Tellus (1997), 49A, 32–53
Printed in UK – all rights reserved

TELLUS
ISSN 0280–6495

Discretization error and signal/error correlation in atmospheric data assimilation

(I). All scales resolved

By HERSCHEL L. MITCHELL* and ROGER DALEY[1], *Direction de la Recherche en Météorologie, Atmospheric Environment Service, 2121 route Transcanadienne, 5e étage, Dorval, Québec, Canada H9P 1J3*

(Manuscript received 28 August 1995; in final form 5 July 1996)

ABSTRACT

Both the numerical models used in atmospheric data assimilation and the forward interpolation from the analysis mesh to the observations are subject to discretization errors. To examine the effect of these errors, a generalized Kalman filter, in which both model and observation errors are functions of the signal, is formulated. Far from the red model error spectrum assumed in many studies, the formulation yields a model error spectrum which increases with wavenumber to reach a maximum at the truncation limit. The resulting second-moment equations are studied in the context of the one-dimensional linear advection equation and, even for this simple equation, found to be quite complex. For example, it is found that signal/error correlations, not normally considered in standard Kalman filter theory, can play an important rôle. Two types of (semi-Lagrangian) model discretization and two types of forward interpolation are examined in this paper. The first type uses Fourier interpolation (and has no error), while the second type uses cubic spline interpolation (and has amplitude and phase errors). To facilitate understanding of the general case, various simpler cases are considered first, e.g., the case of a uniform observation network with the same number of observations as analysis meshpoints reveals important analogies between the forward interpolation error and the model discretization error. It is found that the perfect-model assumption can result in degeneracy for any observation network when the signal/error correlation is properly accounted for. The case of a single observation (coinciding with an analysis gridpoint) strikingly illustrates the importance of the signal/error correlations and suggests that simple model error parametrization based purely on the model discretization error, and neglecting these correlations, would seriously underestimate the forecast and analysis errors. In this paper, it is assumed that the analysis mesh can resolve all scales in the signal. The effect of unresolved scales is considered in a companion paper.

1. Introduction

It is generally recognized that improved discretization schemes for the meteorological equations and improved model resolution have been crucial to the successful development of numerical weather prediction over the past half century (Tribbia and Anthes, 1987; Ritchie et al., 1995, and references therein). Although discretization and resolution are also important for data assimilation, their effect is not explicitly accounted for by current assimilation techniques. In fact, the impact on data assimilation of discretization error and unresolved scales has never been carefully examined.

* Corresponding author.
[1] Current affiliation: Naval Research Laboratory, Monterey, California, USA.

Kalman filter (KF) theory provides a powerful framework not only for developing algorithms for atmospheric data assimilation (Cohn and Parrish, 1991; Cohn et al., 1994; Miller et al., 1994) but also for examining data assimilation problems (Daley 1992a,c). Unlike variational (adjoint) methods, which assume that the assimilating model is perfect, KF theory tries to account explicitly for model error (Ménard and Daley, 1996). In this paper, the KF framework will be used to study the effect of discretization and resolution on atmospheric data assimilation. As standard KF theory is not sufficiently general for our purposes, we will derive and use a more general set of second-moment equations.

In standard KF theory, estimates of the observation error and the model error (or more precisely, their covariance matrices, $\boldsymbol{R}$ and $\boldsymbol{Q}$, respectively) must be specified. In the particular case of atmospheric data assimilation, both types of error can be decomposed into a sum of several subcomponents. Observation error can be decomposed into instrument error and forward interpolation error (Daley, 1993), and the latter of these is subject to the effects of discretization error and limited resolution. Discretization error and limited resolution are also important subcomponents of model error, along with possibly incorrect boundary conditions or inadequate physical parameterizations, for example.

In previous atmospheric data assimilation studies employing Kalman filters (or smoothers), simplifying assumptions have generally been made to facilitate the specification of $\boldsymbol{R}$ and $\boldsymbol{Q}$, (see the above references, and Dee 1991; Ménard 1994; Todling and Ghil 1994). With regard to the observation errors, many studies have assumed that the observed and analyzed variables are the same and that the observation locations coincide with points of the analysis mesh. In that case, the forward interpolation error component of $\boldsymbol{R}$ vanishes, leaving only the externally specified instrument error component. The specification of the model error has been more of a problem, but it has generally been assumed that $\boldsymbol{Q}$ is homogeneous, is similar to the forecast error covariance used in statistical interpolation procedures and, following Phillips (1986), can be obtained from a diagonal spectral matrix, $\hat{\boldsymbol{Q}}$. Following Bennett and Budgell (1987), the elements of $\hat{\boldsymbol{Q}}$ are generally taken to represent a red spectrum, i.e., to decrease as the wavenumber increases.

The forward interpolation error component of $\boldsymbol{R}$, affected as it is by discretization error and limited resolution, will be explicitly retained in the present study. Also, our interest in discretization error, which generally affects the shortest scales most severely, implies that a red spectrum for the model error is inappropriate for our purposes. We will therefore derive an explicit expression for the model error. This expression yields a spectrum consistent with the well-known fact that the spectrum for discretization error increases with wavenumber to reach a maximum at the truncation limit.

Another assumption of standard KF theory is that the observation and model errors are not serially correlated and not correlated with each other. However, if one considers, for example, a thermometer that tends to read low when it is warm, a discretization scheme that poorly handles short scales or a forecast model that tends to underdevelop East Coast lows in winter (due to a deficient convection scheme), it becomes clear that observation and model errors can depend on the atmospheric state, or signal. It follows then that, since the signal itself is serially correlated, these errors are also serially correlated and correlated with each other, as pointed out by Daley (1992b).

In the next section, a general set of second-moment equations are derived. The approach differs from that of the standard KF in that the model and observation errors are not specified externally. Instead, a "true" state or signal (assumed to evolve according to a simple linear advection equation) is postulated, and an assimilating model, $\boldsymbol{M}$, and forward interpolation operator, $\boldsymbol{H}$, are defined. The model discretization and forward interpolation errors, which appear in the second-moment equations, are defined in terms of $\boldsymbol{M}$, $\boldsymbol{H}$ and the signal. Since these errors are defined in terms of the signal, which is itself serially correlated, these errors are also serially correlated and correlated with each other, and the resulting second-moment equations are more complicated than the standard KF equations.

The effect of these signal/error correlations will be examined by numerical experimentation for different models, $\boldsymbol{M}$, and different forward interpolation operators, $\boldsymbol{H}$. In addition, the effect of different truths will be examined, by varying the

Courant number, as will the effect of various types of observation networks. It will be seen that the signal/error correlations can play an important role.

In the present paper, the analysis mesh is assumed to be able to resolve all scales in the signal; the focus is on the forward interpolation error and model discretization error and the effect on the analysis error of their being correlated with the signal. The more general case, where not all scales in the signal are resolvable by the analysis mesh, will be considered in a companion paper (Mitchell and Daley 1997; hereinafter denoted Part II).

2. Theory

Consider the one-dimensional periodic domain $-\pi a \leqslant x \leqslant \pi a$ and the dependent variable $h(x, t)$ representing the geopotential height field. We define the "true" state or signal $h(x, t)$ as a Fourier series, truncated at wavenumber K, which we write in real form

$$h(x,t)=\frac{1}{2}\hat{h}_0^{\mathrm{c}}(t) + \sum_{k=1}^{K}\left[\hat{h}_k^{\mathrm{c}}(t)\cos(kx/a)+\hat{h}_k^{\mathrm{s}}(t)\sin(kx/a)\right], \tag{2.1}$$

where k is the wavenumber. At time t_n, we write these Fourier coefficients as a vector of length $2K+1$,

$$\hat{\boldsymbol{h}}_n^{\mathrm{T}}=[\hat{h}_0^{\mathrm{c}}(t_n)\ \hat{h}_1^{\mathrm{c}}(t_n)\ \hat{h}_1^{\mathrm{s}}(t_n)\ \ldots\ \hat{h}_K^{\mathrm{s}}(t_n)] \tag{2.2}$$

where the superscript "T" indicates transpose.

Define a fixed grid x_j, $1 \leqslant j \leqslant J$ which is symmetric about $x=0$, with J odd and specify that $x=0$ is a gridpoint. It is assumed that the mesh is uniform, that is,

$$x_j=-\pi a+\Delta x_J/2+(j-1)\Delta x_J, \tag{2.3}$$

where $\Delta x_J=2\pi a/J$. *In this paper we assume that the analysis mesh can resolve all scales in the signal, i.e.,* $J=2K+1$. The case of unresolved scales, i.e. $J<2K+1$, will be considered in Part II. Since the equations developed here will also serve in Part II, the notation in the present paper distinguishes between J, the number of points in the analysis mesh, and $2K+1$, the number of components in the signal.

For $x=x_j$, $1 \leqslant j \leqslant J$ and $t=t_n$, we write (2.1) in matrix form as,

$$\bar{\boldsymbol{h}}_n=\boldsymbol{F}\hat{\boldsymbol{h}}_n, \tag{2.4}$$

where $\hat{\boldsymbol{h}}_n$ is defined in (2.2), $\boldsymbol{F}$ is the $J\times(2K+1)$ Fourier transform matrix implied by (2.1) and $\bar{\boldsymbol{h}}_n$ is the vector of length J of values at the gridpoints x_j. The overbar is used to indicate that these are the values at the points of the analysis grid.

The signal is assumed to satisfy the one-dimensional linear advection equation,

$$\frac{\partial h}{\partial t}+U\frac{\partial h}{\partial x}=0, \tag{2.5}$$

where U is a (constant) advecting velocity. We can write the solution to (2.5) at time $t_{n+1}=t_n+\Delta t$ as:

$$h(x, t_{n+1})=h(x-U\Delta t, t_n), \tag{2.6}$$

where $h(-\pi a, t)=h(\pi a, t)$. In spectral form,

$$\hat{\boldsymbol{h}}_{n+1}=\hat{\boldsymbol{M}}\hat{\boldsymbol{h}}_n, \tag{2.7}$$

where $\hat{\boldsymbol{h}}_n$ is defined in (2.2) and $\hat{\boldsymbol{M}}$ is the $(2K+1)\times(2K+1)$ block diagonal matrix with upper left-hand corner element equal to 1 and the remaining 2×2 blocks,

$$\begin{bmatrix}\cos[(kU\Delta t)/a] & -\sin[(kU\Delta t)/a]\\ \sin[(kU\Delta t)/a] & \cos[(kU\Delta t)/a]\end{bmatrix}, \tag{2.8}$$

for $1 \leqslant k \leqslant K$. Here Δt is the timestep. Note that $\hat{\boldsymbol{M}}$ is an orthogonal matrix, i.e.,

$$\hat{\boldsymbol{M}}\hat{\boldsymbol{M}}^{\mathrm{T}}=\boldsymbol{I}, \tag{2.9}$$

where $\boldsymbol{I}$ is the identity matrix. For the signal, we can define an effective gridlength $\Delta x_{\mathrm{s}}=2\pi a/(2K+1)$ and an effective Courant number $C_{\mathrm{s}}=U\Delta t/\Delta x_{\mathrm{s}}=U\Delta t(2K+1)/2\pi a$.

Now consider data assimilation with an assimilating model whose mesh is given by (2.3). Define $\bar{\boldsymbol{h}}_n^{\mathrm{a}}$ and $\bar{\boldsymbol{h}}_n^{\mathrm{f}}$ as the analyzed and forecast values of h on the grid x_j at time t_n. The assimilating model on this mesh will be a discrete approximation to (2.5) and is a $J\times J$ matrix which we denote $\boldsymbol{M}$. Then, the forecast values at time t_{n+1} are obtained from the analyzed values at time t_n by,

$$\bar{\boldsymbol{h}}_{n+1}^{\mathrm{f}}=\boldsymbol{M}\bar{\boldsymbol{h}}_n^{\mathrm{a}}. \tag{2.10}$$

We will define the discrete model $\boldsymbol{M}$ in Section 3.

Our interest is in the analysis and forecast errors. Left multiply (2.7) by $\boldsymbol{F}$ and subtract from (2.10). The result is,

$$\bar{\boldsymbol{e}}_{n+1}^{\mathrm{f}}=\boldsymbol{M}\bar{\boldsymbol{e}}_n^{\mathrm{a}}+[\boldsymbol{MF}-\boldsymbol{F}\hat{\boldsymbol{M}}]\hat{\boldsymbol{h}}_n, \tag{2.11}$$

where $\bar{e}_n^f = \bar{h}_n^f - F\hat{h}_n$ and $\bar{e}_n^a = \bar{h}_n^a - F\hat{h}_n$ are the forecast and analysis errors on the grid. The second right-hand-side term of (2.11) is the discretization error of model M (Cohn and Dee, 1988). In KF theory, this would be referred to as the model error, and it is obviously correlated with the signal in this case. We note from (2.7) that the signal is serially correlated, which means that the model error must also be serially correlated. Serially correlated model error was discussed in Daley (1992b).

Now right multiply (2.11) by its transpose and apply the expectation operator. The result is,

$$\bar{P}_{n+1}^f = M\bar{P}_n^a M^T + Q_n + MX_n[MF - F\hat{M}]^T + [MF - F\hat{M}]X_n^T M^T, \quad (2.12)$$

where $\bar{P}_n^f = \langle \bar{e}_n^f (\bar{e}_n^f)^T \rangle$ and $\bar{P}_n^a = \langle \bar{e}_n^a (\bar{e}_n^a)^T \rangle$ are the forecast and analysis error covariances on the grid, $Q_n = [MF - F\hat{M}]\hat{S}_n[MF - F\hat{M}]^T$ is the (symmetric) model error covariance due to model discretization and $\hat{S}_n = \langle \hat{h}_n \hat{h}_n^T \rangle$ is the signal covariance. $X_n = \langle \bar{e}_n^a \hat{h}_n^T \rangle$ is the covariance between the grid point values of the analysis error and the spectral form of the signal and $X_n[MF - F\hat{M}]^T$ is the covariance between the analysis error and the model discretization error. We assume that the signal covariance in real space is homogeneous and hence $\hat{S}_n$ is diagonal. Then, (2.7) and (2.9) imply that $\hat{S}_n = \hat{S}$ is stationary and therefore also $Q_n = Q$. As we shall see below [eq. (2.16)], the analysis error is in general correlated with the signal, and consequently we cannot ignore the crosscovariance terms involving X_n.

Thus we require predictive equations for X_n. Right multiply (2.11) by the transpose of (2.7) and apply the expectation operator. The result is,

$$Y_{n+1} = \langle \bar{e}_{n+1}^f \hat{h}_{n+1}^T \rangle = MX_n\hat{M}^T + [MF - F\hat{M}]\hat{S}\hat{M}^T, \quad (2.13)$$

where Y_n is the covariance between the forecast error on the grid and the (spectral) signal.

Consider now a time-independent observation network x_i, $1 \leqslant i \leqslant I$, and a column vector of length I of observations h_n^o on this network at time t_n. We will assume that these observations are in situ *point* observations, i.e., radiosondes. Expressions for the analyzed values on the grid x_j of the state variable h at time t_n can be written,

$$\bar{h}_n^a = \bar{h}_n^f + K_n[h_n^o - H\bar{h}_n^f], \quad (2.14)$$

where H is an $I \times J$ time-independent forward interpolation matrix and K_n is a $J \times I$ weight or gain matrix. We will describe the interpolation matrix H in the next section.

Before proceeding, let us define the instrument errors at the observation stations x_i, $1 \leqslant i \leqslant I$ as follows,

$$r_n = h_n^o - G\hat{h}_n, \quad (2.15)$$

where r_n is a vector of length I of instrument errors and $\hat{h}_n$ is the vector of Fourier components of the signal defined in (2.2). G is an $I \times (2K+1)$ Fourier transform matrix between the Fourier coefficients and the observation locations. G is defined as a sum over all K wavenumbers using (2.1) where x is an observation location. We will assume that r_n is stationary, unbiased and neither spatially nor serially correlated and thus r_n is only correlated with itself.

The relation between the analysis and forecast error on the grid can be obtained by subtracting $F\hat{h}_n$ from both sides of (2.14). The result is,

$$\bar{e}_n^a = [I - K_nH]\bar{e}_n^f + K_nr_n + K_n[G - HF]\hat{h}_n. \quad (2.16)$$

The term $[G - HF]\hat{h}_n$ is the forward interpolation error between the analysis grid and the observation network discussed by Lorenc (1986). We note that the forward interpolation error, like the model discretization error of (2.11), is a function of the signal in this case.

Now right multiply (2.16) by its transpose and apply the expectation operator. The result is,

$$\bar{P}_n^a = [I - K_nH]\bar{P}_n^f[I - K_nH]^T + K_n\{R^i + R^H\}K_n^T + [I - K_nH]Y_n[G - HF]^TK_n^T + K_n[G - HF]Y_n^T[I - K_nH]^T, \quad (2.17)$$

where $\bar{P}_n^a$ and $\bar{P}_n^f$ are the analysis and forecast error covariances on the analysis grid (eq. 2.12). $R^i = \langle r_n r_n^T \rangle$ is the *instrument error covariance* which is time independent. R^i is also diagonal because r_n is not spatially correlated. $R^H = [G - HF]\hat{S}[G - HF]^T$ is the (symmetric) *forward interpolation error covariance* discussed by Daley (1993). $R^i + R^H$ is the *observation error covariance.* Y_n is the covariance between forecast error on the grid and (spectral) signal defined in (2.13) and $Y_n[G - HF]^T$ is the covariance between the forecast error and the forward interpolation error. From (2.11), the forecast error is correlated with

the signal, which is itself serially correlated and consequently we cannot ignore the cross-covariance terms involving $\boldsymbol{Y}_n$.

We can also obtain a relation between $\boldsymbol{X}_n$ and $\boldsymbol{Y}_n$ by right multiplying (2.16) by $\hat{\boldsymbol{h}}_n^{\mathrm{T}}$ and applying the expectation operator, giving,

$$\boldsymbol{X}_n=[\boldsymbol{I}-\boldsymbol{K}_n\boldsymbol{H}]\boldsymbol{Y}_n+\boldsymbol{K}_n[\boldsymbol{G}-\boldsymbol{H}\boldsymbol{F}]\hat{\boldsymbol{S}}. \tag{2.18}$$

We consider three choices for the gain matrix $\boldsymbol{K}_n$. The first choice,

$$\begin{aligned}\boldsymbol{K}_n=&[\bar{\boldsymbol{P}}_n^{\mathrm{f}}\boldsymbol{H}^{\mathrm{T}}-\boldsymbol{Y}_n(\boldsymbol{G}-\boldsymbol{H}\boldsymbol{F})^{\mathrm{T}}]\\&\times[\boldsymbol{H}\bar{\boldsymbol{P}}_n^{\mathrm{f}}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}^{\mathrm{i}}+\boldsymbol{R}^{\mathrm{H}}-\boldsymbol{H}\boldsymbol{Y}_n(\boldsymbol{G}-\boldsymbol{H}\boldsymbol{F})^{\mathrm{T}}\\&-(\boldsymbol{G}-\boldsymbol{H}\boldsymbol{F})\boldsymbol{Y}_n^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}]^{-1},\end{aligned} \tag{2.19}$$

is a generalization of the standard KF gain matrix. For given $\bar{\boldsymbol{P}}_n^{\mathrm{f}}$, $\boldsymbol{Y}_n$, $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{H}$, and $\boldsymbol{R}^{\mathrm{i}}$, it minimizes the trace of (2.17). (This result can be obtained by generalizing the derivation of the standard KF (see e.g. Cohn, 1982, pp. 30–33) and requires that the $I\times I$ matrix whose inverse appears in (2.19) be positive definite.) This gain matrix would be unrealistic in an operational setting as it requires knowledge of the covariance between the forecast error and the signal. Consequently, we also define a more conventional gain matrix $\boldsymbol{K}_n$

$$\boldsymbol{K}_n=\bar{\boldsymbol{P}}_n^{\mathrm{f}}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\bar{\boldsymbol{P}}_n^{\mathrm{f}}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}^{\mathrm{i}}+\boldsymbol{R}^{\mathrm{H}}]^{-1}. \tag{2.20}$$

This choice of weight matrix is similar to that used in the standard KF and is a generalization of the weights used in optimum interpolation. Note that the matrix which is inverted in (2.20) contains both instrument and forward interpolation error covariances. The third choice for the gain matrix $\boldsymbol{K}_n$ assumes that there is no prior estimate (i.e., no forecast is available). In this case, for our time-independent observation network

$$\boldsymbol{K}_n=\boldsymbol{K}=[\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}]^{-1}\boldsymbol{H}^{\mathrm{T}}, \tag{2.21}$$

if $I\geqslant J$ and $\boldsymbol{H}$ has full rank, i.e., rank J (Strang, 1988, Section 3.3). Note that these weights also do not depend on the observation error statistics, only on the observation locations and the properties of the forward interpolation matrix $\boldsymbol{H}$. These will be specified in the next section.

Eqs. (2.10) and (2.14) can be integrated to predict the values of $\boldsymbol{h}$ at the gridpoints. The analysis weights required in (2.14) are obtained from (2.19), (2.20) or (2.21), which in general requires $\bar{\boldsymbol{P}}_n^{\mathrm{f}}$ and $\boldsymbol{Y}_n$. These are obtained by integrating the second-moment equations (2.12), (2.13), (2.17) and (2.18). From these equations we also obtain $\bar{\boldsymbol{P}}_n^{\mathrm{a}}$, the analysis error covariance on the grid of the assimilating model. $\bar{\boldsymbol{P}}_n^{\mathrm{a}}$ and $\bar{\boldsymbol{P}}_n^{\mathrm{f}}$ will be examined to determine the effect of forward interpolation error and model discretization error on the assimilation process.

To obtain a continuous representation of the analysis and forecast errors and to facilitate the examination of the unresolved errors case in Part II, we introduce another Fourier transform. Consider an arbitrary vector $\bar{\boldsymbol{z}}_n$ of elements $\bar{z}(x_j, t_n)$, $1\leqslant j\leqslant J$. Consistent with (2.1), we define a discrete Fourier transform by

$$\hat{z}_k^{\mathrm{c}}(t)=\frac{2}{J}\sum_{j=1}^{J}\bar{z}(x_j, t)\cos(kx_j/a), \tag{2.22}$$

$$\hat{z}_k^{\mathrm{s}}(t)=\frac{2}{J}\sum_{j=1}^{J}\bar{z}(x_j, t)\sin(kx_j/a).$$

The first relation is valid for $k\geqslant 0$ and the second for $k\geqslant 1$. In matrix form, we can write (2.22) as,

$$\hat{\boldsymbol{z}}_n=\tilde{\boldsymbol{F}}\bar{\boldsymbol{z}}_n, \tag{2.23}$$

where $\hat{\boldsymbol{z}}_n^{\mathrm{T}}=[\hat{z}_0^{\mathrm{c}}(t_n)\ \hat{z}_1^{\mathrm{c}}(t_n)\ \hat{z}_1^{\mathrm{s}}(t_n)\ \ldots\ \hat{z}_K^{\mathrm{s}}(t_n)]$ and $\tilde{\boldsymbol{F}}$ is the $(2K+1)\times J$ Fourier transform matrix from the grid x_j to Fourier coefficients whose elements are given in (2.22). In the present case, where the analysis mesh resolves all scales, $\boldsymbol{F}\tilde{\boldsymbol{F}}=\tilde{\boldsymbol{F}}\boldsymbol{F}=\boldsymbol{I}$.

From (2.23),

$$\hat{\boldsymbol{e}}_n^{\mathrm{a}}=\tilde{\boldsymbol{F}}\bar{\boldsymbol{e}}_n^{\mathrm{a}}\qquad\text{and}\qquad\hat{\boldsymbol{e}}_{\mathrm{n}}^{\mathrm{f}}=\tilde{\boldsymbol{F}}\bar{\boldsymbol{e}}_{\mathrm{n}}^{\mathrm{f}}, \tag{2.24}$$

where $\hat{\boldsymbol{e}}_n^{\mathrm{a}}=\tilde{\boldsymbol{F}}\bar{\boldsymbol{h}}_n^{\mathrm{a}}-\hat{\boldsymbol{h}}_n$ and $\hat{\boldsymbol{e}}_n^{\mathrm{f}}=\tilde{\boldsymbol{F}}\bar{\boldsymbol{h}}_n^{\mathrm{f}}-\hat{\boldsymbol{h}}_n$ are the (spectral) analysis and forecast error, respectively. Right multiplying each equation in (2.24) by its transpose and applying the expectation operator, yields

$$\hat{\boldsymbol{P}}_n^{\mathrm{a}}=\tilde{\boldsymbol{F}}\bar{\boldsymbol{P}}_n^{\mathrm{a}}\tilde{\boldsymbol{F}}^{\mathrm{T}}\qquad\text{and}\qquad\hat{\boldsymbol{P}}_n^{\mathrm{f}}=\tilde{\boldsymbol{F}}\bar{\boldsymbol{P}}_n^{\mathrm{f}}\tilde{\boldsymbol{F}}^{\mathrm{T}}, \tag{2.25}$$

where $\hat{\boldsymbol{P}}_n^{\mathrm{a}}=\langle\hat{\boldsymbol{e}}_n^{\mathrm{a}}(\hat{\boldsymbol{e}}_n^{\mathrm{a}})^{\mathrm{T}}\rangle$ and $\hat{\boldsymbol{P}}_n^{\mathrm{f}}=\langle\hat{\boldsymbol{e}}_n^{\mathrm{f}}(\hat{\boldsymbol{e}}_n^{\mathrm{f}})^{\mathrm{T}}\rangle$ are the spectral analysis and forecast error covariance matrices, respectively. It might be noted that the forecast error from (2.24) can be inserted into (2.11) to yield,

$$\hat{\boldsymbol{e}}_{n+1}^{\mathrm{f}}=\tilde{\boldsymbol{F}}\boldsymbol{M}\bar{\boldsymbol{e}}_n^{\mathrm{a}}+[\tilde{\boldsymbol{F}}\boldsymbol{M}\boldsymbol{F}-\hat{\boldsymbol{M}}]\hat{\boldsymbol{h}}_n. \tag{2.26}$$

The last term of (2.26) can be interpreted as the spectral form of the model error.

$\hat{\boldsymbol{P}}_n^{\mathrm{a}}$ and $\hat{\boldsymbol{P}}_n^{\mathrm{f}}$ are $(2K+1)\times(2K+1)$ matrices. They can be transformed to spatial form on any grid as described in Appendix A. From (2.25), the resulting error variances coincide with the error variances $\bar{\boldsymbol{P}}_n^{\mathrm{a}}$ and $\bar{\boldsymbol{P}}_n^{\mathrm{f}}$ at the points of the analysis mesh.

Eqs. (2.12), (2.13), (2.17), (2.18) and (2.25) plus an equation for the gain matrix or weights constitute a complete set of second-moment equations. It might be noted that if (2.13) and (2.18) are omitted and the cross-terms are dropped from (2.12) and (2.17), then the equations simplify to the standard KF equations (Cohn, 1982; Daley, 1992a; Ménard, 1994). However, as we shall demonstrate, the standard KF equations are *not* a good approximation to the full set of equations, because the effect of the signal/error correlations is far from negligible.

We now define the characteristics of the signal and describe the assimilation system.

3. The signal and the assimilation system

3.1. Signal covariance

The signal covariance matrix $\hat{\boldsymbol{S}}$, introduced in (2.12), is a $(2K+1)\times(2K+1)$ diagonal matrix whose elements are spectral variances. This implies that the signal covariance is homogeneous. The elements of $\hat{\boldsymbol{S}}$ are given by $\langle h^2\rangle \hat{g}(k)$ with

$$\hat{g}(k)=\frac{c}{[1+l^2(b+k)^2]\,[1+l^2(b-k)^2]}, \tag{3.1}$$

where b and l are constants to be specified and $\langle h^2\rangle$ is the (constant) height variance of the signal. The normalization constant, c, ensures that the integral of $\hat{g}(k)$ with respect to k is equal to one. $\hat{g}(k)$ is the Fourier transform (on a continuous, infinite domain) of the correlation function,

$$\rho(x)=\left[\cos(b|x|)+\frac{\sin(b|x|)}{lb}\right]\exp(-|x|/l). \tag{3.2}$$

$\rho(x)$ is the second-order autoregressive function (Thiébaux, 1976). We set $b=4/a$ and $l=a/3$, as in Daley (1993). $\langle h^2\rangle$ is specified so that the total signal variance at any point in physical space is 10 000 m^2. In Fig. 1a, we plot the spectrum of the signal for these choices (solid line). In Fig. 1b is the corresponding covariance as a function of displacement, x. Note that for the assimilating models and forward interpolation operators defined below, it follows from the expressions for the model discretization and forward interpolation errors in (2.11) and (2.16), that the results are insensitive to the specification of the signal variance for $k=0$.

3.2. Assimilating models

Two assimilating models $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ are considered for the grid x_j, $1\leqslant j\leqslant J$. The first model is the exact Fourier model described in Daley (1992a). This model, which is unconditionally stable, is defined by,

$$\boldsymbol{M}_1=\boldsymbol{F}\hat{\boldsymbol{M}}\tilde{\boldsymbol{F}}, \tag{3.3}$$

where $\boldsymbol{F}$ and $\tilde{\boldsymbol{F}}$ are given by (2.4) and (2.23) respectively. For this model, the matrix $[\boldsymbol{MF}-\boldsymbol{F}\hat{\boldsymbol{M}}]=\boldsymbol{F}\hat{\boldsymbol{M}}(\tilde{\boldsymbol{F}}\boldsymbol{F}-\boldsymbol{I})$ is identically zero and thus *this model is perfect in all scales.*

An imperfect, and therefore more realistic, model $\boldsymbol{M}_2$ is now defined. The solution (2.6) of the linear advection equation (2.5) suggests a semi-Lagrangian time marching algorithm. The point of departure $(x-U\Delta t)$ is always known exactly for this equation, so all that is necessary is to interpolate the values of $h(x-U\Delta t)$ from the values of h at the gridpoints of the mesh. This interpolation is performed using periodic cubic splines (Ahlberg et al, 1967, pp. 9–12), which pass through the gridpoints of the mesh. We note that most numerical algorithms for calculating advection terms assume that the gridpoints represent actual point values of the dependent variables. An expression for $\boldsymbol{M}_2$, analogous to (3.3), is

$$\boldsymbol{M}_2=\boldsymbol{F}\hat{\boldsymbol{M}}\hat{\boldsymbol{D}}\tilde{\boldsymbol{F}}=\boldsymbol{F}\hat{\boldsymbol{D}}\hat{\boldsymbol{M}}\tilde{\boldsymbol{F}}, \tag{3.4}$$

where the matrix $\hat{\boldsymbol{D}}$ is defined in Appendix B. $\hat{\boldsymbol{D}}$ accounts for the damping and phase shift due to the cubic spline interpolation. We note that the model $\boldsymbol{M}_2$ is unconditionally stable for any Courant number.

The Courant number for the models $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ is defined as $C_{\mathrm{M}}=U\Delta t/\Delta x_J=U\Delta tJ/2\pi a$. (When $J=2K+1$ as in the present paper, $C_{\mathrm{M}}=C_{\mathrm{s}}$.) When the Courant number has been chosen to be an integer p, $\hat{\boldsymbol{D}}=\boldsymbol{I}$ and $\boldsymbol{M}_1=\boldsymbol{M}_2$ takes on a particularly simple form. For example, when $p=1$, the only non-zero elements are the elements of the first sub-diagonal and the upper-right-hand corner element, which are all equal to one. When the point of departure is *not* a gridpoint (i.e., C_{M} is not an integer) then the cubic spline method slightly damps and phase shifts the smaller spatial scales (Appendix B). This differs from the Fourier method in which there is no damping for any Courant number.

The matrices $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ are circulant (Cohn and Dee, 1988). As noted by Daley (1992a), the

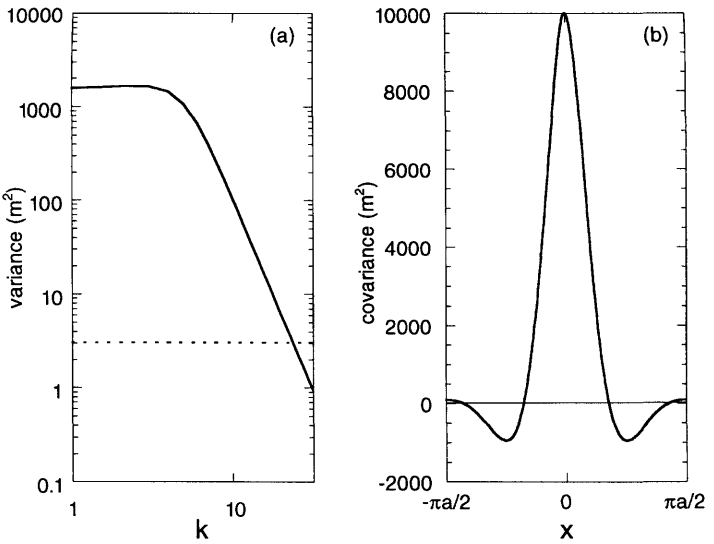

Fig. 1. (a) The spectra of the signal (solid line) and the instrument error (dashed line). (b) The covariance of the signal as a function of displacement, x.

matrix $\boldsymbol{M}_1$ has the properties (i) $\boldsymbol{M}_1\boldsymbol{M}_1^{\mathrm{T}}=\boldsymbol{I}$, and (ii) for any homogeneous covariance matrix $\boldsymbol{B}$, $\boldsymbol{M}_1\boldsymbol{B}\boldsymbol{M}_1^{\mathrm{T}}=\boldsymbol{B}$. The matrix $\boldsymbol{M}_2$ only satisfies these two properties when $\boldsymbol{M}_2=\boldsymbol{M}_1$, i.e., in the case of an integer Courant number. The model $\boldsymbol{M}_2$ cannot satisfy these properties more generally due to its damping of the smaller spatial scales.

3.3. Model error covariance

Substitution of (3.4) into the definition of $\boldsymbol{Q}$, given after (2.12), yields

$$\boldsymbol{Q}=\boldsymbol{F}[\boldsymbol{I}-\hat{\boldsymbol{D}}]\hat{\boldsymbol{S}}[\boldsymbol{I}-\hat{\boldsymbol{D}}]^{\mathrm{T}}\boldsymbol{F}^{\mathrm{T}}. \tag{3.5}$$

The derivation uses (2.9) and the fact that, since $\hat{\boldsymbol{D}}$ is block diagonal with 2×2 blocks given by (B.2),

$$[\boldsymbol{I}-\hat{\boldsymbol{D}}]\hat{\boldsymbol{S}}[\boldsymbol{I}-\hat{\boldsymbol{D}}]^{\mathrm{T}}=\hat{\boldsymbol{S}}[\boldsymbol{I}+\hat{\boldsymbol{D}}\hat{\boldsymbol{D}}^{\mathrm{T}}-(\hat{\boldsymbol{D}}+\hat{\boldsymbol{D}}^{\mathrm{T}})], \tag{3.6}$$

which is diagonal with upper left-hand corner element followed by pairs of equal elements for each wavenumber k. Left and right multiplying (3.5) by $\tilde{\boldsymbol{F}}$ and $\tilde{\boldsymbol{F}}^{\mathrm{T}}$ yields

$$\hat{\boldsymbol{Q}}=\tilde{\boldsymbol{F}}\boldsymbol{Q}\tilde{\boldsymbol{F}}^{\mathrm{T}}=[\boldsymbol{I}-\hat{\boldsymbol{D}}]\hat{\boldsymbol{S}}[\boldsymbol{I}-\hat{\boldsymbol{D}}]^{\mathrm{T}}, \tag{3.7}$$

a simple expression for $\hat{\boldsymbol{Q}}$, the spectral form of the model discretization error. Since $\hat{\boldsymbol{Q}}$ is diagonal, the model discretization error, $\boldsymbol{Q}$, is homogeneous. Of course for $\boldsymbol{M}_1$, $\hat{\boldsymbol{D}}=\boldsymbol{I}$ and so both $\boldsymbol{Q}$ and $\hat{\boldsymbol{Q}}$ vanish.

3.4. Observation networks

We will consider three types of observation networks, (i) uniform, (ii) non-uniform but of homogeneous type, and (iii) a single observation. In some cases we consider a uniform network for which $I=J$, or the more particular network which coincides with the analysis mesh. The non-uniform, homogeneous networks are defined, following Daley (1992c), by dividing the domain into I equal intervals and placing one observation within each interval. We generally require that an observation be located within a certain distance of the centre of its interval, the precise location being determined randomly. Thus although observation networks of this type are non-uniform, the observation density is approximately constant over the domain. It is important to note that these homogeneous networks are not necessarily related to the analysis mesh. In all cases we assume that the observation network measures height only and consists entirely of radiosondes with no bias and constant instrument error variance. Thus, the instrument error covariance matrix $\boldsymbol{R}^{\mathrm{i}}$ is diagonal with all elements equal to $(E^{\mathrm{r}})^2$. We set $E^{\mathrm{r}}=10$ m.

3.5. *Forward interpolation matrices: general case*

We define two forward interpolation matrices. The first, $\boldsymbol{H}_1$, is an exact Fourier procedure, defined in Daley (1992c). Thus,

$$\boldsymbol{H}_1 = \boldsymbol{G}\tilde{\boldsymbol{F}} \tag{3.8}$$

where $\tilde{\boldsymbol{F}}$ is given by (2.23). We note that $[\boldsymbol{G}-\boldsymbol{H}_1\boldsymbol{F}]=\boldsymbol{G}[\boldsymbol{I}-\tilde{\boldsymbol{F}}\boldsymbol{F}]=0$ so that $\boldsymbol{H}_1$ *has no forward interpolation error*. The second forward interpolation matrix $\boldsymbol{H}_2$, like $\boldsymbol{M}_2$, uses periodic cubic spline interpolation.

3.6. *Forward interpolation matrices: uniform observation network*

The case of a uniform observation network with $I=J$ can be analyzed analytically and facilitates the understanding of more general cases. (Note that the observation network and analysis grid need not coincide, but if they do then $\boldsymbol{H}_1=\boldsymbol{H}_2=\boldsymbol{I}$, the identity matrix, so that $[\boldsymbol{G}-\boldsymbol{H}\boldsymbol{F}]=0$.) Moreover, we shall see that the similar structures of the observation network and the analysis mesh in this case result in similarities between the forward interpolation and model discretization errors. For such an observation network,

$$\boldsymbol{H}_1\boldsymbol{H}_1^{\mathrm{T}} = \boldsymbol{G}\tilde{\boldsymbol{F}}\tilde{\boldsymbol{F}}^{\mathrm{T}}\boldsymbol{G}^{\mathrm{T}} = \boldsymbol{I}. \tag{3.9}$$

Comparing (3.8) and (3.3), and keeping in mind (3.4) and Appendix B, by analogy $\boldsymbol{H}_2$ can be written as

$$\boldsymbol{H}_2 = \boldsymbol{G}\hat{\boldsymbol{D}}\tilde{\boldsymbol{F}}. \tag{3.10}$$

When considering this special type of network, it will be convenient to express $\boldsymbol{H}$ using (3.10), with $\hat{\boldsymbol{D}}=\boldsymbol{I}$ in the case of $\boldsymbol{H}_1$ and $\hat{\boldsymbol{D}}$ as defined in Appendix B in the case of $\boldsymbol{H}_2$. Finally, for this special network, it is convenient to define the spectral form of $\boldsymbol{R}^{\mathrm{i}}$, i.e.

$$\hat{\boldsymbol{R}}^{\mathrm{i}} = \tilde{\boldsymbol{F}}\boldsymbol{R}^{\mathrm{i}}\tilde{\boldsymbol{F}}^{\mathrm{T}}. \tag{3.11}$$

It then follows from (3.9) that, in this case,

$$\boldsymbol{R}^{\mathrm{i}} = \boldsymbol{G}\hat{\boldsymbol{R}}^{\mathrm{i}}\boldsymbol{G}^{\mathrm{T}}. \tag{3.12}$$

The constant dashed line in Fig. 1 is the spectrum of the instrument error $\hat{\boldsymbol{R}}^{\mathrm{i}}$. That all diagonal elements (except the first) of the diagonal matrix $\hat{\boldsymbol{R}}^{\mathrm{i}}$ equal $2(E^{\mathrm{r}})^2/J(=3.077\ \mathrm{m}^2)$ follows from the particular structure of $\tilde{\boldsymbol{F}}\tilde{\boldsymbol{F}}^{\mathrm{T}}$ (Appendix A, Part II).

3.7. *Specification of parameters*

In all experiments we set $K=32$. The constant "a" can easily be scaled out of the equations. (This would not be true for a multi-variate system.) Therefore, rather than specify an arbitrary value for a, we set $a=1$. U and Δt will be specified subsequently. We are interested in the stationary statistics which can be obtained by integrating the second-moment equations until stationarity is achieved.

As we saw in Section 2, the stationary statistics are the result of a complex interplay between the model discretization error of (2.12) and (2.13) and the forward interpolation error of (2.17) and (2.18). Both of these errors are correlated with the signal and thus introduce serial correlations into the system which are not properly accounted for in any assimilation system (including the standard KF and the four-dimensional variational algorithms). Since the equations of Section 2 are complex, we will seek to comprehend the system by beginning with simplified special cases and moving into progressively more complex observation networks and models. We begin with the case where there is no prior estimate.

4. No prior estimate (no model)

In this case, the second-moment equations simplify considerably. Only the analysis equations, (2.17)–(2.18), need be considered and the only source of signal/error correlation is through the forward interpolation error.

With $\boldsymbol{K}_n$ defined by (2.21), $\boldsymbol{K}_n\boldsymbol{H}=\boldsymbol{I}$, the identity matrix. Therefore as expected, $\bar{\boldsymbol{h}}_n^{\mathrm{f}}$ drops out of (2.14), i.e., the analyzed value is obtained by interpolating the observed values to the analysis points. The cross-terms in (2.17) also vanish and $\bar{\boldsymbol{P}}_n^{\mathrm{a}}$ and $\boldsymbol{X}_n$ are time-independent, from (2.17) and (2.18), and are given by:

$$\bar{\boldsymbol{P}}^{\mathrm{a}} = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{R}^{\mathrm{i}}+\boldsymbol{R}^{\mathrm{H}}][(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}]^{\mathrm{T}}, \tag{4.1}$$

$$\boldsymbol{X} = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{G}-\boldsymbol{H}\boldsymbol{F}]\hat{\boldsymbol{S}}. \tag{4.2}$$

It follows from (2.25) that $\hat{\boldsymbol{P}}^{\mathrm{a}}$ is also time-independent. Note that $\bar{\boldsymbol{P}}^{\mathrm{a}}$ and $\boldsymbol{X}$ are independent of $\bar{\boldsymbol{P}}^{\mathrm{f}}$ (as expected) and that $\bar{\boldsymbol{P}}^{\mathrm{a}}$ depends only on $\boldsymbol{H}$, the forward interpolation matrix, $\boldsymbol{R}^{\mathrm{i}}$, the instrument error covariance, and $\boldsymbol{R}^{\mathrm{H}}$, the forward inter-

polation error covariance. Since (4.1) and (4.2) do not involve $\boldsymbol{M}$ or $\hat{\boldsymbol{M}}$, all the results in this section are independent of the Courant number.

In this case, the analysis error covariance $\bar{\boldsymbol{P}}^{\mathrm{a}}$ and the signal/analysis error covariance $\boldsymbol{X}$ are independent of each other. Thus, while the analysis error and the signal may be correlated, this does not affect the analysis error, as it does in the more general system of Section 2. However, it is still of interest to calculate $\boldsymbol{X}$ in this case.

We focus on the case where $I=J$, i.e., the observation network and analysis grid have the same number of points. Then $\boldsymbol{H}$ and $\boldsymbol{H}^{\mathrm{T}}$ are square matrices and may be invertible, in which case $\boldsymbol{K}=[\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}]^{-1}\boldsymbol{H}^{\mathrm{T}}=\boldsymbol{H}^{-1}$. Left multiplication of (2.14) by $\boldsymbol{G}\tilde{\boldsymbol{F}}$ and application of (2.23) yields, for Fourier interpolation $\boldsymbol{H}=\boldsymbol{H}_1$,

$$\boldsymbol{G}\hat{\boldsymbol{h}}_n^{\mathrm{a}}=\boldsymbol{G}\tilde{\boldsymbol{F}}\boldsymbol{H}^{-1}\boldsymbol{h}_n^{\mathrm{o}}=\boldsymbol{h}_n^{\mathrm{o}}. \tag{4.3}$$

Thus for the exact forward interpolation, $\boldsymbol{H}_1$, the analysis passes through the observations. It follows from (2.15) that the analysis error, $\hat{\boldsymbol{e}}_n^{\mathrm{a}}$, evaluated at the observation locations equals the instrument error. Eq. (4.3) does not hold for $\boldsymbol{H}_2$, the cubic spline interpolation, since $\boldsymbol{G}\tilde{\boldsymbol{F}}\boldsymbol{H}_2^{-1}\neq\boldsymbol{I}$.

It is possible to obtain simple analytic expressions for (4.1)–(4.2) for the $I=J$ case when the observation network is uniform.

4.1. *Uniform observation networks with $I=J(=2K+1)$*

The general case of a uniform observation network with $I=J\leqslant 2K+1$ is considered in Appendix B of Part II. When all scales are resolved, as is the case here, considerable simplification occurs. Thus (B.2) and (B.3) of Part II reduce to

$$\hat{\boldsymbol{P}}^{\mathrm{a}}=\hat{\boldsymbol{D}}^{-1}\{\hat{\boldsymbol{R}}^{\mathrm{i}}+[\boldsymbol{I}-\hat{\boldsymbol{D}}]\hat{\boldsymbol{S}}[\boldsymbol{I}-\hat{\boldsymbol{D}}]^{\mathrm{T}}\}\hat{\boldsymbol{D}}^{-\mathrm{T}}, \tag{4.4}$$

$$\hat{\boldsymbol{X}}=\tilde{\boldsymbol{F}}\boldsymbol{X}=[\hat{\boldsymbol{D}}^{-1}-\boldsymbol{I}]\hat{\boldsymbol{S}}, \tag{4.5}$$

where $\hat{\boldsymbol{D}}^{-\mathrm{T}}$ denotes $(\hat{\boldsymbol{D}}^{-1})^{\mathrm{T}}$. Note the similarity between the forward interpolation error term in (4.4) and the expression for the model discretization error given in (3.7).

Now the matrix appearing in braces in (4.4) is diagonal and for each wavenumber k has equal sine and cosine components, while $\hat{\boldsymbol{D}}$ is block diagonal with 2×2 blocks given by (B.2). Therefore (4.4) can be decomposed into K sets of 2×2 matrix equations, one for each wavenumber k, following Daley and Ménard (1993). This yields, after matrix multiplication, a diagonal matrix with diagonal elements given by

$$a^2(k)=d^{-2}(k)\ \{r^2(k)+s^2(k)\times[1+d^2(k)-2d(k)\cos\phi(k)]\}, \tag{4.6}$$

where for a given k, $a^2(k)$, $r^2(k)$, and $s^2(k)$ denote the diagonal elements of $\hat{\boldsymbol{P}}^{\mathrm{a}}$, $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{S}}$, respectively, and $d(k)$ and $\phi(k)$ are the damping and phase shift due to the spline interpolation. Since $\hat{\boldsymbol{P}}^{\mathrm{a}}$ is diagonal, $\bar{\boldsymbol{P}}^{\mathrm{a}}$ is homogeneous and the total analysis error, obtained by transforming $\hat{\boldsymbol{P}}^{\mathrm{a}}$ to physical space as described in Appendix A, is also homogeneous.

In a similar way, (4.5) yields

$$x(k)=s^2(k)[\cos\phi(k)-d(k)]/d(k), \tag{4.7}$$

$$\xi(k)=s^2(k)\sin\phi(k)/d(k), \tag{4.8}$$

where $x(k)$ and $\xi(k)$ are the diagonal and off-diagonal elements, respectively, of the kth 2×2 block of the antisymmetric block-diagonal matrix $\hat{\boldsymbol{X}}$. When transformed to physical space (as described in Appendix A), $\hat{\boldsymbol{X}}$ becomes homogeneous. This follows from the fact that the (geometrical) relationship between an analysis point and the observation network is the same for any analysis point.

It should be noted that (4.6)–(4.8) are valid for any forward interpolation operator, it being necessary only to specify the appropriate damping, $d(k)$, and phase shift, $\phi(k)$. These eqs. will be used to examine the effect of the two interpolation operators: the Fourier operator $\boldsymbol{H}_1$, which has no forward interpolation error, and the spline operator $\boldsymbol{H}_2$, which does.

When $\boldsymbol{H}=\boldsymbol{H}_1$, $\hat{\boldsymbol{D}}=\boldsymbol{I}$ and (4.6)–(4.8) reduce to $a^2(k)=r^2(k)$ and $x(k)=\xi(k)=0$. Therefore, all diagonal elements of $\hat{\boldsymbol{P}}^{\mathrm{a}}$ (for $k>0$) equal $2(E^{\mathrm{r}})^2/J$ ($=3.077\ \mathrm{m}^2$ for the parameters specified in Section 3). It follows, from (3.12), that the homogeneous analysis error variance in physical space equals $(E^{\mathrm{r}})^2$ ($=100\ \mathrm{m}^2$). The analysis passes exactly through the observations, from (4.3). The analysis error and signal are uncorrelated, because the instrument error is assumed to be uncorrelated with the signal.

When $\boldsymbol{H}=\boldsymbol{H}_2$, then $\hat{\boldsymbol{D}}\neq\boldsymbol{I}$ (except when the uniform observation network coincides with the analysis grid) and so both the instrument error covariance and the forward interpolation error covariance terms in (4.4) and (4.6) are non-zero.

Each term represents a homogeneous contribution to the analysis error variance and since, from Fig. 10 (Appendix B), $d(k)$ and $\phi(k)$ are functions of the displacement between the observation network and the analysis mesh, each of these terms is also a function of this displacement. Since $d^2(k) \leqslant 1$ and assumes its smallest values for a displacement of half a gridlength, the factor $d^{-2}(k)$ can become quite large especially for a displacement approaching half a gridlength.

In Fig. 2, the spectra of the two analysis error terms of (4.6) are plotted as a function of wavenumber, k, for a displacement of half a gridlength.

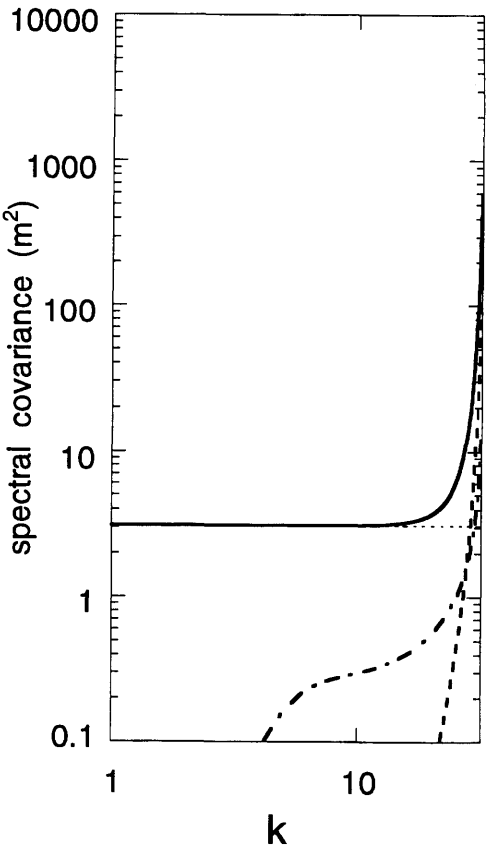

Fig. 2. The spectra of the analysis error variance due to instrument error (solid line), forward interpolation error (dashed line) and signal/error covariance (dashed-dotted line) for $\boldsymbol{H}=\boldsymbol{H}_2$. The displacement between the observation network and the analysis mesh is half a gridlength. The dotted line is the analysis error variance (due to instrument error alone) for $\boldsymbol{H}=\boldsymbol{H}_1$.

The solid line is the instrument error contribution, $r^2(k)/d^2(k)$, while the dashed line is the forward interpolation error contribution, $s^2(k)\ [1+d^2(k) -2d(k)\cos\phi(k)]/d^2(k)$. The dotted line in the figure indicates the constant analysis error value for the case $\boldsymbol{H}=\boldsymbol{H}_1$. Also plotted is the spectrum of $x(k)$, the signal/analysis error covariance (dash-dotted line). (Note that for a displacement of half a gridlength, it follows from Fig. 10b that $\phi(k)$, and hence $\xi(k)$, $=0$, i.e., $\hat{\boldsymbol{X}}$ is diagonal.) The fact that the analysis error and the signal are positively correlated implies that the analysis is an amplified version of the signal. The results are seen to be consistent with (4.6), (4.7) and Fig. 10. Thus, for k less than about 20, $d(k)\approx 1$ and so (i) the analysis error spectra are hardly affected by the fact that $\boldsymbol{H}=\boldsymbol{H}_2$, i.e., these spectra are virtually constant at the values they take in the case $\boldsymbol{H}=\boldsymbol{H}_1$, and (ii) it follows from (4.7) that $x(k)\approx 0$. However, as k approaches the truncation limit, $d(k)\to 0$ and so all of these spectra exhibit an exponential increase with wavenumber.

In Fig. 3, the total contributions (integrated over wavenumber) of the instrument error (solid line) and forward interpolation error (dashed line) are plotted against the displacement of the observation network with respect to the analysis mesh. Consistent with Fig. 10, both terms, and therefore also their sum (the total analysis error variance), exhibit the same behaviour with a pronounced peak at a displacement of half a gridlength. The instrument error covariance term is seen to dominate, as implied by (4.6), [since at large k, $r^2 > s^2(k)$]. For reference, the constant analysis error value for the case $\boldsymbol{H}=\boldsymbol{H}_1$ is plotted as the dotted line in Fig. 3.

For $\boldsymbol{H}=\boldsymbol{H}_2$, we also examined the signal/analysis error covariance in physical space (not shown), by transforming $\hat{\boldsymbol{X}}$ to physical space. This exhibited the same type of dependence on the displacement between the observation network and the analysis mesh as the dashed and solid curves of Fig. 3, but the absolute values were much smaller. The maximum value, which occurred for a half gridlength displacement, was 38 m^2.

4.2. *Homogeneous (nonuniform) observation networks*

In Figs. 2 and 3, the analysis error for the operator $\boldsymbol{H}_2$ (which includes both instrument and

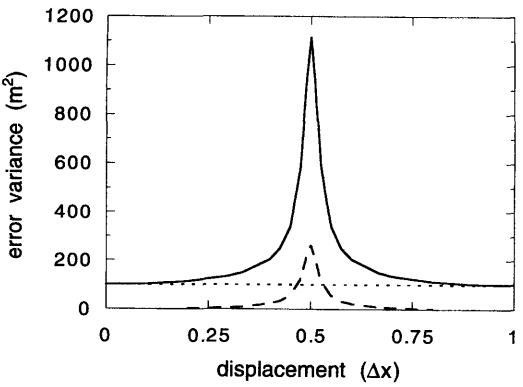

Fig. 3. As in Fig. 2 except for the total instrument and forward interpolation errors (integrated over wavenumber) as a function of the displacement (in gridlengths) between the observation network and the analysis mesh.

forward interpolation error) is larger than that for operator $\boldsymbol{H}_1$ (which has only instrument error). That this is not always the case is illustrated in Fig. 4, where again $I = J$, but now each observation is randomly situated subject to the constraint that it lie not more than one-third of a gridlength from the corresponding point of the analysis mesh. The resulting analysis error variance, when the exact Fourier procedure, $\boldsymbol{H}_1$, is used to perform the forward interpolation, is shown in panel (a). As discussed following (2.26), the total analysis error variance (from $\hat{\boldsymbol{P}}^a$) on the high-resolution output grid (the solid curve) coincides with the analysis error variance at the points of the analysis grid (the diagonal elements of $\bar{\boldsymbol{P}}^a$, represented by the dots). It can be seen that the nonuniformity exhibited by both measures of analysis error is much more pronounced in the case of the total analysis error variance on the high-resolution output grid. (Reducing the degree of nonuniformity in the observation network, by situating each observation closer to the corresponding point of the analysis mesh, was found to reduce the degree of nonuniformity in both measures of analysis error variance (result not shown)). The corresponding analysis error with the cubic spline forward interpolation matrix, $\boldsymbol{H}_2$, is shown in panel (b). The use of $\boldsymbol{H}_2$ is seen to result in a damping of the oscillations of panel (a), even though this interpolation procedure has both instrument and forward interpolation error. For $\boldsymbol{H} = \boldsymbol{H}_2$, the signal/analysis error covariance is non-zero and, consistent with the discussion in Subsection 4.1, the diagonal elements of $\hat{\boldsymbol{X}}$ (not shown) tend to be positive.

The case of networks having more observation

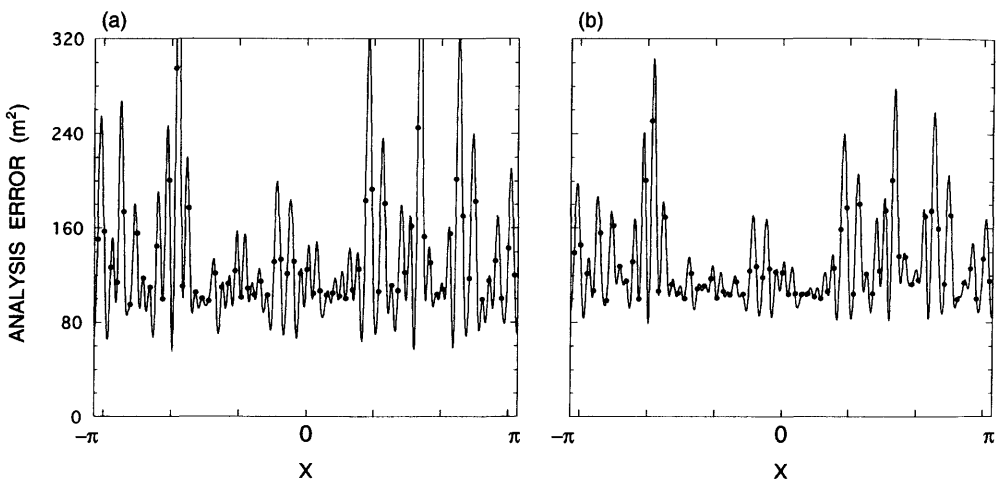

Fig. 4. The analysis error variance as a function of the spatial coordinate, x, for a homogeneous observation network with 65 observations. Each observation is constrained to lie not more than one-third of a gridlength from the corresponding point of the analysis mesh. The dots represent the analysis error variance at the points of the analysis grid (i.e., the diagonal elements of $\bar{\boldsymbol{P}}^a$), while the solid curve represents the total analysis error variance on a high-resolution output grid, i.e., from $\hat{\boldsymbol{P}}^a$ transformed to spatial form as described in Appendix A. $\boldsymbol{H}_1$ is used for the forward interpolation in panel (a), while $\boldsymbol{H}_2$ is used in panel (b).

stations than analysis points ($I > J$) will be discussed more fully in Subsection 4.3 of Part II. Here we simply quote a result which applies when the analysis mesh resolves all scales. In the limit of infinite observation density, the analysis error for $\boldsymbol{H}_1$, the Fourier interpolator, asymptotes to zero, because the only source of error is the spatially uncorrelated instrument error. However, for $\boldsymbol{H}_2$, the cubic spline interpolator, the same asymptotic limit is finite because the additional forward interpolation error is spatially correlated, as discussed by Daley (1993). Therefore, for $\boldsymbol{H}_2$, even letting $I \to \infty$ does not add enough independent information into the system to reduce the analysis error below a certain point.

We will now introduce the models $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ into the assimilation system.

5. The effect of an assimilating model

When an assimilating model is used, the situation becomes more complex: the full set of second-moment equations [(2.12), (2.13), (2.17), (2.18)] must now be included, and signal/error correlations feed back into the system. As in Section 4, we examine a hierarchy of increasingly complex cases. We begin with the case where there are no observations.

5.1. *The no-observation case*

In this case it follows from (2.14) that $\bar{\boldsymbol{h}}_n^{\mathrm{a}} = \bar{\boldsymbol{h}}_n^{\mathrm{f}}$. This implies then that $\bar{\boldsymbol{e}}_n^{\mathrm{a}} = \bar{\boldsymbol{e}}_n^{\mathrm{f}}$ and therefore that $\bar{\boldsymbol{P}}_n^{\mathrm{a}} = \bar{\boldsymbol{P}}_n^{\mathrm{f}}$ and $\boldsymbol{X}_n = \boldsymbol{Y}_n$.

For the perfect model $\boldsymbol{M}_1$, (2.12)–(2.13) become

$$\bar{\boldsymbol{P}}_{n+1}^{\mathrm{a}} = \boldsymbol{M}\bar{\boldsymbol{P}}_n^{\mathrm{a}}\boldsymbol{M}^{\mathrm{T}}, \tag{5.1}$$

$$\boldsymbol{X}_{n+1} = \boldsymbol{M}\boldsymbol{X}_n\hat{\boldsymbol{M}}^{\mathrm{T}}. \tag{5.2}$$

If $\bar{\boldsymbol{P}}_0^{\mathrm{a}}$ is homogeneous, $\bar{\boldsymbol{P}}_\infty^{\mathrm{a}} = \bar{\boldsymbol{P}}_n^{\mathrm{a}} = \bar{\boldsymbol{P}}_0^{\mathrm{a}}$. If $\bar{\boldsymbol{P}}_0^{\mathrm{a}}$ is not homogeneous, $\bar{\boldsymbol{P}}_n^{\mathrm{a}} = \boldsymbol{M}_1\boldsymbol{M}_1 \ldots \boldsymbol{M}_1\bar{\boldsymbol{P}}_0^{\mathrm{a}}\boldsymbol{M}_1^{\mathrm{T}} \ldots \boldsymbol{M}_1^{\mathrm{T}}\boldsymbol{M}_1^{\mathrm{T}}$, i.e., the initial covariance is simply advected around the domain. From (5.2), $\boldsymbol{X}_n$ can be obtained from $\boldsymbol{X}_0$ in a similar manner. Thus, for the perfect model with no observations, the analysis error and the signal/error correlation depend on the specification of the initial conditions.

For the imperfect model $\boldsymbol{M}_2$ (which has discretization error) and no observations, the analysis error and signal/error correlation are somewhat different. Unlike the standard KF with an externally specified model error, $\boldsymbol{Q}$, for which $\bar{\boldsymbol{P}}_n^{\mathrm{a}}$ becomes unbounded as $n \to \infty$, in the present case $\hat{\boldsymbol{X}}_n \to -\hat{\boldsymbol{S}}$ and $\hat{\boldsymbol{P}}_n^{\mathrm{a}} \to \hat{\boldsymbol{S}}$ as $n \to \infty$. A demonstration of this result requires some equations to be derived presently and will therefore be deferred to the end of Subsection 5.3.

5.2. *Perfect model ($\boldsymbol{M} = \boldsymbol{M}_1$) with any observation network*

In this case (2.12)–(2.13) become

$$\bar{\boldsymbol{P}}_{n+1}^{\mathrm{f}} = \boldsymbol{M}\bar{\boldsymbol{P}}_n^{\mathrm{a}}\boldsymbol{M}^{\mathrm{T}}, \tag{5.3}$$

$$\boldsymbol{Y}_{n+1} = \boldsymbol{M}\boldsymbol{X}_n\hat{\boldsymbol{M}}^{\mathrm{T}}, \tag{5.4}$$

while (2.17)–(2.18) remain unchanged.

First, suppose that $\boldsymbol{H} = \boldsymbol{H}_1$, then $[\boldsymbol{G} - \boldsymbol{HF}] = 0$, there is no forward interpolation error, and therefore the definitions of the gain matrix, (2.19) and (2.20), are identical. In fact, $\boldsymbol{Y}_n$ drops out of (2.17) in this case, so our second-moment equations (5.3) and (2.17) reduce to the conventional KF equations for a perfect model. Then in general, $\bar{\boldsymbol{P}}_n^{\mathrm{a}} \to 0$ as $n \to \infty$ for any non-zero Courant number, C. This result does not hold when the system is not observable which can occur when there are very few observations, e.g. the system may not be observable if there is only a single observation and $C > 2$ (Ménard, 1994). However, aside from these exceptional cases, the stationary solution is: $\bar{\boldsymbol{P}}^{\mathrm{f}} = \bar{\boldsymbol{P}}^{\mathrm{a}} = \boldsymbol{X} = \boldsymbol{Y} = \boldsymbol{K} = 0$ and therefore, from (2.25), $\hat{\boldsymbol{P}}^{\mathrm{a}} = \hat{\boldsymbol{P}}^{\mathrm{f}} = 0$. The rate of convergence to the zero stationary solution was also examined. Daley and Ménard (1993) showed that when the observation network and analysis mesh were coincident, the rate of convergence to the zero stationary solution was proportional to $(E^{\mathrm{r}})^2/n$. Our results showed that in the general case, convergence is still proportional to $(E^{\mathrm{r}})^2/n$ but the constant of proportionality is smaller with fewer observations. The dotted curve in Fig. 5, obtained with $\boldsymbol{H}_1$ and a 65-point uniform observation network displaced from the analysis mesh by one-third of a gridlength, exhibits this $(E^{\mathrm{r}})^2/n$ convergence.

When $\boldsymbol{H} = \boldsymbol{H}_2$, then $[\boldsymbol{G} - \boldsymbol{HF}] \neq 0$. Unlike the previous case, $\boldsymbol{X}_n$ and $\boldsymbol{Y}_n$ do not drop out of the second-moment equations [(5.3), (5.4), (2.17), (2.18)] and the generalized and conventional gain matrices are no longer the same. The situation is considerably more complex than was the case with

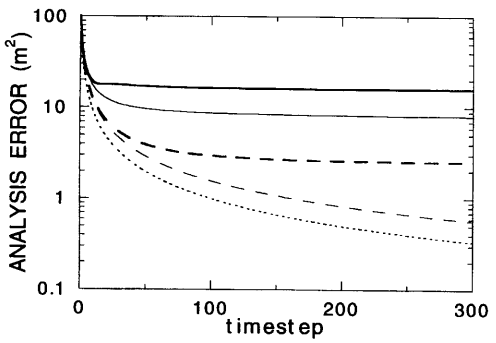

Fig. 5. Analysis error vs. time for the first 300 timesteps of various perfect-model integrations (as described in the text). In all cases: $C=0.8$, the initial analysis error is set to 100 m^2 and the initial signal/error covariance is set to zero.

$\boldsymbol{H}=\boldsymbol{H}_1$ and the results depend on the gain matrix used.

The previous stationary solution, $\bar{\boldsymbol{P}}^{\mathrm{f}}=\bar{\boldsymbol{P}}^{\mathrm{a}}=\boldsymbol{X}=\boldsymbol{Y}=0$, is now not unique. In fact, it can be seen that with the generalized gain matrix, (2.19), and any observation network, if

$$\bar{\boldsymbol{P}}_n^{\mathrm{f}}\boldsymbol{H}^{\mathrm{T}}=\boldsymbol{Y}_n(\boldsymbol{G}-\boldsymbol{H}\boldsymbol{F})^{\mathrm{T}}, \tag{5.5}$$

then $\boldsymbol{K}_n$ is zero, and therefore the solution is independent of the observations and can be described by (5.1) and (5.2). Since (5.5) only determines $\bar{\boldsymbol{P}}_n^{\mathrm{f}}$ and $\boldsymbol{Y}_n$ to within a multiplicative constant, and since the forcing terms (involving the model discretization error) in (2.12) and (2.13) have vanished, *there are in fact an infinite number of stationary solutions with this gain matrix*, i.e., the solution one actually obtains depends on the initial conditions. With the initial signal/error covariance set to zero, it was often observed that the initial analysis error would decrease until eventually $\boldsymbol{K}_n$ approached zero, at which point the observations were disregarded and further reduction of the analysis error stopped. This type of behaviour is exhibited in Fig. 5 by the thin solid curve, which shows the convergence with $\boldsymbol{H}_2$ and the shifted (by one-third of a gridlength) 65-point uniform observation network.

Uniform observation networks with $I=J$, i.e., networks shifted with respect to the analysis mesh, were generally found to be detrimental to filter performance. The filter was much more successful in reducing the initial analysis error towards zero when homogeneous (non-uniform) networks, even having less observations, were used. For example, compare in Fig. 5 the convergence in the case of a homogeneous 65-observation network (the thin dashed curve) with that in the case of the shifted 65-observation uniform network (the thin solid curve). (For the homogeneous network, it is the average analysis error at the gridpoints that is plotted.) This last result is consistent with the findings of Callies and Eppel (1995), who found that uniform observation networks shifted with respect to the analysis mesh give rise to ill-posedness in a 4-d variational context.

Numerical experiments were also done (for $\boldsymbol{H}=\boldsymbol{H}_2$) using the conventional gain matrix, (2.20), in place of the generalized one. This change was found to degrade filter performance. This can be seen in Fig. 5 by comparing the heavy solid and dashed curves, for which $\boldsymbol{K}$ was defined by (2.20), with their thin counterparts.

5.3. *Coincident observation network and analysis mesh with imperfect model ($\boldsymbol{M}=\boldsymbol{M}_2$)*

In this case, we assume that the observation network and analysis grid coincide, and consequently, there is no forward interpolation error. We use the imperfect model $\boldsymbol{M}_2$, thus concentrating entirely on the effect of the model discretization error. Thus $\boldsymbol{F}=\boldsymbol{G}$, $\boldsymbol{H}=\boldsymbol{I}$ and $\boldsymbol{G}-\boldsymbol{H}\boldsymbol{F}=0$. It follows that the definitions of the gain matrices, (2.19) and (2.20), are identical, i.e., in this case the conventional gain matrix minimizes the trace of (2.17). Then (2.12)–(2.13) remain unchanged (with $\boldsymbol{Q}_n=\boldsymbol{Q}$), while (2.17)–(2.18) reduce to:

$$\bar{\boldsymbol{P}}_n^{\mathrm{a}}=[\boldsymbol{I}-\boldsymbol{K}_n]\bar{\boldsymbol{P}}_n^{\mathrm{f}}[\boldsymbol{I}-\boldsymbol{K}_n]^{\mathrm{T}}+\boldsymbol{K}_n\boldsymbol{R}^{\mathrm{i}}\boldsymbol{K}_n^{\mathrm{T}} \tag{5.6}$$

$$\boldsymbol{X}_n=[\boldsymbol{I}-\boldsymbol{K}_n]\boldsymbol{Y}_n. \tag{5.7}$$

As shown in Appendix C, (2.12), (2.13), (5.6), (5.7) together with (2.25) can be decomposed spectrally, yielding a set of 6 scalar equations for each wavenumber. Thus one can either integrate the second-moment equations, or the equivalent scalar equations (C.6)–(C.11). Upon integrating these equations, we found that convergence to the stationary solution was slow for the small wavenumbers, requiring more than a thousand timesteps.

In Fig. 6 is plotted $\bar{\boldsymbol{P}}^{\mathrm{a}}$ at equilibrium (solid curve), which is homogeneous, versus Courant number, C, for $0\leqslant C\leqslant 1$. As noted in Section 3, for C an integer, $\boldsymbol{M}_2=\boldsymbol{M}_1$ and therefore the ana-

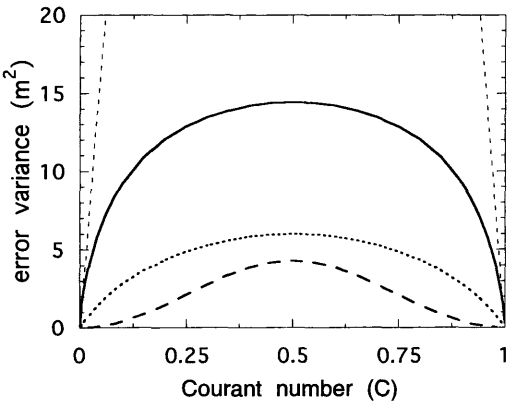

Fig. 6. The analysis (solid curve) and model (heavy dashed curve) error variance due to the discretization error of model $\boldsymbol{M}_2$ as a function of Courant number for an observation network which coincides with the analysis mesh. The thin dashed line is the negative of the signal/analysis error covariance, i.e., $-\hat{\boldsymbol{X}}$ transformed to spatial form. The dotted line is the analysis error variance when the signal/error correlations are neglected.

lysis error is zero for $C=0$ and 1. Since $\boldsymbol{M}_2$ involves interpolating the values of $h(x-U\Delta t)$ from the values of h at the gridpoints of the mesh, the error incurred depends only on the distance of $h(x-U\Delta t)$ from those gridpoints and therefore the analysis error varies periodically with C for C outside of the range shown in the figure. It can be seen that the analysis error is symmetric about $C=0.5$ and rises very sharply as C becomes non-integer quickly reaching values close to the maximum value of 14.4 m^2. The small size of the error explains the slow convergence noted in the previous paragraph (Daley and Ménard, 1993). The corresponding analysis error variance when the signal/error correlations are neglected is plotted as the dotted line in Fig. 6. It can be seen that while the neglect of these correlations does not change the shape of the curve, there is a substantial reduction in the analysis error variance. The analysis system, unable to take the effect of the serial correlations into account, produces an analysis error estimate that is substantially too small.

Also shown in Fig. 6 are the model error $\boldsymbol{Q}=[\boldsymbol{MF}-\boldsymbol{F}\hat{\boldsymbol{M}}]\hat{\boldsymbol{S}}[\boldsymbol{MF}-\boldsymbol{F}\hat{\boldsymbol{M}}]^{\mathrm{T}}$ of (2.12), plotted as the heavy dashed line, and the negative of the signal/analysis error covariance, i.e. $-\hat{\boldsymbol{X}}$ transformed to spatial form, plotted as the thin dashed line. The latter curve, like the others, is symmetric about $C=0.5$ where it reaches its maximum value, which is 76.5 m^2. Note that the analysis error and the signal are now negatively correlated. Assuming that the amplitude of the error is small compared to the amplitude of the signal, this implies that the analysis will be a damped version of the signal. The signal/forecast error covariance, i.e., $\hat{\boldsymbol{Y}}$ transformed to spatial form, is not shown, but it is also negative.

We are also interested in the behaviour of the analysis error as a function of wavenumber, k. The simplest case of (C.6)–(C.11) that can be analyzed occurs when $\phi=0$ which, from Fig. 10b (Appendix B), occurs for a Courant number of 0.5. In this case (C.6)–(C.11) simplify as discussed in Appendix C and, as $n\to\infty$, reduce to

$$x=-s^2(1-d)r^2/[f^2+r^2(1-d)], \tag{5.8}$$

$$f^2=d^2r^2f^2/(f^2+r^2)+s^2(1-d)^2-2xd(1-d), \tag{5.9}$$

where $x=x(k)$ and $f^2=f^2(k)$ are the stationary values of the diagonal elements of $\hat{\boldsymbol{X}}_n$ and $\bar{\boldsymbol{P}}^{\mathrm{f}}_n$, respectively, $s^2=s^2(k)$ and r^2 are the corresponding signal and instrument error variances and $d=d(k)$ is the damping due to the model $\boldsymbol{M}_2$. Eq. (5.8) demonstrates that the analysis error and the signal are negatively correlated, since $0\leqslant d\leqslant 1$.

Substitution of (5.8) into (5.9) yields a fairly general cubic equation in f^2. However, it is straightforward to show that when $d\to 1$ (i.e., the model becomes perfect) that f^2 and $a^2\to 0$. (The a^2 limit follows from (C.9).) When $d\to 0$, then $f^2\to s^2$ and, again from (C.9), $a^2\to s^2r^2/(s^2+r^2)$.

Fig. 10a shows that for $C=0.5$, $d(k)\to 1$ at small k, while $d(k)\to 0$ at large k. Therefore $a^2\to 0$ at small k, and $a^2\to s^2r^2/(s^2+r^2)$ at large k. Now r^2 (the heavy dashed line in Fig. 7a) is independent of k, while s^2 (the heavy solid line) decreases exponentially as k increases. Therefore, for k large, $a^2\to s^2r^2/(s^2+r^2)\approx s^2$. A similar argument shows that $x\to 0$ at small k, and $x\to -s^2$ at large k. It follows that $a^2(k)$ and $-x(k)$ will each have a maximum at some scale larger than the minimum resolvable scale and $\to s^2$ at large k. This behaviour is exhibited by the thin solid and thin dash-dotted curves in Fig. 7a which show $a^2(k)$ and $-x(k)$, respectively, in the case $C=0.5$. Examination of the more general case, (C.6)–(C.11), shows very similar results for $0.25\leqslant C\leqslant 0.75$. However, $a^2(k)$

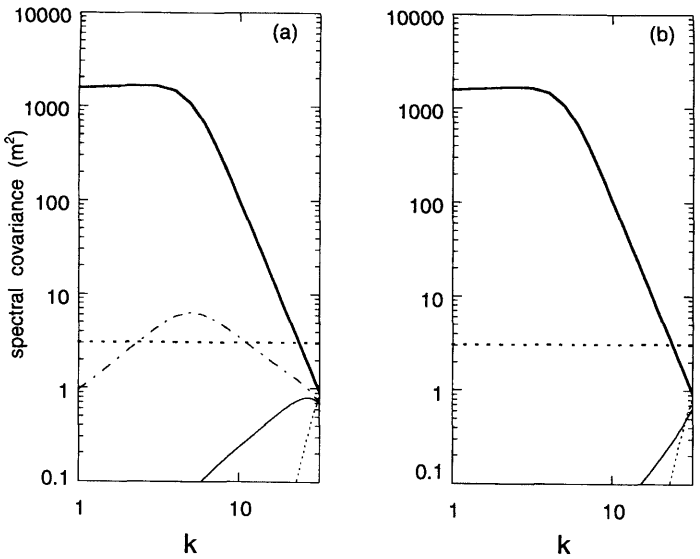

Fig. 7. The analysis (thin solid curve) and model (dotted curve) error variance due to the discretization error of model $\boldsymbol{M}_2$ as a function of wavenumber, k, for $C=0.5$ for an observation network which coincides with the analysis mesh. For reference, the spectra of the signal (heavy solid curve) and of the observation error (heavy dashed curve) are also shown. In panel (a), the complete system is considered and $-x(k)$ is plotted as the thin dash-dotted curve; while in panel (b), the effect of signal/error correlations is neglected.

is rather different when the signal/error correlations are neglected, as seen in Fig. 7b.

The validity of these results is not strictly limited to the case of coincident observation network and analysis mesh. If $\boldsymbol{H}=\boldsymbol{H}_1$ and if the uniform observation network be displaced with respect to the analysis mesh, the results of Figs. 6 and 7 still apply since $\boldsymbol{H}_1$ has no forward interpolation error. $\boldsymbol{H}_2$, the cubic spline forward interpolation matrix, has forward interpolation error. Therefore when $\boldsymbol{H}=\boldsymbol{H}_2$, the results (not shown) are changed somewhat, although overall they remain similar.

We are now in a position to establish the result quoted in Subsection 5.1. At stationarity, in the case of no observations, (C.9)–(C.11) reduce to $a^2=f^2$, $x=y$, and $\xi=v$, respectively. Then (C.6)–(C.8) become

$$(1-d\cos\phi)(x+s^2)+(d\sin\phi)\xi=0, \tag{5.10}$$

$$(d\sin\phi)(x+s^2)-(1-d\cos\phi)\xi=0, \tag{5.11}$$

$$(1-d^2)(a^2+s^2+2x)=0. \tag{5.12}$$

Now in general for $k\geqslant 1$, $d\neq 1$ and $\phi\neq 0$, since $\boldsymbol{M}=\boldsymbol{M}_2$. Therefore the solution of (5.10)–(5.12) is: $\xi=0$, $x=-s^2$ and $a^2=s^2$. Since this is true for every wavenumber, the results of Subsection 5.1 follow directly, i.e., for model $\boldsymbol{M}_2$ with no observations, $\hat{\boldsymbol{X}}_n\to-\hat{\boldsymbol{S}}$ and $\hat{\boldsymbol{P}}_n^{\mathrm{a}}\to\hat{\boldsymbol{S}}$ as $n\to\infty$.

The results of the next subsection will demonstrate the large impact that even a single observation can have.

5.4. *Single observation with imperfect model ($\boldsymbol{M}=\boldsymbol{M}_2$)*

The asymptotic analysis error in the case of a single observation located at $x=0$ is shown in Fig. 8a for $C=0.5$ and 0.9. (For such an observation network, $\boldsymbol{G}-\boldsymbol{HF}=0$ and therefore the gain matrices defined by (2.19) and (2.20) are identical.) The curves for the two Courant numbers have the same shape and indicate that the analysis error is a minimum just downstream of the observation, grows monotonically as one proceeds downstream and then decreases abruptly in the neighbourhood of the observation. The results are similar to those obtained with a single observation by Bennett and Budgell (1987) and Ménard (1994). (In a study of

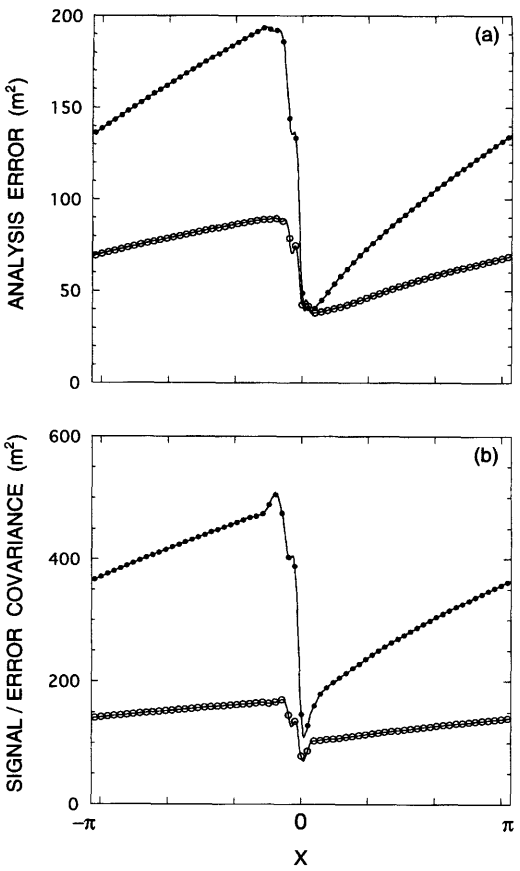

Fig. 8. In panel (a), the dots and open circles represent the analysis error variance on the analysis grid for $C=0.5$ and 0.9, respectively, in the case of a single observation located at $x=0$. The solid curves represent the corresponding analysis error variances on the high-resolution output grid, as in Fig. 4. Panel (b) is the same as panel (a) but for the negative of the diagonal elements of the signal/analysis error covariance, i.e., $-\hat{\boldsymbol{X}}$ transformed to spatial form.

the effect of different observation networks, Daley (1993c) observed that a data void produced a similar analysis error pattern.) Although both curves exhibit error of the same magnitude just downstream of the observation, the error growth is significantly greater for $C=0.5$. This is due to the larger model discretization error in that case, as seen in Fig. 6.

The corresponding signal/analysis error covariance, i.e., $\hat{\boldsymbol{X}}$ transformed to spatial form, is also of interest. The negative of the diagonal elements of this spatial covariance matrix are shown in Fig. 8b for $C=0.5$ and 0.9. (The corresponding signal/forecast error covariance (not shown) is also negative and is very similar.) In general it can be seen that the curves in Fig. 8b have the same form as those in Fig. 8a.

Displacement of the single observation to a location midway between two gridpoints yielded results very similar to those of Fig. 8, irrespective of the choice of $\boldsymbol{H}_1$ or $\boldsymbol{H}_2$ as forward interpolation matrix or (2.19) or (2.20) as gain matrix. *By contrast, dropping the signal/error correlations had a major impact on the results, yielding a much smaller and much more homogeneous analysis error. In fact, the analysis error corresponding to that of Fig. 8a remained below 10 m² at all gridpoints and decreased by only 1 m² in the vicinity of the observation.* (Related results are discussed in Sect. 5.4 of Part II.) As noted above, the results in Fig. 8a are similar to those obtained by Bennett and Budgell (1987) and Ménard (1994) using a red model error spectrum and no signal/error correlation. In the present paper, the model error spectrum is quite different (see Fig. 7) and signal/error correlations are considered. It is interesting then that the similarity between the present results and those in these other studies is strengthened by correctly incorporating the signal/error correlations.

5.5. *Homogeneous observation networks with imperfect model ($\boldsymbol{M}=\boldsymbol{M}_2$)*

Fig. 9 shows two typical analysis error distributions obtained with homogeneous networks. In both cases, the equations were integrated to equilibrium and the choice of $\boldsymbol{H}_1$ versus $\boldsymbol{H}_2$ as forward interpolation matrix produced barely perceptible differences in the results. The upper curve, for which $C=0.5$, was obtained with a network having 15 observations. Each observation gives rise to the characteristic pattern of abrupt decline in the error followed by monotonic growth downstream, which was observed in Fig. 8. The overall effect is a significant reduction in the analysis error in comparison with the case of a single observation. The lower curve shows the analysis error for the same observation network (with $I=65$) as was used for Fig. 4. In this case, $C=1.25$ and the increased number of observations results in a further decrease in the analysis error.

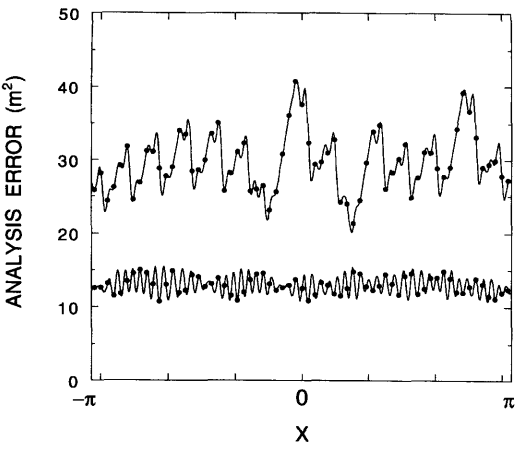

Fig. 9. The analysis errors obtained with forward interpolation matrix $\boldsymbol{H}_2$ and model $\boldsymbol{M}_2$ for two homogeneous observation networks. For the upper curve, $C=0.5$ and the network consists of 15 observations; for the lower curve, $C=1.25$ and the network, consisting of 65 observations, is the same as was used in Fig. 4.

A comparison of Figs. 4 and 9 shows that even an imperfect model can have a substantial impact on data assimilation; despite the model discretization error, the analysis error is substantially reduced by the use of the model. We note also that the analysis error with the 65-point homogeneous observation network (lower curve, Fig. 9) has approximately the same magnitude as that with a uniform network having the same number of observations (cf. Fig. 6). It is also of interest that the analysis error is rather insensitive to the forward interpolation operator, as noted in Subsection 5.3 for uniform observation networks.

6. Summary and conclusions

There are many possible sources of error in atmospheric data assimilation, because imperfect observations are combined with forecasts from imperfect models in less than optimal fashion. This study has focused on three types of errors: (i) model discretization errors, (ii) errors due to the forward interpolation operator (from the discrete model grid to the observation network), and (iii) the enhanced effect of both of these types of errors due to their correlation with the signal (truth). It should be noted, that since the signal is serially correlated, so will be the model discretization and forward interpolation errors.

The study was conducted by applying generalized Kalman filter (KF) equations to a simple one-dimensional linear advection equation. The KF equations were sufficiently general to include evolution equations for the signal/error covariances. The system was completely closed at the second order (with no approximation) and the entire study was performed by examination of the second-moment error statistics. In this paper (Part I), it was assumed that the analysis mesh could resolve all spatial scales.

The second-moment equations derived here are valid for any model discretization or forward interpolation scheme, however, attention has been confined to two model discretization schemes. The first was a semi-Lagrangian technique using Fourier interpolation which had no discretization error. The second algorithm, also semi-Lagrangian but using cubic spline interpolation, had both phase and amplitude errors, with maximum error at the smallest (resolved) scales.

Several types of observation networks were examined, including uniform networks (which, in some cases, coincided with the analysis mesh) and non-uniform, but homogeneous, networks. Two forward interpolation operators were utilized: Fourier interpolation (with no error in the resolved scales) and cubic spline interpolation (with both amplitude and phase errors). Despite the simplicity of the linear advection equation, the data assimilation system yields second-moment error statistics of considerable subtlety and complexity.

In order to provide a basis of comparison for experiments involving the two semi-Lagrangian models, the first experiments (Section 4) were run with no model. This focused attention squarely on the role of forward interpolation error and its correlation with the signal. The particularly simple case of a uniform network with the same number of observation stations as analysis meshpoints (but not necessarily coincident with them) could be analyzed analytically. In this case, the similar structure of the observation network and the analysis mesh produced numerous analogies between the forward interpolation error and the model discretization error (discussed in Section 3 and Appendix C). For this special network, the analysis error for the perfect (Fourier) forward

interpolation operator consisted of instrument error only. However, when the forward interpolation operator was imperfect (cubic spline), the analysis error could be considerably amplified in some cases. In part, this amplification occurred because of the contribution of forward interpolation error, but surprisingly, it was mainly due to the damping properties of the cubic spline operator. Further no-model experiments with more realistic non-uniform (but homogeneous) networks indicated that despite the absence of forward interpolation error, the perfect (Fourier) interpolation operator performed no better than the imperfect (cubic spline) forward interpolator.

The most important experiments included an assimilating model, either perfect (Fourier semi-Lagrangian) or imperfect (cubic spline semi-Lagrangian). Two special cases were examined first: the no-observation case and the perfect model (with any observation network) case. For the no-observation case, the asymptotic limit depended upon whether the model was perfect (in which case the initial analysis error was simply advected around the domain) or imperfect (in which case the analysis error tended to the signal). For the perfect-model case with observations, the behaviour depended on the forward interpolation operator used: with the perfect (Fourier) interpolator, the asymptotic analysis error was zero, but with the cubic spline interpolator, the system could become degenerate yielding an asymptotic analysis error which could depend on the initial conditions. In certain circumstances, smaller analysis errors could be obtained with imperfect models than with perfect models. Consistent with the findings of Callies and Eppel (1995) in a 4-d variational context, the filter was found to be more successful in reducing the initial analysis error when homogeneous (non-uniform) observation networks were used in place of uniform networks shifted with respect to the analysis mesh. A third special case considered the imperfect model and an observation network which coincided with the analysis mesh. In this case an equivalent set of scalar equations for each wavenumber could be obtained, following Daley and Ménard (1993). The principal results of this experiment were that (i) the signal/error correlation was negative (unlike the no-model cases) indicating that the analysis was a damped version of the signal, and (ii) the signal/error correlation tended to greatly increase the (perceived) analysis error and redden its spectrum.

A standard KF experiment (see Bennett and Budgell, 1987, and Ménard, 1994) considers the effect of a single observation (coinciding with an analysis gridpoint). This experiment was also attempted with the more general KF system developed here (and the cubic spline semi-Lagrangian model). Although the discretization error for this model was rather small and concentrated in the smallest scales, the signal/error correlation tended to amplify the effect of this error and produce a large analysis error in the unobserved portion of the domain. This result suggests that simple model error parametrization based purely on the model discretization error (without considering its correlation with the signal) would seriously underestimate the forecast and analysis errors.

Experiments with models and a non-uniform, homogeneous observation network demonstrated the value of even an imperfect model. Despite the presence of a model discretization error correlated with the signal, the analysis error was substantially smaller than in the no-model case. Of course in this paper, the mesh used by the model could resolve all scales in the signal, and therefore the model error tended to be rather small.

Actually, all the results in this paper depend critically on the analysis mesh being able to fully resolve the signal. In reality, this is obviously not the case. We examine the additional complexity due to unresolved scales in Part II.

7. Appendix A

Transformation of spectral covariance matrices to spatial form

Consider $\hat{\boldsymbol{P}}^{\mathrm{a}}_n$, given by (2.25) when all scales are resolved and by (2.7) of Part II in the general case. For purposes of display, it is useful to be able to transform a spectral covariance matrix, such as $\hat{\boldsymbol{P}}^{\mathrm{a}}_n$, to spatial form.

To do this, we define a uniform mesh x_l, $1 \leqslant l \leqslant L$ on the one-dimensional periodic domain $-\pi a \leqslant x \leqslant \pi a$, as in (2.3). Since it takes more than $2K+1$ gridpoints to plot $2K+1$ spectral components faithfully, we generally take $L \gg 2K+1$, in order to see fine spatial structure.

Define $\mathscr{F}$ an $L\times(2K+1)$ Fourier transform matrix between the Fourier coefficients and the output mesh x_l. $\mathscr{F}$ is defined as a sum over all K wavenumbers using (2.1) where x is a point of the output mesh. Then a spectral covariance matrix, such as $\hat{\boldsymbol{P}}^{\mathrm{a}}$ or $\hat{\boldsymbol{X}}$, can be transformed to spatial form on the mesh x_l by left-multiplication by $\mathscr{F}$ and right-multiplication by $\mathscr{F}^{\mathrm{T}}$.

8. Appendix B

Cubic spline interpolation from one uniform mesh to another uniform mesh having the same number of points

Like most interpolation procedures, cubic spline interpolation introduces a spurious damping and phase shift which is a function of wavenumber, k. More precisely, cubic spline interpolation of $h(x)$ has the effect of multiplying its spectral representation $\hat{h}$ by a $(2K+1)\times(2K+1)$ block diagonal matrix with upper left-hand corner element equal to 1 and remaining 2×2 blocks given by

$$d\begin{bmatrix}\cos(kC\Delta x+\phi) & -\sin(kC\Delta x+\phi)\\ \sin(kC\Delta x+\phi) & \cos(kC\Delta x+\phi)\end{bmatrix}, \tag{B.1}$$

where Δx is the gridlength, $d=d(C, k\Delta x)$ and $\phi=\phi(C, k\Delta x)$ are the damping and phase shift due to the spline interpolation, and the constant C will be discussed below. The above matrix represents a rotation through $kC\Delta x+\phi$ (Strang, 1988, p.122) and can be expressed as a product of a rotation through $kC\Delta x$ and a rotation through ϕ, in either order. The rotation through $kC\Delta x$ corresponds to the interpolation by the appropriate distance $C\Delta x$ and the rotation through ϕ corresponds to the spurious phase shift. The effect of the spurious damping and phase shift can be represented as a multiplication by $\hat{\boldsymbol{D}}$, the $(2K+1)\times(2K+1)$ block diagonal matrix with upper left-hand corner element equal to 1 and remaining 2×2 blocks given by

$$\hat{\boldsymbol{d}}=d\begin{bmatrix}\cos\phi & -\sin\phi\\ \sin\phi & \cos\phi\end{bmatrix}. \tag{B.2}$$

We note that when C is an integer, $d=1$ and $\phi=0$ for all k and therefore $\hat{\boldsymbol{D}}=\boldsymbol{I}$, the identity matrix. For the Fourier interpolation, $d=1$ and $\phi=0$ for all k, and therefore $\hat{\boldsymbol{D}}=\boldsymbol{I}$, for *any* C.

From (B.2) it follows that

$$\hat{\boldsymbol{d}}^{-1}=d^{-2}\hat{\boldsymbol{d}}^{\mathrm{T}}, \tag{B.3}$$

and therefore that $\hat{\boldsymbol{D}}\hat{\boldsymbol{D}}^{\mathrm{T}}$ is a diagonal matrix with upper left-hand corner element equal to 1 followed

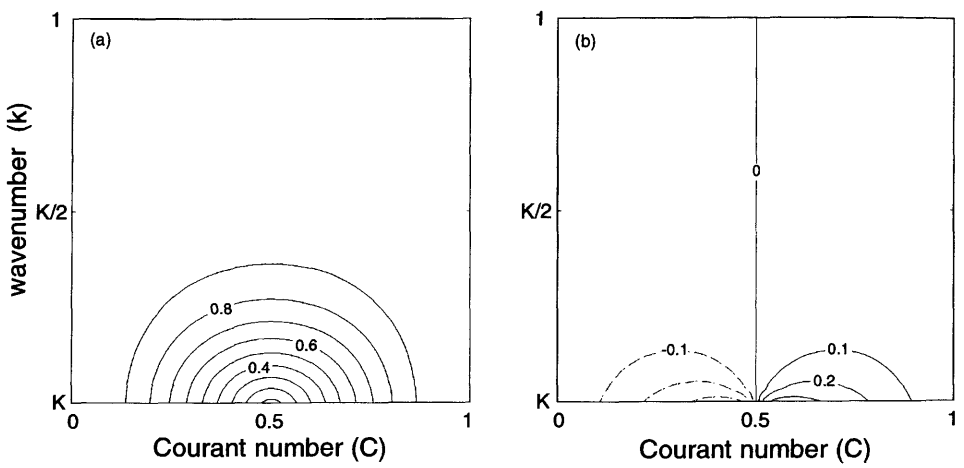

Fig. 10. (a) The damping, d, caused by cubic spline interpolation. The vertical axis is wavenumber, k, for $1\leqslant k\leqslant K$, and the horizontal axis is Courant number, C, for $0\leqslant C\leqslant 1$. The contour interval is 0.1. Panel (b) is the same as panel (a) but for the phase shift, ϕ. In this panel, the contours are in units of π, e.g., the 0.1 contour represents 0.1π, and negative contours are represented by dash-dotted lines.

by pairs of elements equal to $d^2(k)$ for each wavenumber k.

The two panels of Fig. 10 display d and ϕ as a function of k and C as C varies from 0 to 1. (See also Fig. 1 of Pudykiewicz and Staniforth (1984) and note that their eq. (2.11) gives an explicit expression for the response of cubic spline interpolation.) Both d and ϕ vary periodically with C for C outside of the interval [0, 1]. It can be seen that d is a maximum at $C=0.5$ and varies symmetrically about that value, while ϕ is zero for $C=0.5$ and varies antisymmetrically about that value. For large scales the spline interpolation is seen to cause little damping or shifting of phase, but as the resolution limit is approached $d\to 0$ and ϕ becomes large for certain values of C.

The model $\boldsymbol{M}_2$ involves interpolating from the gridpoints to the points of departure. Therefore, setting $C=U\Delta t/\Delta x$ in (B.1) shows that advecting $h(x)$ using the cubic spline model, $\boldsymbol{M}_2$, has the effect of multiplying its spectral representation $\hat{h}$ by the $(2K+1)\times(2K+1)$ block diagonal matrix with upper left-hand corner element equal to 1 and remaining 2×2 blocks

$$\hat{\boldsymbol{m}}_2=\hat{\boldsymbol{m}}\hat{\boldsymbol{d}}=\hat{\boldsymbol{d}}\hat{\boldsymbol{m}}\,, \tag{B.4}$$

where $\hat{\boldsymbol{m}}$ is the matrix defined in (2.8). Eq. (3.4) follows directly.

In a similar way, for a uniform observation network having the same number of points as the analysis mesh, forward interpolation involves interpolating from one uniform mesh to another such mesh having the same number of points. Thus, if $\boldsymbol{H}=\boldsymbol{H}_2$ and C is interpreted as the displacement between the two meshes, the above results apply directly to the forward interpolation.

When there are unresolved scales, as in Part II of this paper, d and ϕ are still given by Fig. 10, if (i) the figure's abscissa is interpreted to be the model Courant number (C_{M}), and (ii) the largest wavenumber on the figure's ordinate (denoted by K) is interpreted to be wavenumber $(J-1)/2$, i.e., the truncation limit. Again d and ϕ vary periodically with C_{M} for C_{M} outside of the interval [0, 1]. Note that as the number of meshpoints (and therefore the truncation limit) decreases, the wavenumbers severely affected by the spline interpolation also decrease. In fact when J is small, even the longest waves are significantly affected by the spline interpolation.

9. Appendix C

Spectral decomposition in the case of model discretization error only

Following Daley and Ménard (1993), we spectrally decompose the second-moment equations in the case of an observation network that coincides with the analysis mesh. Defining

$$\hat{\boldsymbol{X}}_n=\tilde{\boldsymbol{F}}\boldsymbol{X}_n \qquad \text{and} \qquad \hat{\boldsymbol{Y}}_n=\tilde{\boldsymbol{F}}\boldsymbol{Y}_n\,, \tag{C.1}$$

then (2.12), (2.13), (5.6), (5.7) can be rewritten in spectral form as:

$$\hat{\boldsymbol{Y}}_{n+1}=\hat{\boldsymbol{M}}[\hat{\boldsymbol{D}}(\hat{\boldsymbol{X}}_n+\hat{\boldsymbol{S}})-\hat{\boldsymbol{S}}]\hat{\boldsymbol{M}}^{\mathrm{T}}\,, \tag{C.2}$$

$$\hat{\boldsymbol{P}}^{\mathrm{f}}_{n+1}+\hat{\boldsymbol{Y}}_{n+1}+\hat{\boldsymbol{Y}}^{\mathrm{T}}_{n+1}$$

$$=\hat{\boldsymbol{M}}[\hat{\boldsymbol{D}}(\hat{\boldsymbol{P}}^{\mathrm{a}}_n+\hat{\boldsymbol{S}}+\hat{\boldsymbol{X}}_n+\hat{\boldsymbol{X}}^{\mathrm{T}}_n)\hat{\boldsymbol{D}}^{\mathrm{T}}-\hat{\boldsymbol{S}}]\hat{\boldsymbol{M}}^{\mathrm{T}}, \tag{C.3}$$

$$\hat{\boldsymbol{P}}^{\mathrm{a}}_n=[\hat{\boldsymbol{P}}^{\mathrm{f}}_n+\hat{\boldsymbol{R}}^{\mathrm{i}}]^{-1}\hat{\boldsymbol{R}}^{\mathrm{i}}\hat{\boldsymbol{P}}^{\mathrm{f}}_n\,, \tag{C.4}$$

$$\hat{\boldsymbol{X}}_n=[\hat{\boldsymbol{P}}^{\mathrm{f}}_n+\hat{\boldsymbol{R}}^{\mathrm{i}}]^{-1}\hat{\boldsymbol{R}}^{\mathrm{i}}\hat{\boldsymbol{Y}}_n\,. \tag{C.5}$$

Now, $\hat{\boldsymbol{R}}^{\mathrm{i}}$ and $\hat{\boldsymbol{S}}$ are diagonal, while $\hat{\boldsymbol{M}}$ and $\hat{\boldsymbol{D}}$ are block diagonal matrices with 2×2 non-zero blocks along the diagonal. Therefore, the diagonal 2×2 blocks of $\hat{\boldsymbol{P}}^{\mathrm{a}}_n$, $\hat{\boldsymbol{P}}^{\mathrm{f}}_n$, $\hat{\boldsymbol{X}}_n$ and $\hat{\boldsymbol{Y}}_n$ are forced, but the off-diagonal blocks are unforced. Following Ménard (1994), it can be shown that the off-diagonal blocks of $\hat{\boldsymbol{P}}^{\mathrm{a}}_n$, $\hat{\boldsymbol{P}}^{\mathrm{f}}_n$, $\hat{\boldsymbol{X}}_n$ and $\hat{\boldsymbol{Y}}_n$ approach zero as $n\to\infty$, regardless of their initial specification. We therefore concentrate only on the diagonal blocks of $\hat{\boldsymbol{P}}^{\mathrm{a}}_n$, $\hat{\boldsymbol{P}}^{\mathrm{f}}_n$, $\hat{\boldsymbol{X}}_n$ and $\hat{\boldsymbol{Y}}_n$. In fact, since $\bar{\boldsymbol{P}}^{\mathrm{a}}_n$ and $\bar{\boldsymbol{P}}^{\mathrm{f}}_n$ become homogeneous as $n\to\infty$ regardless of their initial specification, $\hat{\boldsymbol{P}}^{\mathrm{a}}_n$ and $\hat{\boldsymbol{P}}^{\mathrm{f}}_n$ become diagonal.

It follows that (C.2)–(C.5) can be decomposed into k sets of 2×2 matrix equations, one set for each wavenumber $k\geqslant 1$. In fact, we can interpret each of the matrices appearing in these equations as a 2×2 matrix. For a given k, define $a^2_n(k)$, $f^2_n(k)$, $r^2(k)$, $s^2(k)$, $x_n(k)$ and $y_n(k)$ as the diagonal elements of $\hat{\boldsymbol{P}}^{\mathrm{a}}_n$, $\hat{\boldsymbol{P}}^{\mathrm{f}}_n$, $\hat{\boldsymbol{R}}^{\mathrm{i}}$, $\hat{\boldsymbol{S}}$, $\hat{\boldsymbol{X}}_n$ and $\hat{\boldsymbol{Y}}_n$, respectively. $\hat{\boldsymbol{X}}_n$, rather than being diagonal, is antisymmetric with equal diagonal elements. The same is true of $\hat{\boldsymbol{Y}}_n$. Let ξ_n and ν_n denote the off-diagonal elements of $\hat{\boldsymbol{X}}_n$ and $\hat{\boldsymbol{Y}}_n$, respectively. Then for each wavenumber k, (C.2)–(C.5) yield the following 6 scalar equations for the 6 variables $a^2_n(k)$, $f^2_n(k)$, $x_n(k)$,

$\xi_n(k)$, $y_n(k)$ and $v_n(k)$:

$$y_{n+1}+s^2=d(x_n+s^2)\cos\phi-d\xi_n\sin\phi, \quad \text{(C.6)}$$

$$v_{n+1} \quad =d(x_n+s^2)\sin\phi+d\xi_n\cos\phi, \quad \text{(C.7)}$$

$$f_{n+1}^2+s^2+2y_{n+1}=d^2(a_n^2+s^2+2x_n), \quad \text{(C.8)}$$

$$(f_n^2+r^2)a_n^2=r^2f_n^2, \quad \text{(C.9)}$$

$$(f_n^2+r^2)x_n=r^2y_n, \quad \text{(C.10)}$$

$$(f_n^2+r^2)\xi_n=r^2v_n, \quad \text{(C.11)}$$

where the index k has been dropped. Eq. (C.8) can also be written in the form,

$$f_{n+1}^2=d^2a_n^2+q^2+2d[x_n(d-\cos\phi)+\xi_n\sin\phi], \quad \text{(C.12)}$$

where $q^2=s^2(1+d^2-2d\cos\phi)$ is the model error spectral variance corresponding to $\hat{\boldsymbol{Q}}$ of (3.7). Note the similarity between the form of q^2 here and the forward interpolation error term of (4.6).

Eqs. (C.6)–(C.11) can be used to examine any model discretization, once $d(k)$ and $\phi(k)$ have been determined. The cubic spline semi-Lagrangian scheme has both phase errors ($\phi\neq0$) and amplitude errors ($d\leqslant1$). Other schemes, such as the Crank-Nicholson, are undamped ($d=1$) but have larger phase errors. We note that at stationarity, (C.6)–(C.11) constitute 6 equations in the 6 unknowns a^2, f^2, x, ξ, y and v, where the latter are the stationary values of a_n^2, f_n^2, x_n, ξ_n, y_n and v_n, respectively.

The simplest case of (C.6)–(C.11) that can be analyzed occurs when $\phi=0$, i.e., (from Fig. 10b) for a Courant number of 0.5. In this case substituting (C.7) into (C.11) yields

$$v_{n+1}=dr^2v_n/(f_n^2+r^2). \quad \text{(C.13)}$$

This equation has the form $v_{n+1}=c_nv_n$, with $0\leqslant c_n<1$, as $d\leqslant1$. This means that as $n\to\infty$, v_n (and ξ_n) will approach zero and this process will be accelerated as $d\to0$. Thus (for $\phi=0$) as $n\to\infty$, (C.6) simplifies while (C.7) and (C.11) degenerate and can be dropped.

REFERENCES

Ahlberg, J., Nilson, E. and Walsh, J. 1967. *The theory of splines and their applications.* Academic Press, New York, 284 pp.

Bennett, A.F. and Budgell, W.P. 1987. Ocean data assimilation and the Kalman filter: Spatial regularity. *J. Phys. Oceanogr.* **17**, 1583–1601.

Callies, U. and Eppel, D.P. 1995. On the sensitivity of a least-squares fit of discretized linear hyperbolic equations to data. *Tellus* **47A**, 80–102.

Cohn, S.E. 1982. *Methods of sequential estimation for determining initial data in numerical weather prediction.* PhD thesis. New York University, 183 pp.

Cohn, S.E. and Dee, D.P. 1988. Observability of discretized partial differential equations. *SIAM J. Numer. Anal.* **25**, 586–617.

Cohn, S.E. and Parrish, D.F. 1991. The behavior of forecast error covariances for a Kalman filter in two dimensions. *Mon. Wea. Rev.* **119**, 1757–1785.

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

Daley, R. 1991. *Atmospheric data analysis,* Cambridge University Press, Cambridge, 457 pp.

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

Daley, R. 1992b. The effect of serially correlated observation and model error on atmospheric data assimilation. *Mon. Wea. Rev.* **120**, 164–177.

Daley, R. 1992c. Forecast-error statistics for homogeneous and inhomogeneous observation networks. *Mon. Wea. Rev.* **120**, 627–643.

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

Daley, R. and Ménard, R. 1993. Spectral characteristics of Kalman filter systems for atmospheric data assimilation. *Mon. Wea. Rev.* **121**, 1554–1565.

Dee, D.P. 1991. Simplification of the Kalman filter for meteorological data assimilation. *Quart. J. Roy. Meteor. Soc.* **117**, 365–384.

Lorenc, A.C. 1986. Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.* **112**, 1177–1194.

Ménard, R. 1994. *Kalman filtering of Burgers' equation and its application to atmospheric data assimilation.* PhD thesis. McGill University, 211 pp.

Ménard, R. and Daley, R. 1996. The application of Kalman smoother theory to the estimation of 4DVAR error statistics. *Tellus* **48A**, 221–237.

Miller, R.N., Ghil, M. and Gauthiez, F. 1994. Advanced data assimilation in strongly nonlinear dynamical systems. *J. Atmos. Sci.* **51**, 1037–1056.

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*, this issue.

Phillips, N.A. 1986. The spatial statistics of random geostrophic modes and first-guess errors. *Tellus* **38A**, 314–332.

Pudykiewicz, J. and Staniforth, A. 1984. Some properties and comparative performance of the semi-Lagrangian method of Robert in the solution of the advection-diffusion equation. *Atmos.-Ocean* **22**, 283–308.

Ritchie, H., Temperton, C., Simmons, A., Hortal, M., Davies, T., Dent, D. and Hamrud, M. 1995. Implementation of the semi-Lagrangian method in a high-resolution version of the ECMWF forecast model. *Mon. Wea. Rev.* **123**, 489–514.

Strang, G. 1988. *Linear algebra and its applications.* (3rd edition). Harcourt Brace Jovanovich, San Diego, 505 pp.

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

Todling, R. and Ghil, M. 1994. Tracking atmospheric instabilities with the Kalman filter. Part I: Methodology and one-layer results. *Mon. Wea. Rev.* **122**, 183–204.

Tribbia, J.J. and Anthes, R.A. 1987. Scientific basis of modern weather prediction. *Science* **237**, 493–499.