Approximate Data Assimilation Schemes for Stable and Unstable Dynamics

By Stephen E. Cohn

Data Assimilation Office, NASA/Goddard Space Flight Center, Greenbelt, Maryland, U.S.A.

and

Ricardo Todling

Universities Space Research Association, NASA/GSFC/DAO, Greenbelt, Maryland, U.S.A.

(Manuscript received 11 July 1995, in revised form 3 November 1995)

Abstract

Two suboptimal data assimilation schemes for stable and unstable dynamics are introduced. The first scheme, the partial singular value decomposition filter, is based on the most dominant singular modes of the tangent linear propagator. The second scheme, the partial eigendecomposition filter, is based on the most dominant eigenmodes of the propagated analysis error covariance matrix. Both schemes rely on iterative procedures like the Lanczos algorithm to compute the relevant modes.

The performance of these schemes is evaluated for a shallow-water model linearized about an unstable Bickley jet. The results are contrasted against those of a reduced resolution filter, in which the gains used to update the state vector are calculated from a lower-dimensional dynamics than the dynamics that evolve the state itself. The results are also contrasted against the exact results given by the Kalman filter. These schemes are validated for the case of stable dynamics as well.

The two new approximate assimilation schemes are shown to perform well with relatively few modes computed. Adaptive tuning of a modeled trailing error covariance for all three of these low-rank approximate schemes enhances performance and compensates for the approximation employed.

1. Introduction

A number of recent studies have made clear that rapid error growth in regions of baroclinic and barotropic instability is nonmodal, and therefore can generally be explained by the singular values and singular vectors, rather than the eigenvalues and eigenvectors, of the tangent linear dynamics (Farrell 1989; Trefethen et al. 1993; Moore and Farrell 1994). Modern four-dimensional (4D) data assimilation methods, such as 4D variational algorithms and nonlinear Kalman filter (KF) schemes, offer the potential to exploit this fact through their direct use of the tangent linear dynamics. Such methods would assign more weight to observations in unstable regions than those in more quiescent regions, all else being equal.

It is well known that practical implementation of Kalman filter schemes requires sensible approximation. Evaluation of several possible approximations, known as suboptimal schemes (SOS’s), was carried out by Todling and Cohn (1994; TC94 hereafter) for a stable 2D linear barotropic model. Results were encouraging for this highly idealized model. However, when these SOS’s were evaluated recently for a barotropically unstable version of that model, we found that all these schemes failed to capture the instability, and generally diverged. On the other hand, Todling and Ghil (1994) have shown that the complete KF is well-behaved in the presence of instability, even when the number of observations is very limited. This motivates the need for more reliable approximations to the KF.

The purpose of this paper is to formulate and evaluate three approximations potentially capable of handling non-normal, unstable, linear and non-
linear dynamics. Results with a linear, barotropically unstable model are reported here (see also Cohn and Todling 1995). One scheme, which we call the reduced resolution filter (RRF), calculates the complete forecast/analysis error covariance evolution according to the standard KF equations, but at a resolution lower than that of the model that evolves the state. The resulting Kalman gain is then interpolated to the model grid and the analysis proceeds as usual. Other reduced resolution filters have been studied by Le Moine and Alvarez (1991) and by Fukumori and Malanotte-Rizzoli (1995).

The other two schemes are new. The first one, called the partial singular value decomposition filter (PSF), utilizes a partial singular value decomposition (SVD) of the tangent linear dynamics between consecutive observation times: the tangent linear propagator is approximated by the leading part of its SVD. This scheme assumes that most of the propagated analysis error covariance is due to a small collection of the model’s most rapidly growing (“leading”) singular modes. It has much in common with ensemble forecasting schemes based on similar ideas (Toth and Kalnay 1993, Molteni et al. 1994). While this will be seen to be a powerful assumption, it does neglect the fact that the propagated analysis error covariance could also be large in more stable, quiescent regions, where too few observations have been available in the recent past to reduce analysis errors.

Our second scheme, called the partial eigendecomposition filter (PEF), generalizes the PSF in an attempt to more fully account for the dependence of the propagated analysis error covariance matrix upon the observations themselves. Rather than calculating a partial SVD of the tangent linear dynamics, the PEF calculates a partial eigendecomposition of that portion of the forecast error covariance matrix that arises from the propagation of the previous analysis error covariance matrix by the complete tangent linear dynamics and its adjoint. The propagated analysis error covariance is then replaced by the leading part of this decomposition. A scheme similar in spirit to the PEF has been proposed by Courtier (1993) as a means of cycling 4D variational algorithms. That scheme employs a partial eigendecomposition of the analysis error covariance matrix (inverse of the Hessian of the cost function) itself. Such an eigendecomposition has been investigated independently by Sheinbaum (1995) as an approach to error analysis in 4D variational methods. A square-root implementation of another PEF-like scheme, without the adaptive tuning described below, has been suggested by Verlaan and Heemink (1995).

In their simplest versions, the schemes studied here — RRF, PSF and PEF — are all low-rank approximations of the full KF. In the analysis step of any low-rank approximation, the only errors that can be reduced by the observations are those that lie in the space spanned by the approximate low-rank forecast error covariance matrix. Errors not in that space cannot be reduced, and in fact can grow in the case of unstable dynamics, or in the presence of model error. We therefore enhance the performance of these schemes by incorporating an additive “trailing” error covariance matrix, resulting in a full rank approximation of the forecast error covariance matrix. The magnitude of the trailing covariance matrix is tuned adaptively, according to the scheme suggested by Dee (1995).

In Sections 2 and 3 we describe briefly the PSF and PEF, respectively. Numerical results are reported in Section 4 and conclusions drawn in Section 5.

2. The partial singular value decomposition filter (PSF)

The forecast error covariance evolution equation for the linear KF, the nonlinear extended KF and second-moment closure nonlinear KF is given in discrete form by

\[ S_k' = S_{k,k-\ell} + Q_{k,k-\ell}, \quad (2.1) \]

where \( Q_{k,k-\ell} \) is the \( n \times n \) model error covariance matrix accumulated between consecutive observation times \( t_{k-\ell} \) and \( t_k \), and \( S_{k,k-\ell}' \) arises from propagation of the \( n \times n \) analysis error covariance matrix \( S_{k,k-\ell}' \).

\[ S_{k,k-\ell}' = \Psi_{k,k-\ell} S_{k,k-\ell}^p \Psi_{k,k-\ell}^T, \quad (2.2) \]

Here \( \Psi_{k,k-\ell} \) denotes the \( n \times n \) propagator, or tangent linear model (TLM), from time \( t_{k-\ell} \) to time \( t_k \), and \( \Psi_{k,k-\ell}^T \) is its adjoint. The calculation represented by (2.2) is extremely expensive for large \( n \). To approximate it, first write the complete SVD of \( \Psi_{k,k-\ell} \) as

\[ \Psi_{k,k-\ell} = (UDV^T)_{k,k-\ell}. \quad (2.3) \]

The singular values, or elements of the \( n \times n \) diagonal matrix \( D \), are real and nonnegative. A Lanczos algorithm (e.g., Golub and Van Loan 1989) can be used to calculate efficiently just some number \( L \) of the largest (“leading”) singular values, entered into an \( L \times L \) diagonal matrix \( D_L \), along with the corresponding left and right singular vectors, forming \( n \times L \) matrices \( U_L \) and \( V_L \), respectively. Having done so, write (2.2) as

\[ S_{k,k-\ell}' = \tilde{\Psi}_{k,k-\ell} S_{k,k-\ell}^p \tilde{\Psi}_{k,k-\ell}^T + T_{k,k-\ell}, \quad (2.4) \]

where

\[ \tilde{\Psi}_{k,k-\ell} = (U_L D_L V_L^T)_{k,k-\ell} \quad (2.5) \]
is a partial SVD of the propagator, and the “trailing” error covariance $T_{k,k-\ell}$ can be written explicitly as the sum of three terms:

$$T_{k,k-\ell} = \begin{bmatrix} \Psi - \tilde{\Psi} \end{bmatrix} S^a \begin{bmatrix} \Psi - \tilde{\Psi} \end{bmatrix}^T + \tilde{\Psi} S^a \begin{bmatrix} \Psi - \tilde{\Psi} \end{bmatrix}^T + \begin{bmatrix} \Psi - \tilde{\Psi} \end{bmatrix} S^a \begin{bmatrix} \Psi - \tilde{\Psi} \end{bmatrix}^T \right)_{k,k-\ell},$$

(2.6)

where the last two terms are the transposes of each other. They have the property of having zero trace:

$$\text{Tr} [\Psi - \tilde{\Psi} S^a] = \text{Tr} [\tilde{\Psi} S^a] = \text{Tr} [\begin{bmatrix} \Psi - \tilde{\Psi} \end{bmatrix} S^a] = 0, \quad (2.7)$$

The PSF is based on (2.4) and must therefore approximate the trailing error covariance matrix (2.6). The simplest PSF scheme takes $T_{k,k-\ell} = 0$, for all times. Another possibility is discussed in Section 4.

Observe from (2.5) that calculating the first term in (2.4) requires calculating only the $L \times L$ covariance matrix $V_{L,k,k-\ell} S^a_{L,k,k-\ell}$ of $T_{L,k,k-\ell}$ at time $t_{k-\ell}$, instead of the complete analysis error covariance matrix. Using the standard KF analysis error covariance formulas, this requires $L$ linear system solves. The first term in (2.4) is then conveniently available as an operator, at considerably less expense (for reasonably small $L$) than the complete formula (2.2). The overall computational cost of the PSF algorithm, relative to that of a standard KF implementation, is proportional to $L/n$.

### 3. The partial eigendecomposition filter (PEF)

Write the complete eigendecomposition of $S_{k,k-\ell}$ in (2.2) as

$$S_{k,k-\ell} = \begin{bmatrix} \Psi_{k,k-\ell} \end{bmatrix} S^a_{k,k-\ell} \begin{bmatrix} \Psi_{k,k-\ell} \end{bmatrix}^T = \begin{bmatrix} W_{L,k,k-\ell} S_{L,k,k-\ell} W_{L,k,k-\ell}^T \end{bmatrix}_{k,k-\ell},$$

(3.1)

Since $S^a_{k,k-\ell}$ is a covariance matrix, its eigenvalues (diagonal elements of the diagonal matrix $S$) are real and nonnegative, and may decrease rapidly, possibly more so than the singular values of the propagator (see Section 4). We choose the dependent variables of our model in such a way that the sum of their squares is proportional to the energy density, hence each eigenvalue can be interpreted as the variance, in units of energy, contained in the corresponding eigenmode.

Now write (3.1) as

$$S^a_{k,k-\ell} = \begin{bmatrix} W_{L,k,k-\ell} S_{L,k,k-\ell} W_{L,k,k-\ell}^T \end{bmatrix}_{k,k-\ell} + T_{k,k-\ell},$$

(3.2)

where the first term is the leading part of the eigendecomposition in (3.1) and the second term $T_{k,k-\ell}$ is the trailing part to be approximated by the PEF. Thus $S_{L,k} = L \times L$ diagonal matrix of the $L$ largest eigenvalues, with $L$ a parameter generally distinct from that of the previous section, and the columns of the $n \times L$ matrix $W_{L,k}$ are the corresponding eigenvectors. The ratio of the trace of $S_{L,k}$ to that of $S$ measures the proportion of propagated analysis error energy explained directly by the PEF dynamics.

The matrices $S_{L,k}$ and $W_{L,k}$ can be calculated efficiently by a Lanczos algorithm, provided the propagator and its adjoint, as well as the approximate analysis error covariance $S_{k,k-\ell}$ are available as operators. The latter can be treated as an operator by incorporating the KF analysis equations directly into the Lanczos algorithm. The computational cost of the PEF algorithm is comparable to that of the PSF, for fixed $L$. Relative to the cost of a standard KF implementation, both are proportional to $L/n$.

An alternative to approximating the propagated error covariance matrix is to approximate the analysis error covariance matrix itself, as suggested by Courtier (1993) and studied independently by Sheinbaum (1995) in the context of 4D variational algorithms. In this alternative, the leading $L$ eigenmodes of the analysis error covariance matrix, explaining some fraction of the total analysis error variance, are retained. Thus in the KF context, decomposing $S_{k,k-\ell}$ as

$$S_{k,k-\ell}^a = \begin{bmatrix} W_{n,k,k-\ell} S_{n,k,k-\ell} W_{n,k,k-\ell}^T \end{bmatrix}_{k,k-\ell},$$

(3.3)

the propagated analysis error covariance matrix (2.2) can be written as

$$S^a_{k,k-\ell} = \begin{bmatrix} W_{n,k,k-\ell} S_{n,k,k-\ell} W_{n,k,k-\ell}^T \end{bmatrix}_{k,k-\ell} + T_{k,k-\ell}^a,$$

(3.4)

where $W_{n,k,k-\ell} = \Psi_{k,k-\ell}^T \tilde{W}_{k,k-\ell} \widetilde{\Psi}_{k,k-\ell}^T$ and $\tilde{W}_{k,k-\ell}$ yields an approximation requiring $L$ TLM integrations between consecutive analysis times, as indicated by the last equality in (3.4), in addition to the availability of $S^a_{k,k-\ell}$ as an operator. Note that (3.4) does not require the adjoint of the TLM whereas (3.2) does, if each is implemented with an iterative Lanczos-type algorithm.

The qualitative difference between approximations (3.2) and (3.4) is that in the former, $S^a_{k,k-\ell}$
The simplest PEF scheme takes \( T_{k,k} = 0 \) (or \( T_k = 0 \)). An alternative model is introduced in the following section, and yet a third possibility has been evaluated by Cohn and Todling (1995).

4. Numerical experiments

a. Model and assimilation setup

To evaluate the schemes described in the previous sections we employ the performance analysis technique described in TC94, along with a modification of the two-dimensional, linear shallow-water model utilized there. The modified model is barotropically unstable: it is linearized about a meridionally-dependent squared-hyperbolic secant jet (Bickley jet; Haltiner and Williams 1980, p. 75) given by

\[
U(y) = U_0 + U_1 \text{sech}^2 \left( \frac{2\pi}{L'} (y - L_y/2) \right),
\]

where \( L_y = 3000 \) km is the meridional extent of the model domain, \( U_0 = 20 \text{ ms}^{-1} \) and \( U_1 = 12.5 \text{ ms}^{-1} \).

A choice of \( L' = 2000 \) km is small enough to provoke instability, given that \( \beta = 1.82 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1} \). This is illustrated in Fig. 1, where the velocity profile is displayed in panel (a) and the gradient \( \beta - U_{yy} \) of the absolute vorticity profile is displayed in panel (b): the absolute vorticity gradient vanishes within the domain as required for instability to occur (Pedlosky 1987, p. 505).

Todling and Ghil (1994) showed, for a similar problem, that observations within the strongest part of the jet are more efficient in reducing analysis errors than observations outside the jet, when the KF is used. In all assimilation experiments performed here, we use an observing network extracted from the radiosonde network of TC94. To create a stringent test, as shown in Fig. 2 we take 12-hour height and wind observations only from those radiosondes positioned outside the jet of Fig. 1a. Parallel lines in Fig. 1a delimit this region. The assumed observation error variances are those of TC94. Also, to isolate the effect of approximating the propagated analysis error covariance we make the perfect model assumption, so that \( Q_{k,k} = 0 \) at all times. The model takes 9 time steps per hour, so that the 12-hour propagator is \( I_{108} \), the single time step unstable dynamics matrix.

Except for experiments in which we take the trailing error covariance to vanish identically, we use here an intentionally crude model for the trailing error covariance matrix analogous to the model error covariance matrix of Cohn and Parrish (1991). We define \( T_{k,k} (or T'_{k,k} or T''_{k,k}) \) to be given by

\[
T_{k,k} = \alpha Z_S \hat{T}_S Z_S^T, \tag{4.2}
\]

where \( \alpha \) is a parameter to be tuned adaptively, \( \hat{T}_S \) is a diagonal matrix of positive elements corresponding to the spectral components of \( T_{k,k} \), and \( Z_S \) is the matrix whose columns are the slow eigenmodes of the unstable dynamics matrix \( \Psi \). That is, \( Z_S \) is determined from the eigendecomposition

\[
\Psi = Z \hat{\Psi} Z^{-1}, \tag{4.3}
\]

where \( \hat{\Psi} \) is the diagonal matrix of eigenvalues of \( \Psi \), and \( Z \) is the matrix of eigenvectors partitioned as

![Fig. 1. Meridionally-dependent unstable Bickley jet (a) and its corresponding absolute vorticity gradient (b), scaled by \( 10^{10} \).](image)

![Fig. 2. Observation network consisting of 33 radiosondes observing outside the jet of Fig. 1a.](image)
Z = (Z_S Z_F),  \quad (4.4)

with the subscripts S and F referring to the slow and fast eigenvectors, respectively. The spectrum $T_S$ is chose in such a way that the $hh$ correlations are roughly Gaussian in physical space; it is identical to the Cohn and Parrish (1991) model error spectrum. Notice that the only time dependence in the model (4.2) for the trailing error covariance comes from the tuning of the parameter $\alpha_t$.

**b. The unstable dynamics case**

1) **Performance of the TC94 schemes**

Before evaluating the performance of the schemes developed in Sections 2 and 3, we evaluate the behavior of one of the schemes studied in TC94, namely the balanced-simplified Kalman filter (balanced-SKF), but now for the unstable dynamics case. Among all the suboptimal schemes in TC94, this one had the best performance for stable dynamics. Recall that this scheme consists of advection of the height-height error covariance matrix and a dynamically-balanced cross-covariance generation of the wind-height and wind-wind cross- and autocovariances. The advective speed (4.1) is a function of latitude and the balanced-SKF for the unstable case uses this expression for advecting the height-height error covariance. We start the experiment from an *almost* asymptotic analysis error covariance matrix obtained from a prior KF run taken up to 10 days.

Figure 3 shows the domain-averaged, expected root-mean-square (ERMS) error in the total energy, in relative units, for two SKF runs. The exploding, unlabeled curve is the result from the balanced-SKF, and the curve labeled T corresponds to a balanced-SKF run in which we attempt to account for the SKF approximation imperfections by adding to the propagated error covariance matrix the covariance matrix defined in (4.2). The parameter $\alpha_t$ is tuned adaptively using the technique of Dee (1995). The KF result is also shown in the figure by the curve labeled KF, and is seen to be in an *almost* periodic steady state.

The balanced-SKF is unable to prevent the estimation errors from growing in this unstable dynamical system. This is a consequence of the fact that the SKF uses incorrect dynamics, which accounts for advection but has no knowledge of the unstable behavior of the system. Evidently the divergence term in the continuity equation, neglected by the SKF, is important in the unstable case. The *adaptive* balanced-SKF, on the other hand, does recognize some of the unstable nature of the system, since it uses the innovation vectors explicitly, although it still diverges. This adaptive result, however, depends on the realization and should be interpreted with caution. In any case, other experiments (not shown) have demonstrated that the qualitative behavior of the adaptive scheme is the same for different realizations. The failure of these schemes is the motivation behind developing the schemes described in Sections 2 and 3.

2) **Performance of the PSF**

The linearity of the dynamical model under consideration and the fixed interval (12 hours) between analysis times allows us to compute the singular values/vectors of the 12-hour propagator off-line and store them for use in the data assimilation experi-
ments. For large-scale nonlinear dynamics the SVD would have to be performed iteratively, for the first few singular vectors, using a Lanczos-type algorithm at each analysis time. The singular values/vectors would be state-dependent, to the same extent that the tangent linear propagator and its corresponding adjoint would be state-dependent.

For our unstable shallow-water dynamics, the singular values of the slow propagators [defined by setting to zero the elements of $\psi$ in (4.3) corresponding to $Z_F$ in (4.4)] for four different time intervals $t_k-t_k-\ell$ are displayed in Fig. 4: 12 hours (solid line), 24 hours (dotted line), 36 hours (dashed line) and 48 hours (dash-dotted line). Only 325 singular values are displayed in panel (a) since this is the number of slow modes of the dynamics (see TC94). The first 36 singular values are larger than unity for all four propagators, while the first 54 are larger than or equal to unity. Detail of the first ten singular values is shown in panel (b). The larger the time interval, the more significant is the difference between the few gravest singular values and the others. A similar plot is given by Buizza (1994), obtained from the tangent linear propagator of a primitive equation model at truncation T21 with 19 layers in the vertical. Buizza’s results also show that a clear distinction between the gravest singular vectors and others occurs at times 36 hours or longer. Since the radiosonde network observes at 12-hour intervals, we use singular values/vectors computed from the 12-hour propagator in the assimilation experiments to follow.

To study the PSF approximation, first we examine the 12-hour forecast error covariance starting from the almost asymptotic analysis error covariance obtained after 10 days of data assimila-
Table 1. Percentage error variance against number of modes retained in the PSF.

<table>
<thead>
<tr>
<th># Singular Modes</th>
<th>Variance (%)</th>
<th># Singular Modes</th>
<th>Variance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>44.1</td>
<td>60</td>
<td>71.0</td>
</tr>
<tr>
<td>20</td>
<td>47.2</td>
<td>70</td>
<td>88.1</td>
</tr>
<tr>
<td>30</td>
<td>53.8</td>
<td>80</td>
<td>91.2</td>
</tr>
<tr>
<td>40</td>
<td>56.4</td>
<td>90</td>
<td>92.3</td>
</tr>
<tr>
<td>50</td>
<td>59.6</td>
<td>100</td>
<td>93.3</td>
</tr>
</tbody>
</table>

Fig. 6. Performance of the partial singular value decomposition filter (PSF): Curves are forecast/analysis expected root-mean-square (ERMS) errors in the total energy, computed from the performance analysis equations (TC94). Curve labeled 10 is the result from a PSF experiment with ten leading singular vectors; curve labeled 10T is for an identical experiment but including a trailing error covariance model with adaptively tuned magnitude; and curves labeled 36 and 54 are for the untuned PSF experiments including 36 and 54 singular vectors, respectively.

Fig. 7. Spectra of error covariances. Solid curve is for the initial analysis error covariance matrix used to start the KF run; the dash-dotted curve is for the forecast error covariance at 10.5 days. Eigenvalues are normalized by $10^6$. The cross-term (c) and (d). The cross-term (e) contributes very little to the total variance (b), the values in panel (e) being almost two orders of magnitude smaller than those in panel (b). This can also be seen by comparing panel (b) with panel (f), which is the sum of the variance fields in panels (c) and (d). Thus, the cross-covariance terms in $T_{k,k-l}$ only redistribute variance as discussed in Section 2, and in fact contribute little to the total variance. Numerical confirmation of the trace property in (2.7) is obtained to machine precision when summing the values of the field in panel (e) over all grid points. This result suggests that modeling the trailing covariance will be an important step in the PSF algorithm, since the leading part is a significant underestimate of the total error variance field, at least with $L=10$, while modeling the leading/trailing cross-covariance may be unnecessary.

Table 1 shows the percentage error in the total energy field accounted for in 12-hour forecasts performed using the PSF with different values of $L$, starting again from the 10-day KF analysis error covariance. The entries in the table are the domain-averaged forecast errors in the total energy obtained by the PSF for $L = 10, 20, \ldots, 100$, normalized by the same quantity obtained using the full KF com-
putation. Only 44.1% of the total error variance is accounted for with 10 modes (see panel (c) of Fig. 5). However, a considerable jump is seen from 50 to 70 modes, with 88.1% of the error variance being accounted for by 70 modes. This indicates that including at least all singular vectors with singular values larger than or equal to unity (54 here) is important. This will be evident in the experiments to follow.

Figure 6 shows the performance of four suboptimal runs utilizing the PSF. These runs again start from the almost asymptotic analysis error covariance obtained from the 10-day KF run. The domain-averaged forecast/analysis ERMS errors in the total energy for the subsequent 10-day assimilation period are displayed in the figure. The top-most curve, labeled 10, gives the performance for a PSF experiment using only the first $L = 10$ singular values/vectors in computing the leading error covariance matrix and no modeling of the trailing error covariance ($T_{k,k-l} = 0$). The PSF with so few singular vectors and no accounting for the trailing error covariance is unstable, with variances growing without bound. The corresponding KF result is shown as the unlabeled (lowest) curve for comparison. The curve labeled 36 gives the PSF result when the first $L = 36$ singular vectors are used, again with $T_{k,k-l} = 0$. These singular vectors correspond to all of those with singular values larger than one (see Fig. 4a). The ERMS curve in this case indicates that this PSF is still slightly unstable (verified with a run to 20 days). The curve labeled 54 refers to the PSF experiment using all 54 singular modes with singular values larger than or equal to unity. A 20-day run showed this to be a stable case with error variances kept bounded. The performance of the PSF with 54 modes is comparable to that of the KF.

Table 2. Percentage error variance against number of modes retained in the PEF.

<table>
<thead>
<tr>
<th># Eigenmodes</th>
<th>Variance (%)</th>
<th># Eigenmodes</th>
<th>Variance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>62.7</td>
<td>60</td>
<td>89.4</td>
</tr>
<tr>
<td>20</td>
<td>72.4</td>
<td>70</td>
<td>91.8</td>
</tr>
<tr>
<td>30</td>
<td>78.6</td>
<td>80</td>
<td>93.8</td>
</tr>
<tr>
<td>40</td>
<td>83.0</td>
<td>90</td>
<td>95.4</td>
</tr>
<tr>
<td>50</td>
<td>86.5</td>
<td>100</td>
<td>96.8</td>
</tr>
</tbody>
</table>

From the results in Fig. 5, we have seen that modeling the trailing error covariance matrix may be important in the PSF algorithm. In Fig. 6 the curve labeled 10T shows the ERMS error evolution for the same PSF experiment indicated by the curve labeled 10, but with a trailing error covariance matrix modeled as described in (4.2), the magnitude parameter $\alpha_t$ being estimated adaptively during the experiment. The result is much improved in comparison with the one obtained when the trailing error covariance is ignored, but the ERMS error still grows. While this adaptive result has to be interpreted with caution, additional realizations (not shown) indicated no qualitative differences. Moreover, adaptively tuning the trailing error covariance model for the PSF with $L = 36$ or 54 modes (results not shown) did not yield any significant improvement over the untuned results displayed in Fig. 6. This leads to the conclusion that accounting for all singular vectors with singular values larger than or equal to unity is more important than our simple accounting for the trailing error covariance.

3) Performance of the PEF

By analogy with looking first at the singular values of the propagator for the PSF scheme, we now examine the spectrum of eigenvalues of a forecast error covariance matrix, to clarify the approximation made by the PEF scheme. Figure 7 shows two covariance spectra from a KF run. Only 325 eigenvalues are shown since this is the number of slow modes in our unstable model. The remaining eigenvalues are all zero, as explained in Cohn and Parrish (1991). The solid line is for the initial analysis error covariance used to start the KF run, and the dash-dotted line is for the forecast error covariance at 10.5 days. The important feature to notice from the figure is that the forecast error covariance spectrum has only about 10 significantly nonzero eigenvalues. A PEF approximation could be aimed toward capturing only these first few modes.

Table 2 is analogous to Table 1 and it shows, starting from the 10-day KF run analysis error covariance, the percentage of forecast error variance explained by the first $L = 10, 20, \ldots, 100$ eigenmodes of $S_k^f$ at 10.5 days, corresponding to a 12-hour forecast with the PEF. That is, the table displays the ratio of the sum of the first $L$ eigenvalues to the sum of all 325 nonzero eigenvalues. It is clear that most of the variance is explained by relatively few eigenmodes. When compared to the results in Table 1 for the PSF, we see the PEF capturing more variance with fewer modes. Moreover, the amount of variance explained by the PEF increases smoothly with the number of modes $L$.

In Fig. 8 we plot maps of the 12-hour forecast error variance in the energy density field obtained
by keeping \( L = 10, 36 \) and 54 modes in the PEF scheme. The initial analysis error covariance is the same used to produce the maps in Fig. 5, that obtained at the end of a 10-day assimilation experiment using the KF, the corresponding variance map being shown in Fig. 5a. There is little difference among all three maps in Fig. 8, as one might expect from Table 2. The maps in Fig. 8 also agree rather well with the exact result displayed in Fig. 5b. There is, however, a striking difference between Fig. 8a for the PEF with \( L = 10 \) eigenmodes and Fig. 5c for the PSF with \( L = 10 \) singular modes.

The PEF captures much more error variance than the PSF for fixed \( L \), at similar cost. This also indicates that accounting for the trailing part in the PEF might be less important than in the PSF.

Figure 9 shows the performance evaluation for the PEF starting from the 10-day KF result, to be compared with Fig. 6 for the PSF, discussed in Section 4b.2. Results for the PEF with \( L = 10 \) and 36 are indicated in Fig. 9 by the corresponding labels. The KF result (unlabeled) is also displayed, nearly superimposed on the \( L = 36 \) PEF result. Adaptive tuning of an additive error covariance matrix to account for the trailing part has shown no significant improvement over these results. In contrast to the PSF result with 36 singular vectors, the PEF result with 36 modes does not diverge.

Experiments performed with the alternative PEF-type scheme described in Section 3 (not shown) produced results very similar to those just discussed. A full comparison between these two schemes, for example documenting the influence of observation frequency and pattern, has not yet been carried out.
4) Non-steady state initial covariance

The results of the previous subsections all started from an almost asymptotic analysis error covariance matrix. Here we study the more transient behavior of the PSF and PEF when the assimilation starts from an initial error covariance matrix similar to that of TC94. The modal structure of this covariance matrix is defined by (4.2), with a prescribed multiplicative parameter.

In addition to the PSF and PEF schemes, we also evaluate the performance of a reduced resolution filter (RRF), in which the covariances are evolved at a lower resolution than that used to evolve the state vector. This covariance evolution differs from that of Fukumori and Malanotte-Rizzoli (1995) in that it uses directly a coarse-grid discretization of the governing differential equations. The RRF scheme produces a gain matrix at low resolution which is interpolated using cubic splines (spline interpolant with periodic boundary conditions in the east-west direction and an Akima spline in the north-south direction) to the full resolution of the model. These full resolution gains are then initialized by using the energy-projector described in TC94 to filter out gravity waves introduced by the interpolation. The initialized gains are used to update the state vector and to compute the actual analysis error covariance matrices for the performance analysis. The full model resolution is $25 \times 16$, and in the experiments to follow we vary the resolution in the meridional direction while keeping the resolution in the zonal direction at approximately half of the original resolution, i.e., 13 grid points. Varying the meridional resolution is a more significant change due to the meridional dependence of the basic state velocity profile, and the resulting dependence of the instability properties of the model on the resolution. The computational expense of the RRF, relative to that of the full KF, is about $(m/n)^2$, $m$ being the number of degrees of freedom of the reduced resolution dynamics.

Figure 10 shows the ERMS error in the total energy for the KF, the RRF with $13 \times 12$ resolution, the PSF with $L = 10$, and the PEF with $L = 10$. All the SOS's perform poorly, and in fact diverge. The RRF performs reasonably well for a short time, but ultimately does not prevent the growth of unstable modes because, as we shall see later, it cannot resolve them (see also Yanai and Nitta 1968). Its cost greatly exceeds the cost of the PSF and PEF in this case. The PSF and PEF behave similarly to each other for the first few days, with the latter performing slightly better than the former, but as time progresses they are unable to prevent the unstable modes from growing, the PSF analyses ultimately being worse than the PSF forecasts. Examination of the curves S10 and E10 shows that the PEF makes much better use of the observations than does the PSF, which shows no impact of the observations during most of the 10-day period.

In Fig. 11 we show results analogous to those of Fig. 10, but with increased RRF resolution and number of modes kept in the PSF and PEF. The RRF now has resolution $13 \times 16$. The PSF retains 54 singular modes, which includes all those with singular values larger than or equal to one (see Fig. 4), and the PEF computes 54 eigenmodes, which in the almost asymptotic KF regime would account for about 65 % and 88 % of the total variance in the energy (see Tables 1 and 2), respectively. The KF result (unlabeled bottom curve) is also plotted as a basis for comparison. The most striking result is the performance of the RRF, which is now extremely close to that of the KF itself. This RRF has the same meridional resolution as the underlying model, though half the zonal resolution, and therefore resolves the unstable jet fully. Contrary to the behavior presented in Fig. 10, both the PSF and the PEF are now stable. For the PSF, this indicates once again that accounting for all growing and neutrally stable modes is fundamental for its stability; for the PEF, it is not clear what criterion determines its stability properties. The PSF performs better.
than the PEF at 10 days, although the transient behavior (up to 5 days) of the PEF is better than that of the PSF. The transient behavior may be more important than the long-term behavior in nonlinear systems. The behavior of these schemes is considerably more complex than what Tables 1 and 2 seem to suggest, when the approximations are employed recursively in assimilation cycles.

Next, we compare all three SOS’s with adaptive tuning. In the case of the PSF and PEF we take the trailing part to be the covariance matrix described previously, with magnitude adaptively tuned. In the RRF case, we add to the propagated error covariance a similar “trailing” covariance matrix but generated at the resolution of the RRF. Its magnitude is also adaptively tuned. Figure 12 displays the results obtained for the same three SOS’s of Fig. 10, but with adaptive tuning. The KF result is shown once again for reference. Comparing Fig. 12 with Fig. 10 shows that adaptive tuning improves the performance of all the schemes dramatically, although the PSF and RRF still diverge. As in Fig. 10, the PEF shows the best performance, but now its errors are kept bounded. In fact, the performance of the adaptively tuned PEF is fairly close to that of the full KF. It is also far superior to that of the non-adaptive PEF with 54 modes (see Fig. 11).

c. The stable dynamics case

The motivation for developing the PSF and PEF algorithms was the failure of the SOS’s introduced in TC94 for the present unstable dynamical model. Thus one question left to be answered is whether these new schemes also perform appropriately for stable dynamics. In this section we return to the TC94 dynamics, of a shallow water model linearized about a constant velocity profile for which $U_0 = 25 \text{ ms}^{-1}$ and $U_1 = 0$ in (4.1). We start the assimilation runs here using the initial error covariance matrix of TC94, but we use the observation network introduced in Section 4a. The runs are all taken up to 10 days.

In Fig. 13 we show the performance analysis results, analogously to Figs. 10–12, for three different approximate schemes and compare them against the corresponding KF result. The SOS’s used are the PSF with 36 singular vectors, the PEF with 36 eigenvectors, the RRF with $13 \times 12$ resolution, and the balanced-SKF, all labeled accordingly in the figure. The unlabeled curve refers to the KF result. There are no surprises here, in view of the corresponding results for the unstable dynamics case shown in Fig. 11. All the schemes perform roughly as well as they did in the unstable case, with the exception of the balanced-SKF, which diverged in the unstable case but has the best performance in the stable case. The schemes are all stable, in contrast to some of the SOS’s tested in TC94.

No trailing part is taken into account in the results of Fig. 13. In Fig. 14 we show results for the tuned PSF and PEF, with 36 singular vectors and eigenvectors, respectively. We also show the RRF result at $13 \times 16$ resolution, as well as the KF result. The RRF uses no tuning since this gave no significant improvement over the untuned experiment. The tuned SKF also gave no improvement over the untuned SKF, and is not shown in Fig. 14. Comparison with Fig. 13 indicates that the $13 \times 16$ RRF has better performance than the SKF, although the RRF with this resolution involves more computation than the SKF. The PSF and PEF results are substantially improved over the untuned results. They also present better performance than the RRF with $13 \times 12$ resolution, and are about equal in performance with the SKF. However, no SOS matches the performance of the RRF with complete meridional resolution, as was also the case for unstable dynamics.
5. Conclusions

We have introduced two new approximate schemes for error covariance propagation in the Kalman filter, namely the partial singular value decomposition filter (PSF) and the partial eigendecomposition filter (PEF). We have studied their performance for a barotropically unstable model, and compared their performance against that of the full Kalman filter as well as a reduced resolution filter (RRF). Based on the experiments conducted here, our main conclusions are as follows.

- The RRF diverges unless it fully resolves the barotropic instability, in which case it performs quite well.
- The PSF must account for all modes with singular values larger than or equal to unity, for otherwise the PSF diverges.
- A stability criterion for the PEF in the absence of adaptive tuning is not obvious, although with adaptive tuning the PEF is stable and performs well with very few modes retained.
- In most cases, adaptive tuning enhances the performance of all three schemes considerably; in very few cases, it makes no significant improvement.

An adaptively tuned PEF, perhaps combined with the RRF idea, shows promise for practical implementation with large-scale, nonlinear models. The fraction of propagated analysis error variance it accounts for directly could be calculated through a randomized trace estimation scheme (Girard 1991). Considerable effort would have to be undertaken to learn how best to approximate the trailing error covariance.

One great advantage of both the PSF and the PEF is that they only require operation of the analysis error covariance on a limited number of vector quantities, thus avoiding the need for calculating the entire analysis error covariance matrix. The Physical-space Statistical Analysis System (da Silva et al. 1995) is being designed to provide the necessary operator for either one of these schemes.

Acknowledgments

It is a pleasure to thank A.M. da Silva, D.P. Dee, I.M. Navon, and J.J. Tribbia for helpful discussions throughout the course of this research. D. Sorensen provided the iterative eigensystem package ARPACK used here. The numerical results were obtained on the Cray C90 through cooperation of the NASA Center for Computational Sciences at the Goddard Space Flight Center. This research was supported by the NASA EOS Interdisciplinary Project on Data Assimilation (SEC) and by a fellowship from the Universities Space Research Association (RT).

References


安定及び不安定な力学のための近似的なデータ同化スキーム

Stephen E. Cohn
(Data Assimilation Office, NASA/Goddard Space Flight Center, U.S.A.)

Ricardo Todling
(Universities Space Research Association, NASA/GSFC/DAO, U.S.A.)

安定及び不安定な力学に対する、2つの副次的な最適データ同化スキームが提出される。最初のスキームである「部分的特異値分解フィルター」は、線形モデルの最も卓越する特異モードに基づいている。二番目のスキームである「部分的固有分解フィルター」は、解析誤差と共分散行列の時間発展において最も卓越する固有モードに基づいている。2つのスキームは、それぞれのモードを計算法するために、ランダム・アルゴリズムに似た反復法を用いる。

これらのスキームの性能が、安定なビックレイ・ジェットを基本場とする、線形浅水モデルに対して評価される。結果は「縮小分解能フィルター」の性能と対比される。このフィルターでは、状態ベクトルを更新するために用いる利得が、系の状態その自身を時間発展させる力学ではなく、低次元の力学から計算される。結果は、カルマン・フィルターによって与えられる正確な結果とも対比される。これらのスキームは、安定な力学の場合に対しても検証される。

2つの新しい近似的な同化スキームは、計算された比較的少数のモードを用いても、満足に揺る舞うことが示される。これらの2つの近似スキームに対して、モデル化された誤差共分散の副次的成分に適切なチューニングを行うと、性能が向上し、用いられている近似的による誤差が減少する。