \leqslant K_{\mathrm{m}}$ of the corresponding $\mathbf{G}^{\mathrm{t}}$, $\mathbf{F}^{\mathrm{t}}$ and $\mathbf{W}^{\mathrm{t}}$. Both $\mathbf{H}_{\mathbf{W}}$ and $\mathbf{H}_{\mathbf{I}}$ allow us to transform directly the spectral coefficients to the values at any observation positions and no intermediate interpolation process is required. For $\mathbf{H} = \mathbf{H}_{\mathbf{W}}$, the variance of the representativeness error given by

$$\sigma_{\mathbf{H}_{\mathbf{W}}}^2 = \sigma_{\mathrm{t}}^2 \sum_{|k|>K_{\mathrm{m}}} \hat{w}_k^2 \hat{\rho}_k^{\mathrm{t}} \tag{16}$$

contains only the contribution of the unresolved scales. For $\mathbf{H} = \mathbf{H}_{\mathbf{I}}$, the variance of the representativeness error given by

$$\sigma_{\mathbf{H}_{\mathbf{I}}}^2 = \sigma_{\mathrm{t}}^2 \left[\sum_{|k|>K_{\mathrm{m}}} \hat{w}_k^2 \hat{\rho}_k^{\mathrm{t}} + \sum_{|k|\leqslant K_{\mathrm{m}}} (1 - \hat{w}_k)^2 \hat{\rho}_k^{\mathrm{t}} \right] \tag{17}$$

contains both the contribution of the unresolved scales and the misrepresentation of the observation operator in the resolved scales. In practice, the second term stands for the difference between the model representation and the observation representation of the signal in model-resolved scales. Equations (16) and (17) are, practically, the sum of spectral variance components of the representativeness error (making use of the fact that the diagonal covariance matrix in spectral space corresponds to an isotropic covariance in physical space and that the sum of the spectral variance components is equal to the average of the error variance in physical space). They will be used in section 3 for explanations of the representativeness error. As a result of using the Fourier-transform matrix, both $\sigma_{\mathbf{H_W}}$ and $\sigma_{\mathbf{H_I}}$ are independent of the position of observations, and of their density, for both regular and irregular observation-networks (Daley 1993).

(*e*) *Analysis schemes*

Applying the variational formulation (Lorenc 1986) to our 1-D problem, the optimal estimation of the model spectral-coefficients is to minimize a cost function given by

$$J(\hat{\mathbf{x}}) = \tfrac{1}{2}(\hat{\mathbf{x}} - \hat{\mathbf{x}}_{\mathrm{b}})^{*}\hat{\mathbf{B}}^{-1}(\hat{\mathbf{x}} - \hat{\mathbf{x}}_{\mathrm{b}}) + \tfrac{1}{2}(\mathbf{H}\hat{\mathbf{x}} - \mathbf{y})^{*}\mathbf{R}^{-1}(\mathbf{H}\hat{\mathbf{x}} - \mathbf{y}). \tag{18}$$

We recall that $\hat{\mathbf{x}}$ is the spectral vector to be analysed (truncated at wave number K_{m}) and $\hat{\mathbf{x}}_{\mathrm{b}}$ is the background vector. $\mathbf{y}$, $\mathbf{H}$, $\hat{\mathbf{B}}$ and $\mathbf{R}$ are respectively the observation vector, the observation operator, the (spectral) background-error covariance and the observation-error covariance. The star stands for the complex conjugate transpose. In the case of a linear observation-operator, the direct solution of minimizing (18) is given by

$$\hat{\mathbf{x}}_{\mathbf{a}} = \hat{\mathbf{x}}_{\mathbf{b}} + \mathbf{K}(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_{\mathbf{b}}), \tag{19}$$

with the gain matrix

$$\mathbf{K} = \hat{\mathbf{B}}\mathbf{H}^{*}(\mathbf{H}\hat{\mathbf{B}}\mathbf{H}^{*} + \mathbf{R})^{-1} = (\hat{\mathbf{B}}^{-1} + \mathbf{H}^{*}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{*}\mathbf{R}^{-1}. \tag{20}$$

For any gain matrix $\mathbf{K}$, the general analysis-error covariance (Ghil and Malanotte-Rizzoli 1991) is given by

$$\hat{\mathbf{A}} = (\mathbf{I} - \mathbf{KH})\hat{\mathbf{B}}^{\mathrm{t}}(\mathbf{I} - \mathbf{KH})^{*} + \mathbf{K}\mathbf{R}^{\mathrm{t}}\mathbf{K}^{*}, \tag{21}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{H}$ is the specified observation operator. The superscript t is used to indicate 'true' $\hat{\mathbf{B}}$ and $\mathbf{R}$, which should be distinguished from the specified ones included in the gain matrix $\mathbf{K}$. Here we recall that the true observation-error covariance includes the contribution of both the instrument and representativeness errors ((14) and (15)). The quality of the analysis scheme depends strongly on the observation operator $\mathbf{H}$, the observation-error covariance $\mathbf{R}$ and the background-error covariance $\hat{\mathbf{B}}$ being correctly specified. If this is so, then the scheme is optimal and the general error-covariance equation (21) reduces to a simpler form as

$$\hat{\mathbf{A}} = (\hat{\mathbf{B}}^{-1} + \mathbf{H}^{*}\mathbf{R}^{-1}\mathbf{H})^{-1}. \tag{22}$$

That is, the analysis precision (inverse of error) is the sum of the precision of the background and the observations. In fact, the general expression (21) can be used in all cases to compute the analysis error. In the optimal case, one would obtain exactly the same analysis error from (21) and (22). The spectral covariance can be transformed to grid-point space using the inverse Fourier-transform matrix.

In this study, we focus on the effect of different $\mathbf{H}$ and $\mathbf{R}$, whereas $\hat{\mathbf{B}}$ is always assumed to be specified correctly. We consider the following schemes with different configurations:

- **scheme $\mathbf{H_W R}$**: with the 'weighted' observation-operator $\mathbf{H} = \mathbf{H_W}$ and the complete observation-error covariance $\mathbf{R} = \mathbf{R}^{\mathrm{i}} + \mathbf{R}^{\mathbf{H_W}}$, the corresponding analysis-error covariance can be written as

$$\hat{\mathbf{A}} = [\hat{\mathbf{B}}^{-1} + \mathbf{W}^*\mathbf{F}^*(\mathbf{R}^{\mathrm{i}} + \mathbf{R}^{\mathbf{H}_{\mathrm{W}}})^{-1}\mathbf{F}\mathbf{W}]^{-1}; \tag{23}$$

- **scheme $\mathbf{H}_{\mathbf{I}}\mathbf{R}$**: with the 'simplified' observation operator $\mathbf{H} = \mathbf{H}_{\mathbf{I}}$ and the complete observation-error covariance $\mathbf{R} = \mathbf{R}^{\mathrm{i}} + \mathbf{R}^{\mathbf{H}_{\mathrm{I}}}$, the corresponding analysis-error covariance is

$$\hat{\mathbf{A}} = [\hat{\mathbf{B}}^{-1} + \mathbf{F}^*(\mathbf{R}^{\mathrm{i}} + \mathbf{R}^{\mathbf{H}_{\mathrm{I}}})^{-1}\mathbf{F}]^{-1}; \tag{24}$$

- **scheme $\mathbf{H}_{\mathbf{W}}\mathbf{R}_{\mathrm{diag}}$**: with the 'weighted' observation operator $\mathbf{H} = \mathbf{H}_{\mathbf{W}}$ and the diagonal observation-error covariance $\mathbf{R} = \mathbf{R}^{\mathrm{i}} + \mathrm{diag}(\mathbf{R}^{\mathbf{H}_{\mathrm{W}}})$. In this case, the data with correlated error are assimilated as if they were uncorrelated;
- **scheme $\mathbf{H}_{\mathbf{I}}\mathbf{R}_{\mathrm{diag}}$**: with the 'simplified' observation operator $\mathbf{H} = \mathbf{H}_{\mathbf{I}}$ and the diagonal observation-error covariance $\mathbf{R} = \mathbf{R}^{\mathrm{i}} + \mathrm{diag}(\mathbf{R}^{\mathbf{H}_{\mathrm{I}}})$.

Note that $\mathbf{R}^{\mathbf{H}_{\mathrm{W}}}$ and $\mathbf{R}^{\mathbf{H}_{\mathrm{I}}}$ are the realistic representativeness-error covariances obtained for the corresponding observation-operators $\mathbf{H}_{\mathbf{W}}$ and $\mathbf{H}_{\mathbf{I}}$. Our study is purely statistical and thus there is no need to simulate the signal and the observations, but only to manipulate the covariance matrices. In the next section, we attempt to answer the questions posed in the introduction by examining these analysis schemes with various configurations of the model resolution and the observation resolution and density.

3. RESULTS

In this section, the representativeness-error covariance is estimated for various configurations and then used in the analysis schemes. The analysis-error variance is used as the criterion of the analysis quality. Two kinds of observation network are considered in practical computation: uniform (equally-spaced) and homogeneous. The latter is achieved by dividing the domain into equal intervals and placing one observation station randomly within each interval (Daley 1993). As explained in subsection 2(d), the two observation operators used in our computations are of Fourier type so that the representativeness-error variance is independent of the location and density of observations. This results in the analysis error having no large differences between uniform and homogeneous networks having the same number of observations (the analysis error for the homogeneous network is slightly bigger). Consequently, only the results for the uniform observation network are shown below.

(*a*) *Case with uniform weighting-function*

This kind of observation is defined in (10) and shown in Figs. 1(a,b) (dashed line). The representativeness r.m.s. error $\sigma_{\mathbf{H}}$ as a function of the observation resolution L_{o} is given in Fig. 2(a) for $\mathbf{H} = \mathbf{H}_{\mathbf{W}}$ and $\mathbf{H} = \mathbf{H}_{\mathbf{I}}$. The computational configuration is the following: wave-number limit of the true signal $K_{\mathrm{t}} = 200$, standard deviation of true signal $\sigma_{\mathrm{t}} = 10$, model truncation $K_{\mathrm{m}} = 39$ (analysis mesh $\Delta r = 100$ km) and observation number $N = 79$ (the same as the dimension of the model). As expected, at $L_{\mathrm{o}} = 0$ (*in situ* measurement) where $\hat{w}_k$ is equal to one for all values of k and the two observation operators are identical, the representativeness error is the same. Moreover, we note that the representativeness error for $\mathbf{H} = \mathbf{H}_{\mathbf{W}}$ decreases monotonously with the increase of L_{o}. For $\mathbf{H} = \mathbf{H}_{\mathbf{I}}$, a minimum r.m.s. error is situated at $L_{\mathrm{o}} = 80$ km. These features are consistent with those used by Daley (1993, Fig. 2) and can be explained by (16), (17) and Fig. 1. For $\mathbf{H} = \mathbf{H}_{\mathbf{W}}$, the representativeness error contains only the contribution of the unresolved scale, which decreases as L_o increases ($\hat{w}_k$ will be narrower). For $\mathbf{H} = \mathbf{H_I}$, the representativeness error contains both the contribution of the unresolved and resolved scales. As L_o increases, narrower spectral weight increases the distance between the cardinal and the uniform spectral weighting-curves (see Fig. 1(b)). This causes the errors in resolved scales to increase. The minimum of the representativeness error for $\mathbf{H} = \mathbf{H_I}$ is the result of the compromise between the errors resulting from the unresolved and resolved scales. Moreover, L_o (80 km) at the minimum position is comparable to the analysis mesh size (100 km). That is, the most representative observation is the uniform average of the signal within the analysis mesh box. This also indicates that it is reasonable to use the average of values within the analysis mesh box as an approximation to the value at a grid point in the model. Note that the overall observation-error correlation is the result of the sum of the uncorrelated instrument-error and the correlated representativeness-error. In the case shown in Fig. 2, the maximal representativeness r.m.s. error is 0.53 so that the maximal possible correlation for adjacent observations is not greater than $0.53^2/(1 + 0.53^2) \approx 0.22$ (diagonal elements of total observation-error covariance are $1 + 0.53^2$ and off-diagonal elements from the representativeness-error covariance are smaller than 0.53^2). This weak observation-error correlation implies that the schemes $\mathbf{H_W R}_\mathrm{diag}$ and $\mathbf{H_I R}_\mathrm{diag}$ are very similar to that of the schemes $\mathbf{H_W R}$ and $\mathbf{H_I R}$. Consequently, in this subsection, only the results of the schemes $\mathbf{H_W R}$ and $\mathbf{H_I R}$ are discussed. The effect of the correlation will be treated in subsection 3(c).

Analysis r.m.s. errors as a function of the observation resolution L_o are shown in Fig. 2(b). Here we recall that both the background- and the instrument-error variances are taken as 1. The correlation length of the background error L_b is 208 km. One sees that for the scheme $\mathbf{H_I R}$ the variation of the analysis error with L_o is consistent with that of the corresponding representativeness-error. A minimum analysis-error appears in the same L_o position (80 km) as that of the representativeness error. However, this is not the case for the scheme $\mathbf{H_W R}$ where a minimum analysis-error at $L_\mathrm{o} = 60$ km is not associated with a minimum representativeness-error. In addition, we also note that the scheme $\mathbf{H_W R}$ is not always better than the scheme $\mathbf{H_I R}$ which seems to be counter-intuitive. It seems reasonable to believe that a more accurate observation-operator $\mathbf{H_W}$ would result in a better analysis than the approximate one $\mathbf{H_I}$. Indeed, the representativeness error is smaller for $\mathbf{H} = \mathbf{H_W}$ than for $\mathbf{H} = \mathbf{H_I}$. However, the observation operator also plays a role in projecting the instrument and representativeness errors to spectral space. In fact, if we examine only the observation term in (23) and (24), i.e., compare the projection of the observation precision (inverse of error) with the model space $\mathbf{W}^*\mathbf{F}^*(\mathbf{R}^\mathrm{i} + \mathbf{R}^{\mathbf{H_W}})^{-1}\mathbf{FW}$ and $\mathbf{F}^*(\mathbf{R}^\mathrm{i} + \mathbf{R}^{\mathbf{H_I}})^{-1}\mathbf{F}$, we find that the scheme $\mathbf{H_W R}$ is less accurate than the scheme $\mathbf{H_I R}$ for all values of L_o (Fig. 3). To explain this, we show in Fig. 4 the spectral precision (reciprocal of spectral variance) of the observation term as a function of wave number for $L_\mathrm{o} = 20$ km and $L_\mathrm{o} = 100$ km. One can see that the reduction of spectral precision of the observation term with the increase of wave number is larger for $\mathbf{H} = \mathbf{H_W}$ (solid line) than for $\mathbf{H} = \mathbf{H_I}$ (dashed line). This is because the presence of the spectral weighting-matrix $\mathbf{W}$ produces an additional damping effect for $\mathbf{H} = \mathbf{H_W}$. The order of the damping at wave number k would be $\hat{w}_k^2$. This suggests that, to choose a good observation-operator, we should examine not only its ability to reduce the representativeness error, but also its projection effect onto the model space. One also notes that the variation range of the analysis error in Fig. 2(b) is much smaller than that in Fig. 3 (particularly for the scheme $\mathbf{H_W R}$). This is because the background error has a red spectrum (spectral errors decrease with increase of wave number), which means that the background has a very high spectral precision for high wave-numbers

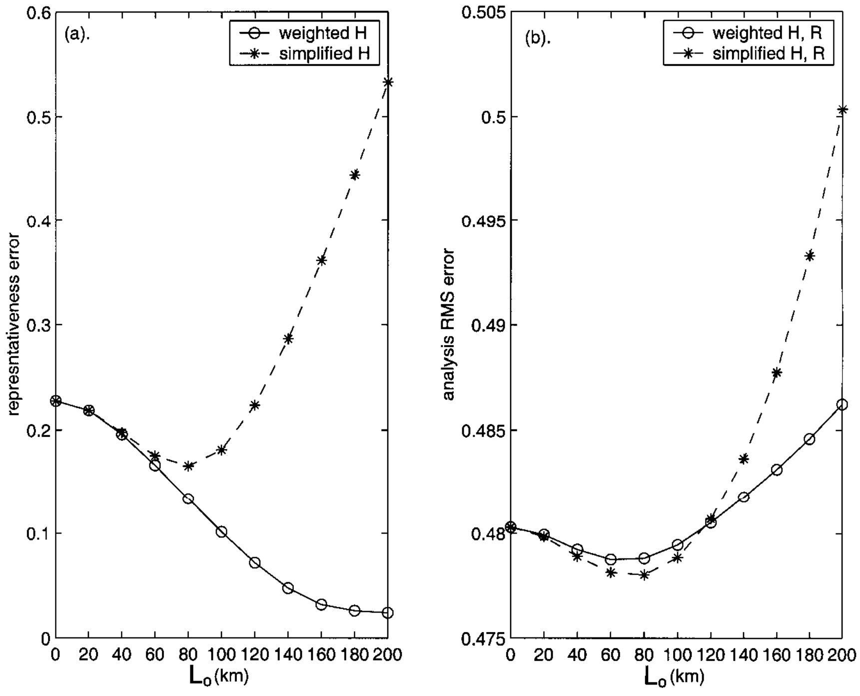

Figure 2. Errors as a function of the observation resolution L_o for the observation with the uniform weighting-function: (a) representativeness error; (b) analysis error. In each panel the two curves are respectively for the 'weighted' observation operators $\mathbf{H}_\mathrm{W} = \mathbf{G}$ (solid line) and the 'simplified' one $\mathbf{H}_\mathrm{I} = \mathbf{F}$ (dashed line). Configuration is the following: wave-number limit of the true signal $K_t = 200$, standard deviation of true signal $\sigma_t = 10$, model truncation $K_m = 39$ (analysis mesh $\Delta r = 100$ km), background- and observation-error variances $\sigma_b = \sigma_o = 1$, correlation length of the background error $L_b = 208$ km and number of observations $N = 79$ (the same as the dimension of the model).

and can compensate the degradation in small scales induced by the damping effect of the spectral weighting-matrix **W** (the contribution of the observation term is very small for small scales). In consequence, the difference between the two schemes in Fig. 2(b) will come mainly from the contribution of large scales (e.g., wave number $|k| < 20$). This is why the scheme $\mathbf{H}_\mathrm{W}\mathbf{R}$ is better than the scheme $\mathbf{H}_\mathrm{I}\mathbf{R}$ for some large L_o where spectral precision in large scales is higher for the scheme $\mathbf{H}_\mathrm{W}\mathbf{R}$ than for the scheme $\mathbf{H}_\mathrm{I}\mathbf{R}$ as shown in Fig. 4(b).

In the next experiment, two fields were created by doubling and halving the initial number of observations, whilst keeping an equal, but different, spacing between the observations in each of the new fields. The result is shown in Fig. 5. Note that the variance of the representativeness error is independent of the observation density in our context and therefore Fig. 2(a) is still valid. For more observations (Fig. 5(a)), we note that, for both schemes, a value of L_o is preferred which is smaller than that in Fig. 2(b). The optimal observation-resolution (40 km) for the scheme $\mathbf{H}_\mathrm{I}\mathbf{R}$ becomes half of that in Fig. 2(b)—it simply corresponds to the increase of the observation number by two. For fewer observations (Fig. 5(b)), the optimal L_o for the scheme $\mathbf{H}_\mathrm{I}\mathbf{R}$ is still 80 km as in Fig. 2(b). Moreover, a minimum analysis-error for the scheme $\mathbf{H}_\mathrm{W}\mathbf{R}$ appears at about $L_o = 130$ km which is greater than the analysis mesh.

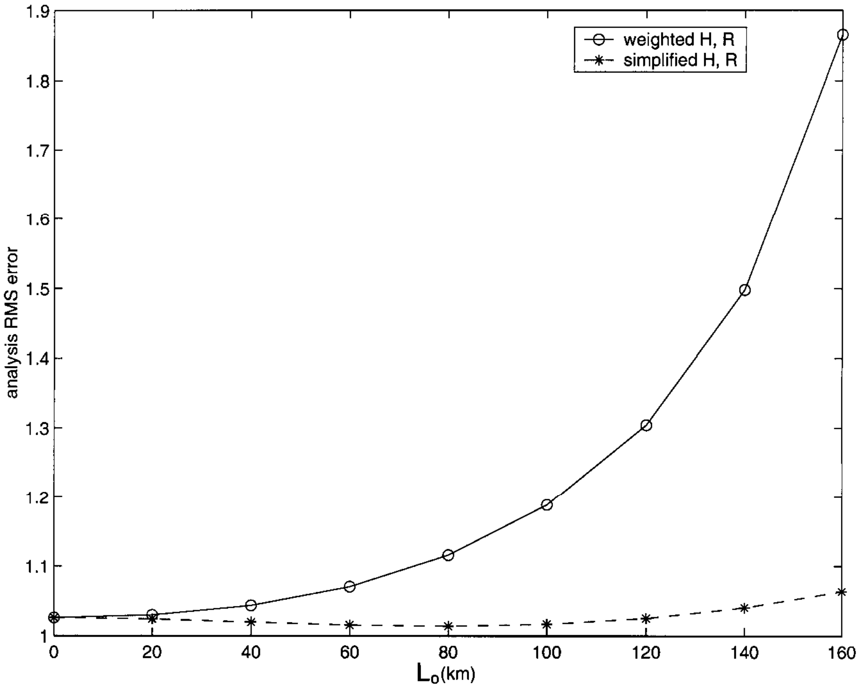

Figure 3. As Fig. 2(b), except that the background term is not considered.

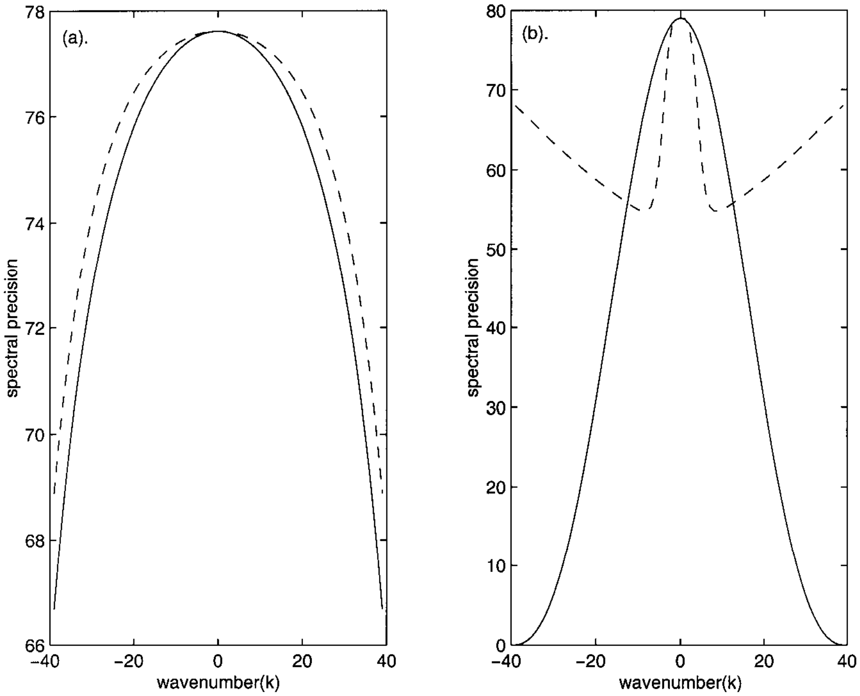

Figure 4. Spectral error precision (reciprocal of spectral variance) as a function of wave number: (a) $L_o = 20$ km; (b) $L_o = 200$ km. Configurations are as Fig. 2. Solid line: diagonal elements of the matrix $\mathbf{W}^*\mathbf{F}^*(\mathbf{R}^{\mathrm{i}} + \mathbf{R}^{\mathbf{H}_{\mathrm{W}}})^{-1}\mathbf{F}\mathbf{W}$. Dashed line: diagonal elements of the matrix $\mathbf{F}^*(\mathbf{R}^{\mathrm{i}} + \mathbf{R}^{\mathbf{H}_{\mathrm{I}}})^{-1}\mathbf{F}$.

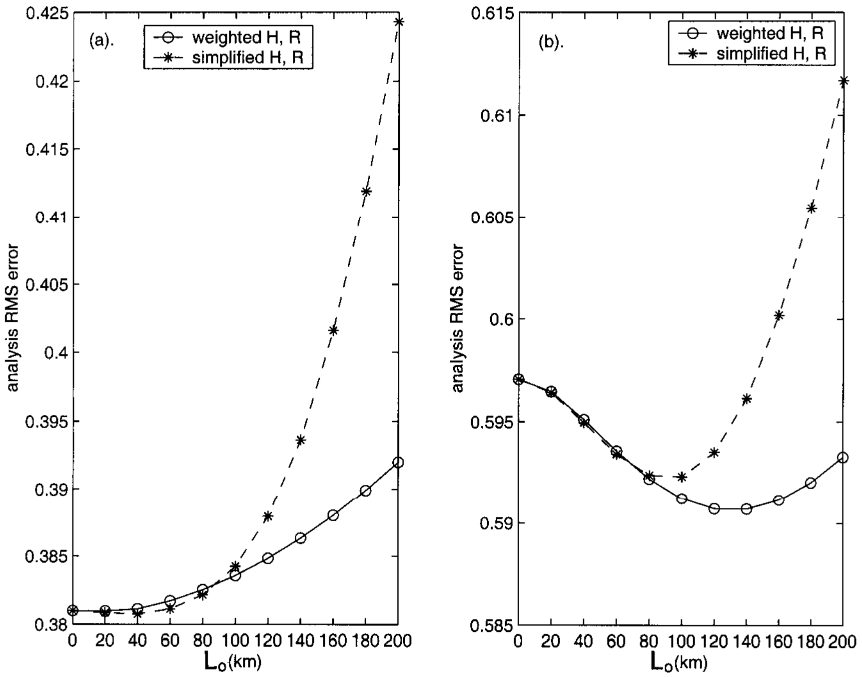

Figure 5. As Fig. 2(b), except for observation number: (a) $N = 159$; (b) $N = 39$.

These results suggest that in the case where the observation number N is larger than the model grid-point number M, the optimal observation-resolution for a uniform weighting-function can be obtained by an approximate relation $L_o \approx \Delta r \frac{M}{N}$, with Δr the analysis mesh size. That is, the denser the observation network, the higher the preferred observation resolution. Moreover, the simplified observation-operator $\mathbf{H}_\mathrm{I}$ can be used with good analysis accuracy in this case. For sparser observations, it is possibly more beneficial to have a larger value for the observation 'extent' or 'resolution' L_o than for the analysis mesh size Δr. In this case, the weighted observation-operator $\mathbf{H}_\mathrm{W}$ should be adopted. For example, in the case where a remote-sensing measurement covers multiple analysis mesh boxes, one should average these grid-point values (this can be regarded as an approximation of our weighted observation-operator) rather than interpolate them for comparing to the real observation.

Changing the model resolution leads to similar results. For example, for the scheme $\mathbf{H}_\mathrm{I}\mathbf{R}$, one finds an optimal observation-resolution $L_o = 160$ km and $L_o = 80$ km in the case $\Delta r = 200$ km respectively for $N = M$ and for $N = 2M$. Another parameter possibly influencing the results is the correlation length of the background error L_b. Figure 6 gives the results for $L_\mathrm{b} = 100$ km and $L_\mathrm{b} = 300$ km. Other parameters are the same as those in Fig. 2(b). One sees that the correlation length of the background error affects to a large extent the overall analysis accuracy but does not modify the overall shape of the curves. In particular, the optimal observation resolution for the scheme $\mathbf{H}_\mathrm{I}\mathbf{R}$ is not sensitive to the variation of L_b.

Let us now turn to study the effect of the different observation weighting-functions.

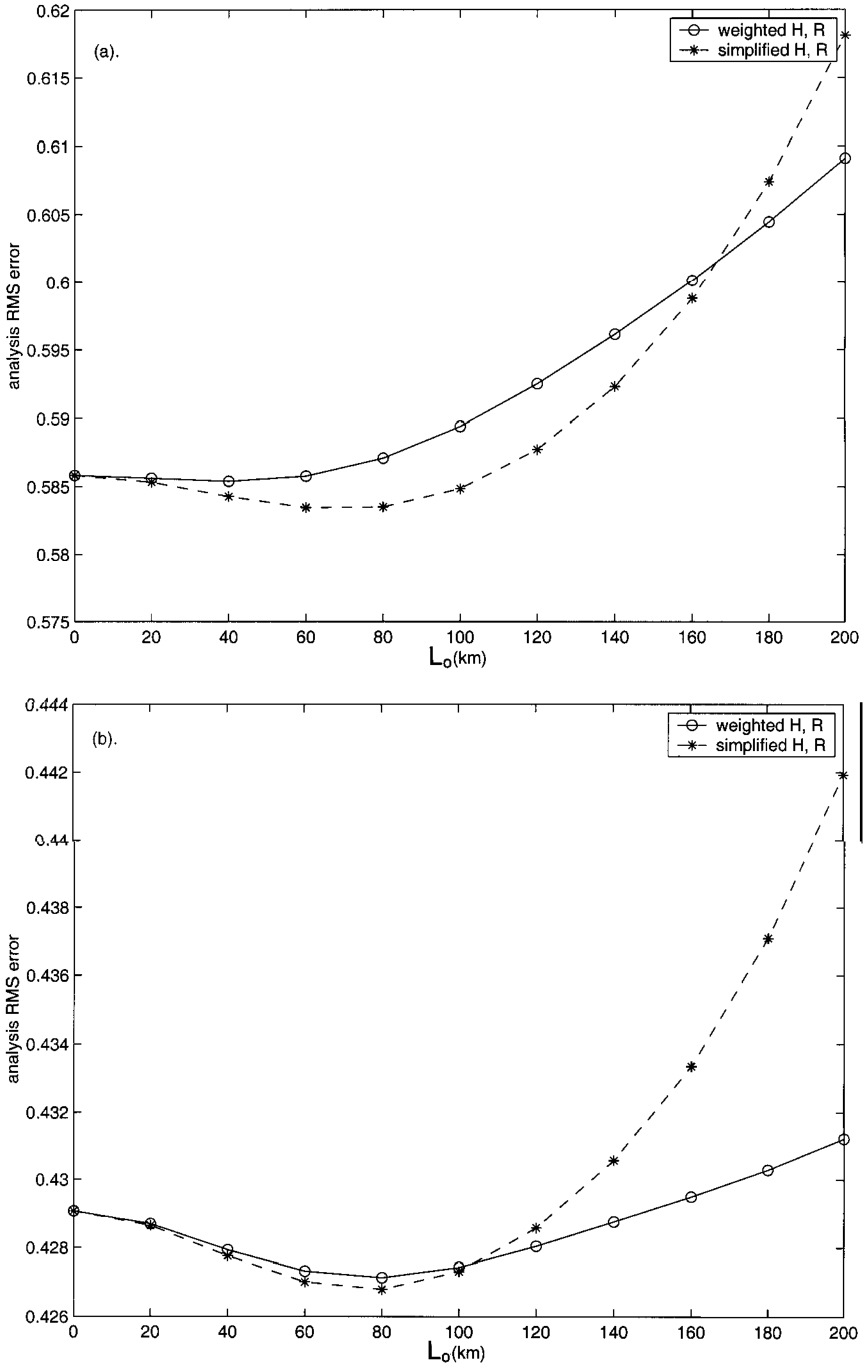

Figure 6. As Fig. 2(b), except for the correlation length of the background error: (a) $L_b = 100$ km; (b) $L_b = 300$ km.

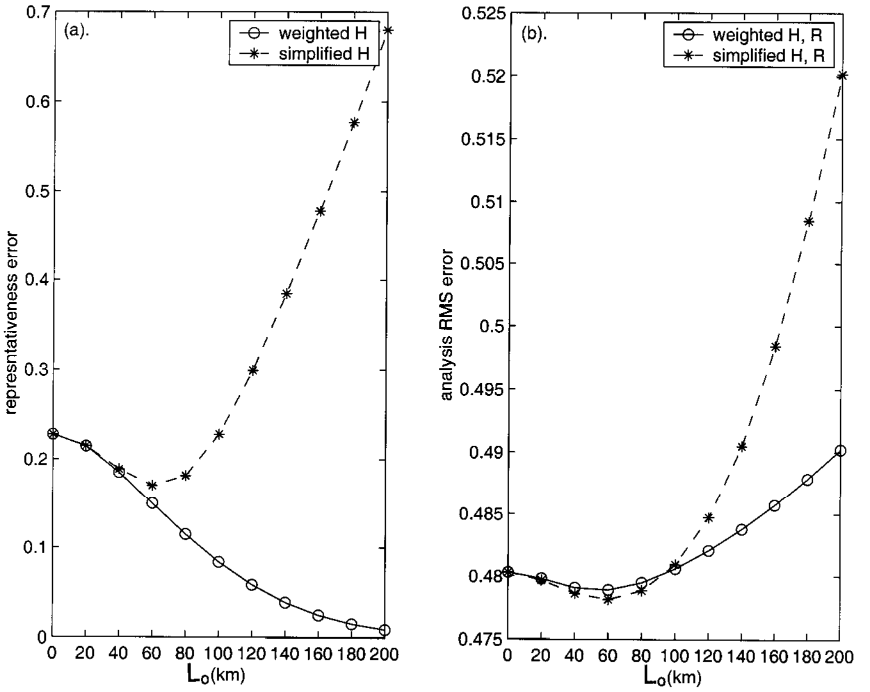

Figure 7. Errors as a function of observation resolution L_o for observations using the Gaussian weighting-function: (a) representativeness error; (b) analysis error. Configuration is as Fig. 2.

(*b*) *Case with the Gaussian weighting-function*

This kind of observation is defined in (11). It can be used to approximate the horizontal weighting-function of some limb- and occultation-sounding instruments. Figures 7(a,b) show the variation of the representativeness error and the analysis error as a function of the observation resolution L_o. The configurations are the same as in Fig. 2. One recalls the analysis mesh size $\Delta r = 100$ km and the correlation length of the background error $L_b = 208$ km. As expected, the qualitative aspects are the same as for the uniform weighting-function. For example, the representativeness error (Fig. 7(a)) for the scheme $\mathbf{H}_W\mathbf{R}$ decreases monotonously with the increase of L_o. For the scheme $\mathbf{H}_I\mathbf{R}$, a minimum representativeness error also appears (for $L_o = 60$ km). However, the two observation types lead to different results in quantitative aspects. From Fig. 1, one can see that for the same L_o the realistic observation-extent of the Gaussian instrument should be slightly larger than that of the uniform instrument. Moreover, the attenuation of the signal towards small scales is more rapid for the Gaussian weighting-function. This leads to such results that, for the observation operator $\mathbf{H}_W$, the representativeness error of the Gaussian observation is smaller than that of the uniform observation, and inversely for the operator $\mathbf{H}_I$. Correspondingly, the optimal observation-resolution L_o is 60 km which corresponds to a realistic extent of the observation close to 80 km—the optimal L_o in the case of the uniform observation. Similar relation between the analysis mesh size, the optimal L_o and the observation number is also obtained when we vary the model resolution and the observation number as was done for the uniform observation.

(*c*) *Thinning and observation-error correlation*

So far, the observation-error correlation resulting from representativeness error is so weak that the results of the schemes $\mathbf{H}_\mathbf{W}\mathbf{R}_{\mathrm{diag}}$ and $\mathbf{H}_\mathbf{I}\mathbf{R}_{\mathrm{diag}}$ cannot be distinguished from those of the schemes $\mathbf{H}_\mathbf{W}\mathbf{R}$ and $\mathbf{H}_\mathbf{I}\mathbf{R}$. To examine the impact of the suboptimal schemes when error correlation is strong, we consider a case of uniform observations with $L_\mathrm{o} = 400$ km. The analysis mesh is taken as 100 km, and the other configurations, such as the true signal-covariance as well as the background and the instrument errors, are kept the same as before. For the simplified observation-operator $\mathbf{H}_\mathbf{I}$, the representativeness r.m.s. error reaches 1.54, and the maximal correlation allowed for two adjacent observations is $1.54^2/(1 + 1.54^2) = 0.7$. For the weighted observation-operator $\mathbf{H}_\mathbf{W}$, the representativeness error almost vanishes so that the overall observation-error correlation is negligible. Figure 8(a) shows the observation-error correlation relative to $r = 0$ for an observation interval of 50 km for $\mathbf{H}_\mathbf{I}$. The error correlation for two adjacent observation reaches 0.6, and a weak negative correlation is present for distances greater than 200 km. Figure 8(b) shows the variation of the analysis r.m.s. error as a function of the observation interval for the four schemes. For the suboptimal scheme $\mathbf{H}_\mathbf{I}\mathbf{R}_{\mathrm{diag}}$, which ignores the observation-error correlation, the analysis r.m.s. error does not always decrease with decrease of the observation interval (which means that more observations are inserted), and an optimal observation-interval producing the minimum analysis-error is about 80 km. The scheme $\mathbf{H}_\mathbf{I}\mathbf{R}$ can still benefit from the dense observation since it takes into account the observation-error correlation correctly. The vanishing error correlation for $\mathbf{H}_\mathbf{W}$ leads to identical results for the schemes $\mathbf{H}_\mathbf{W}\mathbf{R}$ and $\mathbf{H}_\mathbf{W}\mathbf{R}_{\mathrm{diag}}$. They are much better than the schemes $\mathbf{H}_\mathbf{I}\mathbf{R}$ and $\mathbf{H}_\mathbf{I}\mathbf{R}_{\mathrm{diag}}$. In addition, one notes that for the schemes $\mathbf{H}_\mathbf{W}\mathbf{R}$ and $\mathbf{H}_\mathbf{W}\mathbf{R}_{\mathrm{diag}}$ the analysis error seems to tend to zero with decrease of the observation interval whereas this is not the case for the scheme $\mathbf{H}_\mathbf{I}\mathbf{R}$. This is consistent with the theoretical results of Mitchell and Daley (1997, Fig. 6 and Appendix C): (i) with a perfect observation operator (e.g., $\mathbf{H}_\mathbf{W}$, which makes the representativeness error vanish in the resolved scales), the analysis error (in the resolved scales) approaches zero asymptotically as observation density increases, but (ii) an imperfect observation-operator (e.g., $\mathbf{H}_\mathbf{I}$, which makes the representativeness error not equal to zero in the resolved scales) gives rise to strongly correlated error, and the analysis error does not vanish as observation density increases.

The results above indicate that thinning of observations is necessary for a dense observation-network with strongly correlated error if a suboptimal analysis-scheme such as $\mathbf{H}_\mathbf{I}\mathbf{R}_{\mathrm{diag}}$ is used. This is already done in most NWP centres (e.g., Järvinen and Undén 1997). Let us now consider a realistic thinning of the observation system to better understand the optimality of this 'thinning' process. The observations for each configuration are extracted from a common network. The elements of the corresponding error-covariances are also extracted from a complete error covariance for all observations of this network. The analysis mesh (100 km) and the background error are the same as before. We suppose that the complete network has an observation interval of 20 km (401 observations). Thinned observations are still equally-spaced, and the thinned intervals are taken as multiples of 20 km. The observation-error variance (including the instrument and representativeness errors) is taken as one. For the observation-error correlation, we assume that it is of Gaussian-type $\exp(-r^2/2L_\mathrm{c}^2)$ with a correlation scale $L_\mathrm{c} = 100$ km. The observation operator $\mathbf{H}_\mathbf{I}$ is used in the computation. Note that no specific observation-type is assumed (it may be any observation type that can produce the assumed observation-error covariance with $\mathbf{H}_\mathbf{I}$).

Figure 9 shows the results for the optimal (modelling the observation-error correlation) and suboptimal schemes (ignoring the observation-error correlation). Similarly to

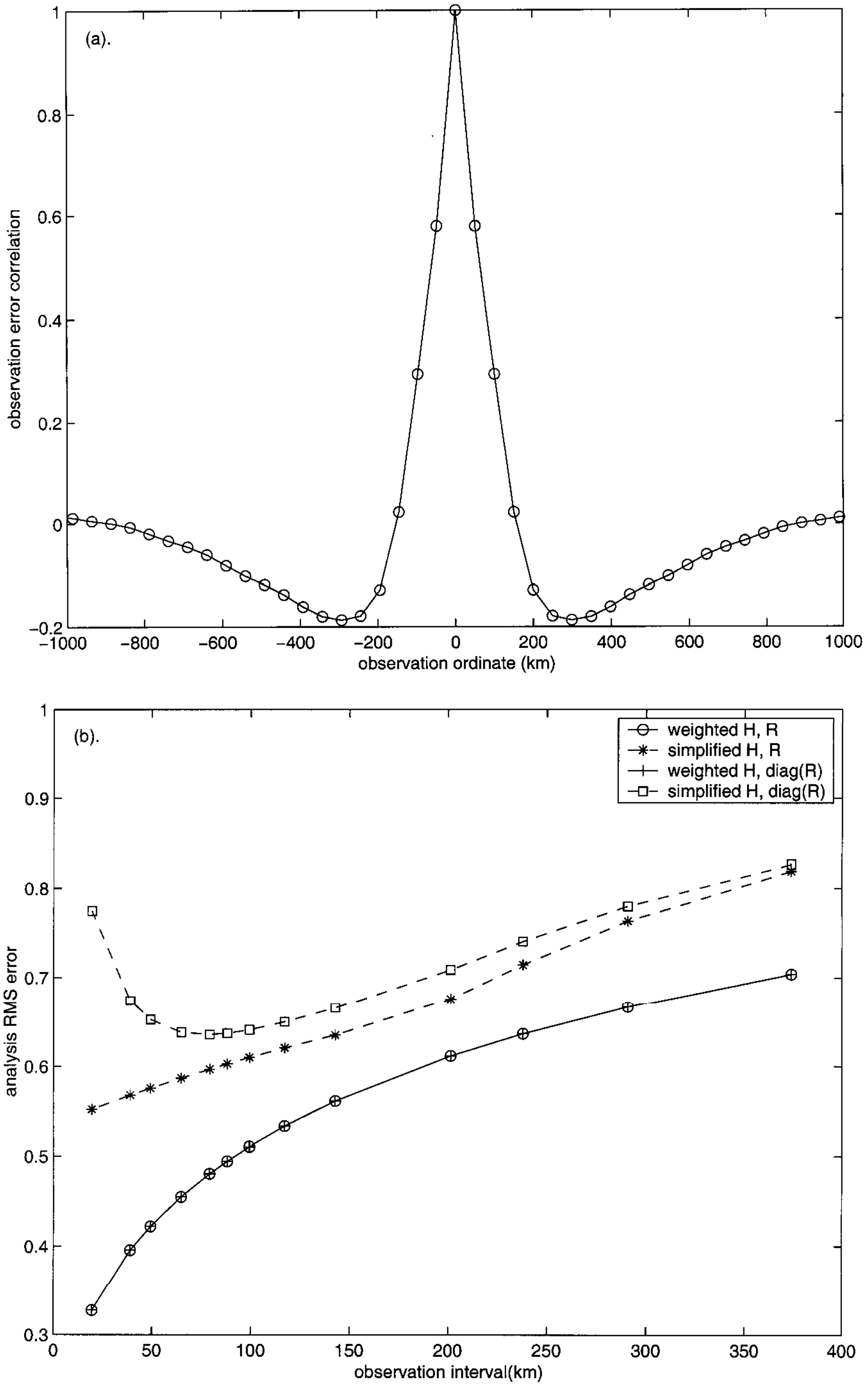

Figure 8. Errors for a uniform observation with $L_o = 400$ km and $K_m = 39$ ($\Delta r = 100$ km): (a) correlation of the observation error for the observation interval of 50 km; (b) analysis error as a function of the observation interval for the four schemes.

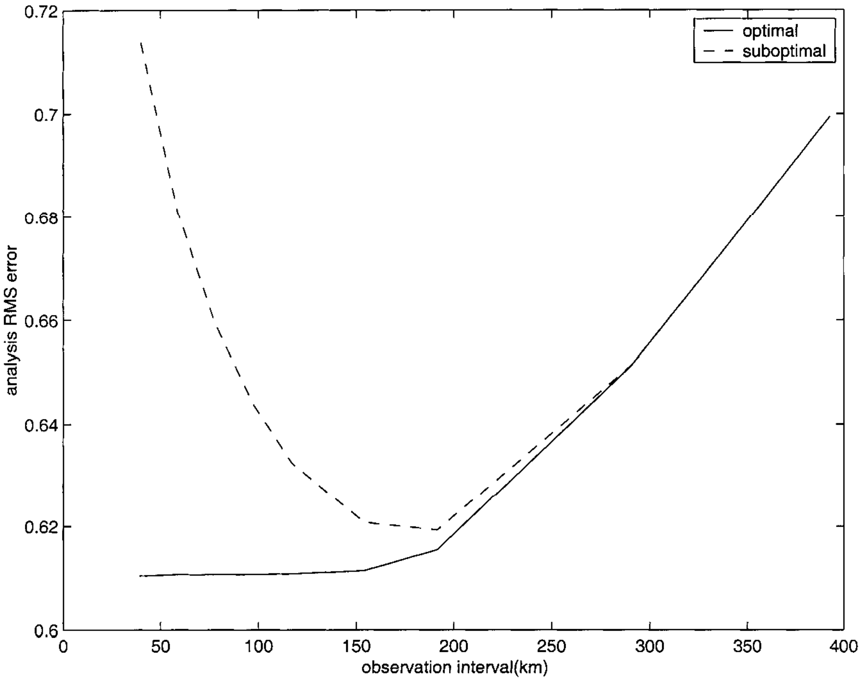

Figure 9. Thinning of observations with spatially correlated errors. The analysis mesh $\Delta r = 100$ km, the correlation length of the background error $L_b = 208$ km. The background- and observation-error variances have a value of one. The observation-error correlation is of Gaussian type with a correlation length of 100 km. The observations and their error covariances for each configuration are extracted from an equally-spaced observation network with an interval of 20 km.

Fig. 8(b), the suboptimal scheme has a optimal thinning interval (200 km), for which the analysis error is minimum. There, the best balance is reached between the advantage of increasing the number of observations and the disadvantage of strengthening their correlation. These results, and others (not shown), indicate that a small correlation-coefficient 0.15 between two adjacent measurements is allowed to reach the best balance. This critical correlation could be considered a safe threshold for the validity of the suboptimal scheme. As expected, both schemes obtain the same results on the side of heavy thinning where the error correlation is exactly negligible. Moreover, even for the optimal scheme, increasing the density of observation beyond a certain threshold value (an interval of about 150 km in this case) will yield little improvement in analysis accuracy. This contrasts with observations with spatially uncorrelated errors, when the improvement of analysis accuracy with increase of observation density is rapid (not shown). The result is consistent with that of Bergman and Bonner (1976), who used satellite temperature-soundings with spatially correlated errors in the optimal interpolation analysis. More interestingly, the optimal thinning of observations yields an analysis accuracy close to that of the optimal scheme with all data (0.62 compared to 0.61). It then seems to extract most independent information included in the complete observation-network.

4. CONCLUSIONS

The representativeness error is one of the main unknowns in present data-assimilation systems. In a simple 1-D framework, one can estimate and model it in analysis schemes, and the exact analysis-error covariance can be computed accurately. Different analysis-schemes and observation weighting-functions were used in this study.

General conclusions can be drawn about the optimal observation resolution and the choice of observation operator. For the dense observation-network (i.e., the number of observations N is greater than the number of analysis grid-points M), increasing the observation resolution (smaller L_o) is beneficial. The results suggest that an optimal observation-resolution L_o is approximately determined by $L_\mathrm{o} \approx \Delta r \frac{M}{N}$, with Δr the analysis mesh size. If this is not the case, we should design our remote-sensing instrument to have its resolution as close as possible to that of the model, or larger in some cases of very sparse observation-network. For the observation extent a few times as large as the analysis mesh size (as for meso-scale models and some satellite observations), an average-type observation operator (as $\mathbf{H_W}$), rather than an interpolation-type (as $\mathbf{H_I}$), should be used.

In general, for observations with spatially uncorrelated error, increasing the observation density always significantly improves the analysis accuracy. For observations with spatially correlated error and using the optimal analysis scheme, which models the observation-error correlation in the cost function, increasing the observation density beyond a threshold value will yield little or no improvement in analysis accuracy. Moreover, the suboptimal analysis scheme (which assimilates the data with correlated errors as if they were uncorrelated), will limit the use of high-density observations. In this case, an optimal thinning of observations, which compromises between the risks of having too small a data density and being hit by correlated observation errors, can extract most independent information included in the complete observation-network and yield an analysis accuracy close to that of the optimal scheme with all data. In our case, a threshold correlation value 0.15 between two adjacent observations is found for this optimal thinning of observations.

The current variational data-assimilation schemes tend to adopt the incremental algorithm (Courtier *et al.* 1994) which allows the analysis to be performed at a reduced resolution to hasten computation. In this case, the above remarks would be valid, except that they should be related to the resolution of the increments rather than to the one associated with the comparison of the atmospheric state to observations. Finally, it should be mentioned that this study is purely statistical, not especially targeted towards real atmospheric situations, and consequently can only provide general guide-lines. The real representativeness-error will be flow-dependent and vary with space and time. Obviously, an *in situ* measurement might be sufficient to represent the atmospheric state in a wide area around it in settled weather, but not during stormy conditions. Further studies should be performed for real situations, particularly for some 'difficult' atmospheric cases using a practical 3-D NWP model.

APPENDIX

The cardinal weighting-function given by (5) is well known in classical sampling theory (see Petersen and Middleton (1963) for an application in meteorology). For a general weighting function $w(r)$ defined in a 1-D periodical domain, we have

$$\begin{aligned} y^{\mathrm{t}}(r_{\mathrm{o}}) &= \int_{r=-\pi a}^{\pi a} x^{\mathrm{t}} w(r-r_{\mathrm{o}})\,\mathrm{d}r \\ &= \sum_{k=-\infty}^{+\infty} \hat{x}_k \left[\int_{r=-\pi a}^{\pi a} \exp(ikr/a) w(r-r_{\mathrm{o}})\,\mathrm{d}r \right] \\ &= \sum_{k=-\infty}^{+\infty} \hat{x}_k \left[\int_{s=-\pi a - r_{\mathrm{o}}}^{\pi a - r_{\mathrm{o}}} \exp\{ik(s+r_{\mathrm{o}})/a\} w(s)\,\mathrm{d}s \right] \\ &= \sum_{k=-\infty}^{+\infty} \hat{x}_k \exp(ikr_{\mathrm{o}}/a) \left[\int_{s=-\pi a - r_{\mathrm{o}}}^{\pi a - r_{\mathrm{o}}} \exp(iks/a) w(s)\,\mathrm{d}s \right]. \end{aligned} \tag{A.1}$$

Now the question is to perform the integral within square brackets. For the cardinal weighting-function (5), the spectral weighting-coefficients are following a rectangular window ($\hat{w}_k = 1$ for $|k| \leqslant K_{\mathrm{m}}$ and $\hat{w}_k = 0$ for $|k| > K_{\mathrm{m}}$). It is more convenient to perform the derivation from spectral space to physical space. Extending the limit of integration within square brackets to infinity, we have

$$\begin{aligned} &\int_{s=-\pi a - r_{\mathrm{o}}}^{\pi a - r_{\mathrm{o}}} \exp(iks/a) w(s)\,\mathrm{d}s \\ &\qquad \approx \int_{s=-\infty}^{\infty} \exp(iks/a) w(s)\,\mathrm{d}s \\ &\qquad = \hat{w}_k. \end{aligned} \tag{A.2}$$

Extending the integration to infinity allows us to use the Fourier transformation formula, namely

$$\begin{aligned} w(s) &= \frac{1}{2\pi a} \int_{k=-\infty}^{\infty} \exp(-iks/a) \hat{w}_k\,\mathrm{d}k \\ &= \frac{1}{2\pi a} \int_{k=-K_{\mathrm{m}}}^{K_{\mathrm{m}}} \exp(-iks/a)\,\mathrm{d}k \\ &= -\frac{a}{2\pi a i s} \exp(-iks/a)\Big|_{k=-K_{\mathrm{m}}}^{K_{\mathrm{m}}} \\ &= \frac{\sin(K_{\mathrm{m}} s/a)}{\pi s}. \end{aligned} \tag{A.3}$$

This is the cardinal weighting-function (5). Integrating numerically

$$\int_{s=-\pi a}^{\pi a} \exp(iks/a) \frac{\sin(K_{\mathrm{m}} s/a)}{\pi s}\,\mathrm{d}s$$

for the given K_{m} and k, we indeed obtain a rectangular window in spectral space. The formula is approximate (because of the extension of the limits of integration). However, this does not affect our results for representativeness errors and analysis errors since this cardinal weighting-function was not used explicitly in the computation. The essential motivation of giving this cardinal weighting-function is to provide a physical interpretation of the analysis mesh values for this 1-D spectral ‘model’. This allows us to compare explicitly the representation of model and observations to true signal in physical space (Fig. 1(a)). Equations (10) and (11) can also be obtained by integrating the corresponding weighting-functions.

ACKNOWLEDGEMENTS

The authors want to thank Jean-Noël Thépaut, Gérald Desroziers and Olivier Talagrand for fruitful discussions. Two anonymous referees are gratefully acknowledged for some useful comments on an earlier version of the manuscript.

REFERENCES

Andersson, E., Haseler, J., Undén, P., Courtier, P., Kelly, G., Vasiljević, D., Branković, C., Cardinali, C., Gaffard, C., Hollingsworth, A., Jakob, C., Janssen, P., Klinker, E., Lanzinger, A., Miller, M., Rabier, F., Simmons, A., Strauss, B., Thépaut, J.-N. and Viterbo, P. 1998 The ECMWF implementation of three-dimensional variation assimilation (3D-Var). III: Experimental results. *Q. J. R. Meteorol. Soc.*, **124,** 1831–1860

Bergman, K. H. and Bonner, W. D. 1976 Analysis error as a function of observation density for satellite temperature soundings with spatially correlated errors. *Mon. Weather Rev.*, **104,** 1308–1316

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

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

Ghil, M. and Malanotte-Rizzoli, P. 1991 Data assimilation in meteorology and oceanography. *Advances in geophysics*, **33,** 141–266

Järvinen, H. and Undén, P. 1997 'Observation screening and first guess quality control in the ECMWF 3D-VAR data assimilation system'. *ECMWF Tech. Memo. 236.*, available from the ECMWF, Reading, UK

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

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

Lorenc, A. C., Ballard, S. P., Bell, R. S., Ingleby, N. B., Andrews, P. L. F., Barker, D. M., Bray, J. R., Clayton, A. M., Dalby, T., Li, D., Payne, T. J. and Saunders, F. W. 2000 The Met. Office global three-dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc.*, **126,** 2991–3012

Miller, M. 1999 'Resolution studies'. *ECMWF Tech. Memo. 299*, available from the ECMWF, Reading, UK

Mitchell, H. L. and Daley, R. 1997 Discretization error and signal/error correlation in atmospheric data assimilation. Part II: The effect of unresolved scales. *Tellus*, **49A,** 54–73

Parrish, D. F. and Derber, J. C. 1992 The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Weather Rev.*, **120,** 1747–1763

Petersen, D. P. and Middleton, D. 1963 On representative observations. *Tellus*, **15,** 387–405

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

Thiébaux, H. J. 1976 Anisotropic correlation functions for objective analysis. *Mon. Weather Rev.*, **104,** 994–1002

Thiébaux, H. J., Mitchell, H. L. and Shantz, D. W. 1986 Horizontal structure of hemispheric forecast error correlations for geopotential and temperature. *Mon. Weather Rev.*, **114,** 1048–1066