The University of Reading

Very Large Inverse Problems in Atmosphere and Ocean Modelling

C. Johnson[1], N.K.Nichols[2] and B.J. Hoskins[2]

NUMERICAL ANALYSIS REPORT 5/04

[1]*Mesa Lab*
National Center for Atmospheric Research
PO Box 3000
Boulder, CO 80307
U.S.A.

[2]*Departments of Mathematics*
and Meteorolgy
The University of Reading
Box 220, Whiteknights
Reading, Berkshire RG6 6AX,

Presented at the ICFD conference on Numerical Methods for Fluid Dynamics,
Oxford, UK, 29 March - 1 April 2004.

Department of Mathematics

Very Large Inverse Problems in Atmosphere and Ocean Modelling *

C. Johnson, N.K. Nichols and B.J. Hoskins

Abstract

For the very large nonlinear dynamical systems that arise in a wide range of physical, biological and environmental problems, the data needed to initialize a numerical forecasting model are seldom available. To generate accurate estimates of the expected states of the system, both current and future, the technique of 'data assimilation' is used to combine the numerical model predictions with observations of the system measured over time. Assimilation of data is essentially an ill-posed inverse problem. In four dimensional variational assimilation schemes, the dynamical model equations provide constraints that act to spread information into data sparse regions, enabling the state of the system to be reconstructed accurately. The mechanism for this is not well understood. Singular value decomposition techniques are applied here to analyse the critical features in this process. Simplified models are used to demonstrate how information is propagated from observed regions into unobserved areas. The impact of the size of the observational noise and the temporal position of the observations is examined. The best signal-to-noise ratio needed to extract the most information from the observations is estimated using Tikhonov regularization theory.

1 INTRODUCTION

Accurate prediction of the behaviour of very large evolutionary systems requires both accurate numerical models for simulating the system dynamics and accurate data for initializing the forecast. In practice, precise data describing the current state of a system are not available, and uncertainties in the initial data lead to significant errors between the predicted states

---

*Presented at the ICFD International Conference on Numerical Methods for Fluid Dynamics, Oxford, March, 2004

and the actual states of the system. To generate improved estimates of the expected states, both current and future, the technique of 'data assimilation' is used to combine numerical model predictions with observations of the system measured over time. The data assimilation problem can be expressed as: *Given a discrete model of the dynamics of a system, a (noisy) estimate of the current state and (noisy) observations of the system over time, find accurate estimates of the system states.*

Variational data assimilation techniques are attractive because they deliver the best statistically linear unbiased estimate of the system states given the available observations and their error covariances [1, 2]. The problem is formulated as an optimization problem where the objective function measures the mismatch between the model predictions and the observed system states, weighted by the inverse of the error covariance matrices. In four-dimensional schemes, the objective function is minimized over a time interval, and the model equations are treated as strong constraints [3]. These variational assimilation schemes are applicable to a wide range of physical modelling problems, including oil recovery, coastal flow and sediment transport, flood prediction, and traffic flow problems. With the proliferation of observational data from expensive satellites and other instruments, techniques of data assimilation are needed increasingly to extract the best value from the information provided.

For the very large systems that arise in meteorology and oceanography, the data assimilation problem is an ill-posed inverse problem. The available observations are not generally sufficient to determine all of the degrees of freedom in the problem, and often there are data sparse areas where good state estimates are needed. In four dimensional variational assimilation schemes (4DVar), the dynamical model equations act to spread information into unobserved regions [4], but the mechanism for this is not well understood. Here we apply singular value decomposition (SVD) techniques to analyse the critical features in this process. Idealized model studies are used to demonstrate how information is propagated from observed regions into unobserved regions.

Using identical twin experiments with a two-dimensional Eady model [5], we show that by observing only the lower level temperature at two points in time, 4DVar can reconstruct the upper level temperature wave needed for the growth or decay of a baroclinic wave. Using the SVD, the optimal state estimate is written as a linear combination of the right singular vectors of the observability matrix. The properties of the SVD are then used to understand how information is propagated from observed regions to unobserved regions. The impact of varying the relative weight given to the background (predicted) states, the noise on the observations and the position of the observations in the assimilation time interval is examined. By writing the problem in the form of a Tikhonov regularization problem, it is shown that the best signal-to-noise ratio needed to extract the most

information from the observations can be determined.

In the next section we present the variational data assimilation method. In Section 3 we describe the test model and show experimental results obtained by variational assimilation. In Section 4 the information content of the observations is analysed and the critical features of the mechanism for reconstructing the states of the system are examined. In Sections 5 and 6, the application of Tikhonov regularization to the problem is described and conclusions are drawn.

2 FOUR DIMENSIONAL VARIATIONAL DATA ASSIMILATION

Variational Data assimilation schemes are described here for a system modelled by the discrete nonlinear equations

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k), \quad k = 0, \dots, N - 1, \tag{1}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the model state vector and $\mathbf{f}_k : \mathbb{R}^n \times \mathbb{R}^{m_k} \to \mathbb{R}^n$ is a nonlinear function describing the evolution of the states from time t_k to time t_{k+1}.

The observations are related to the system state by the equations

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \boldsymbol{\delta}_k, \quad k = 0, \dots, N - 1, \tag{2}$$

where $\mathbf{y}_k \in \mathbb{R}^{p_k}$ is a vector of p_k observations at time t_k and $\mathbf{h}_k : \mathbb{R}^n \to \mathbb{R}^{p_k}$ is a nonlinear function that includes transformations and grid interpolations. The observational errors $\boldsymbol{\delta}_k \in \mathbb{R}^{p_k}$ are assumed to be unbiased, serially uncorrelated, Gaussian random vectors with covariance matrices $R_k \in \mathbb{R}^{p_k \times p_k}$.

A prior estimate, or 'background estimate,' $\mathbf{x}_0^b$ of the initial state $\mathbf{x}_0$ is assumed to be known and the initial random errors $(\mathbf{x}_0 - \mathbf{x}_0^b)$ are assumed to be Gaussian with covariance matrix $B_0 \in \mathbb{R}^{n \times n}$. The observational errors and the errors in the prior estimates are assumed to be uncorrelated.

The aim of the data assimilation is to find the maximum likelihood Bayesian estimate of the system states given the observations and the prior estimate of the initial state. This problem reduces to minimizing the square error between the model predictions and the observed system states, weighted by the inverse of the covariance matrices, over the assimilation interval [1]. The model is assumed to be 'perfect' and the system equations are treated as strong constraints on the objective function. The model states that satisfy the system equations are uniquely determined on the assimilation interval by the initial states of the system. The initial states can thus be treated as the required control variables in the optimization. The data assimilation problem is defined explicitly as follows.

Problem 1 *Minimize, with respect to* $\mathbf{x}_0$, *the objective function*

$$\mathcal{J} = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T B_0^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}\sum_{j=0}^{N-1}(\mathbf{h}_j(\mathbf{x}_j) - \mathbf{y}_j)^T R_j^{-1}(\mathbf{h}_j(\mathbf{x}_j) - \mathbf{y}_j) \quad (3)$$

subject to the system equations (1).

In practice the constrained minimization problem is solved iteratively by a gradient method. The problem is first reduced to an unconstrained problem using the method of Lagrange. Necessary conditions for the solution to the unconstrained problem then require that a set of adjoint equations together with the system equations must be satisfied. The adjoint equations are given by

$$\boldsymbol{\lambda}_N = 0, \quad (4)$$

$$\boldsymbol{\lambda}_k = F_k^T(\mathbf{x}_k)\boldsymbol{\lambda}_{k+1} - H_k^T R_k^{-1}(\mathbf{h}_k(\mathbf{x}_k) - \mathbf{y}_k), \quad k = N-1, \ldots, 0, \quad (5)$$

where $\boldsymbol{\lambda}_k \in \mathbb{R}^n$, $j = 0, \ldots, N$, are the adjoint variables and $F_k \in \mathbb{R}^{n \times n}$ and $H_k \in \mathbb{R}^{n \times p_k}$ are the Jacobians of $\mathbf{f}_k$ and $\mathbf{h}_k$ with respect to $\mathbf{x}_k$.

The gradient of the objective function (3) with respect to the initial data $\mathbf{x}_0$ is then given by

$$\nabla_{\mathbf{x}_0} \mathcal{J} = B_0^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) - \boldsymbol{\lambda}_0. \quad (6)$$

At the optimal, the gradient (6) is required to be equal to zero. Otherwise this gradient provides the local descent direction needed in the iteration procedure to find an improved estimate for the optimal initial states. Each step of the gradient iteration process requires one forward solution of the model equations, starting from the current best estimate of the initial states, and one backward solution of the adjoint equations. The estimated initial conditions are then updated using the computed gradient direction. This process is expensive, but it is operationally feasible, even for very large systems, such as weather and ocean systems, which may involve as many as 10^7 state variables.

3   APPLICATION TO THE EADY MODEL

The success of the 4DVar assimilation technique is largely due to the action of the dynamical model equations, which spread information from the available observations into sparse data regions, enabling the states to be reconstructed accurately everywhere in the domain. To investigate the critical features of the reconstruction process, we conduct identical twin experiments using a simple two-dimensional Eady model of baroclinic instability. The model describes the evolution of perturbations to a basic state, given by a linear zonal wind shear with height in a domain between two rigid horizontal

boundaries. The density, static stability and Coriolis parameter are taken to be constants, and it is assumed that the interior quasi-geostrophic potential vorticity is zero.

Perturbations to the basic state are advected zonally by the basic shear flow. The system dynamics are described by the non-dimensional equations

$$(\frac{\partial}{\partial t} + z\frac{\partial}{\partial x})b = \frac{\partial \psi}{\partial x}, \qquad \text{for} \ \ z = \pm\frac{1}{2}, \quad x \in [0, X], \tag{7}$$

where b is the buoyancy and the geostrophic streamfunction ψ satisfies

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial z^2} = 0, \qquad z \in [-\frac{1}{2}, \frac{1}{2}], \quad x \in [0, X], \tag{8}$$

with boundary conditions

$$\frac{\partial \psi}{\partial z} = b, \qquad \text{for} \ \ z = \pm\frac{1}{2}, \quad x \in [0, X]. \tag{9}$$

The buoyancy and streamfunction are assumed to be periodic in x on $[0, X]$.

3.1 Experiments

The aim of the experiments is to reconstruct the buoyancy wave on the upper boundary of the region from observations of the buoyancy on the lower boundary at the beginning and end of the assimilation interval using 4DVar. The model equations are obtained by discretizing the system equations (7) - (9) using a leap-frog advection scheme with 11 vertical levels and 40 grid points in the horizontal. Perfect observations representing the 'truth' are generated by model runs over a time interval corresponding to 6 hours, initiated with the most rapidly growing (or decaying) normal mode of the system. The initial fields have a tilt with height that is associated with vertical coupling between the upper and lower waves, leading to exponential growth (or decay) of the solution. Uncorrelated random noise with variance $\sigma_o = 1$ is added to the observations. The prior estimate of the state at time t_0, known as the initial *background state*, is equal to the true state with a phase shift. This represents a displacement in the estimate of the current system state. The resulting background errors are assumed to be random and uncorrelated with variance σ_b. In the objective function (3) of the 4DVar scheme, the covariance matrices are thus defined to be $B_0^{-1} = \sigma_b^{-2}I_n$ and $R_j^{-1} = \sigma_o^{-2}I_p$, $j = 0, 1$. The error variance ratio is defined to be $\mu^2 = \sigma_o^2/\sigma_b^2$.

In practice, the error characteristics of the observations are generally known quite well from the instrumentation data and the positioning of the observations. Much less confidence can be placed in the error statistics assumed for the initial background state. We investigate here the effects on the estimated states, known as the *analysis*, produced in 4DVar assimilation by changes in the variance ratio and in the positions of the observations in time.
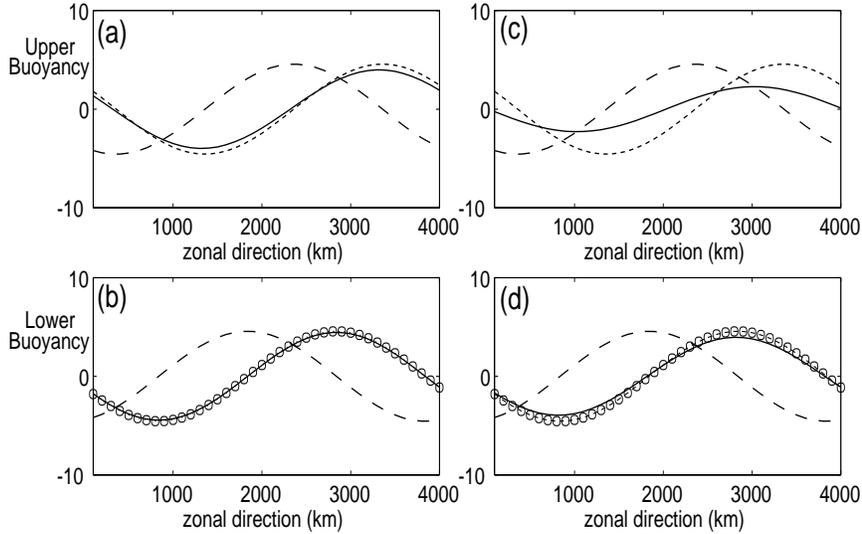
Figure 1: 4DVar analyses (solid line) shown at the final time of the assimilation interval in the cases (a)-(b) $\mu^2 = 0.1$ and (c)-(d) $\mu^2 = 0.01$. The truth and the background state are shown by the dotted and the dashed lines, respectively. In both cases perfect observations, shown by the circles, are given at the beginning and end of the interval.

3.2 Results

In Figure 1, the background and analysis of the buoyancy are shown in the case where the observations are exact. The variance ratio is selected to be $\mu^2 = 0.01$ (Figure 1(a)) and $\mu^2 = 0.1$ (Figure 1(b)). The lower boundary wave, where the solution is measured, is estimated accurately in both cases. The upper boundary is also reconstructed accurately in the first case, but where a larger weight is placed on the background, the amplitude and phase of the upper wave are estimated less accurately, and the growth rate is no longer determined well. We see that weighting the background too heavily causes useful information to be lost.

In Figure 2, the effects of noise on the background and analysis are shown in the cases where $\mu^2 = 0.08$ and $\mu^2 = 0.01$. In the first case the weights are statistically correct and the lower wave is reconstructed accurately, but the reconstructed upper wave contains phase and amplitude errors, again due to the weighting on the background term. If less weighting is applied to the background now, however, the noise on the observations leads to small oscillations in the reconstructed lower wave, and causes a non-physical wave to be generated by the assimilation on the upper boundary. We see that with less weighting on the background, the results become sensitive to the noise in the observations.

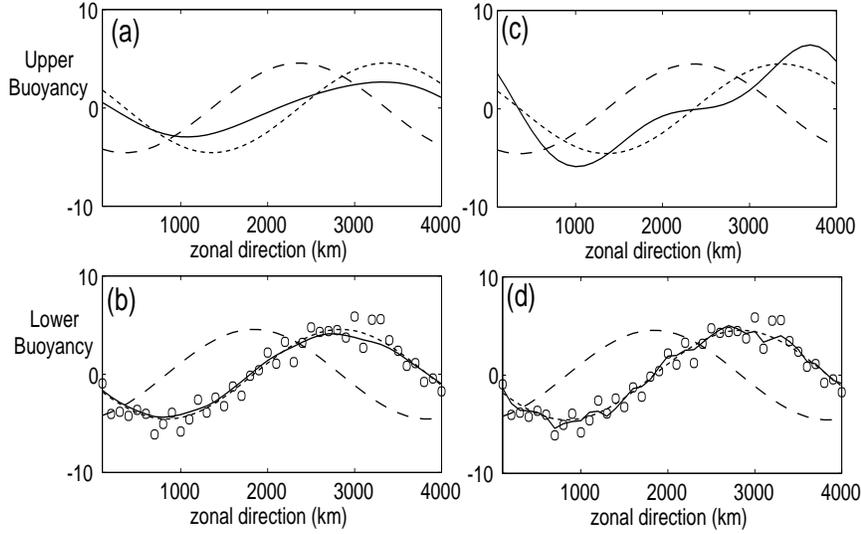The effects of positioning the observations at different times in the as-

Figure 2: 4DVar analyses (solid line) shown at the final time of the assimilation interval in the cases (a)-(b) $\mu^2 = 0.08$ and (c)-(d) $\mu^2 = 0.01$. The truth and the background are shown by the dotted line, and the dashed line, respectively. In both cases, observations, shown by the circles, are given at the beginning and end of the interval and contain random noise with a Gaussian distribution and standard deviation $\sigma_o = 1$.
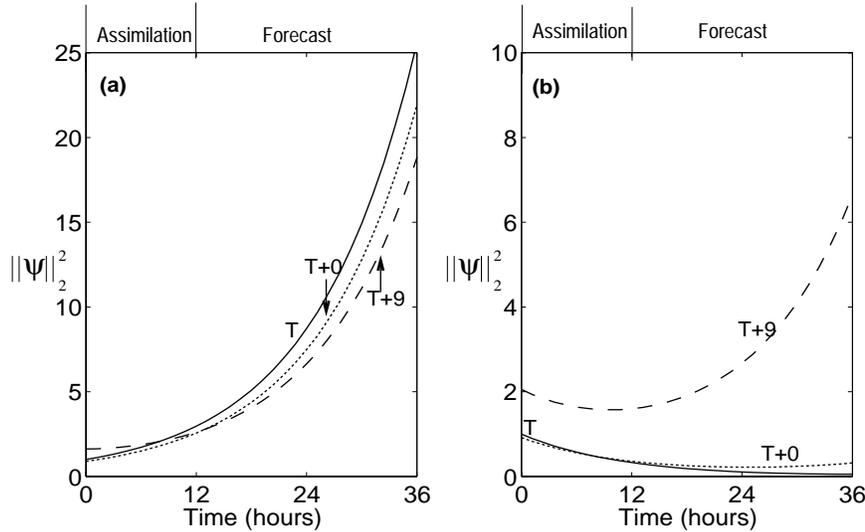

Figure 3: Euclidean norm of the streamfunction for 4DVar analyses and forecasts in the case of the most rapidly (a) growing and (b) decaying normal modes. Observations are given at $t = t_f$, the end of the assimilation interval and at either $t = 0$ (dotted) or at $t = (3/4)t_f$ (dashed). The noise on the observations has standard deviation $\sigma_o = 1.0$ and the variance ratio is specified as $\mu^2 = 0.04$. The 'truth' (solid) is also shown.

8

similation interval are shown in Figure 3. The observations are assimilated over an interval of length t_f, corresponding to twelve hours, and a forecast over a time interval corresponding to 24 hours is produced from the analysis at time t_f, the end of the assimilation interval. Observations at the end time $t = t_f$ are used, together with observations at either the initial time $t = 0$, or at the time $t = (3/4)t_f$. The norm of the streamfunction is plotted for the case where the model is initiated by a growing mode (Figure 3(a)) and by a decaying mode (Figure 3(b)). It can be seen that the results of the assimilation are more accurate in both cases for observations more widely separated in time. Moreover, when the observations are close together in time, the decaying mode is very badly reconstructed, and in the forecast the analysis is actually growing rather than decaying.

3.3 Summary

The conclusions of the experiments are summarized as follows.

(i) Weighting the background state too heavily may filter information needed to reconstruct the state in unobserved regions.

(ii) In unobserved regions, the analysis may be sensitive to noise in the observations if the background is not weighted heavily enough.

(iii) Selecting the appropriate value for $\mu^2 = \sigma_o^2/\sigma_b^2$ is critical for extracting the maximum amount of useful information from the observations.

(iv) A better analysis is achieved if the observations are placed as far apart as possible in the assimilation time interval.

In the next section we analyse the mechanism producing these effects, using a singular vector approach.

4 INFORMATION CONTENT OF OBSERVATIONS IN 4D-VAR

To analyse the critical features in the 4DVar assimilation process, we use a singular value decomposition technique. The model and observation operators are assumed to be linear and the prior (background) error and observational errors are assumed to be uncorrelated with fixed variance. We write

$$\mathbf{x}_{k+1} = M(t_{k+1}, t_k)\mathbf{x}_k, \tag{10}$$

$$\mathbf{y}_k = H_k\mathbf{x}_k + \boldsymbol{\delta}_k, \tag{11}$$

for $k = 0, \ldots, N - 1$. The objective function (3) is written as

$$\mathcal{J} = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T B_0^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}(\hat{\mathbf{H}}\mathbf{x}_0 - \hat{\mathbf{y}})^T \hat{\mathbf{R}}^{-1}(\hat{\mathbf{H}}\mathbf{x}_0 - \hat{\mathbf{y}}), \tag{12}$$

9

where

$$\hat{\mathbf{H}} = \left[H_0^T, (H_1 M(t_1, t_0))^T, \ \ldots, \ (H_{N-1} M(t_{N-1}, t_0))^T \right]^T, \qquad (13)$$

$$\hat{\mathbf{y}}^T = \left[\mathbf{y}_0^T, \ \mathbf{y}_1^T, \ \ldots, \ \mathbf{y}_{N-1}^T \right], \qquad (14)$$

and $\hat{\mathbf{R}}$ is a block diagonal matrix with diagonal blocks equal to R_j. The matrix $\hat{\mathbf{H}}$ is known as the *observability* matrix ([6, 7]). The solution to the optimization problem is then given explicitly by

$$\mathbf{x}_0 = \mathbf{x}_0^b + \left(B_0^{-1} + \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}} \right)^{-1} \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{d}}, \qquad \hat{\mathbf{d}} = \left(\hat{\mathbf{y}} - \hat{\mathbf{H}} \mathbf{x}_0^b \right). \qquad (15)$$

4.1   Singular Value Decomposition

We assume now that $B_0 = \sigma_b^2 I$ and $\hat{\mathbf{R}} = \sigma_o^2 I$ and define the singular value decomposition of the observability matrix $\hat{\mathbf{H}}$ to be

$$\hat{\mathbf{H}} = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^T, \qquad (16)$$

where $\mathbf{\Lambda} = \mathrm{diag}\{\lambda_j\}$. The scalars λ_j are the singular values of $\hat{\mathbf{H}}$, and the left and right singular vectors $\mathbf{v}_j$ and $\mathbf{u}_j$ are given by the columns of $\mathbf{V}$ and $\mathbf{U}$, respectively. Applying the SVD in (15) enables us to write the optimal analysis as

$$\mathbf{x}_0 = \mathbf{x}_0^b + \sum_j \frac{\lambda_j^2}{\mu^2 + \lambda_j} \frac{\mathbf{u}_j^T \hat{\mathbf{d}}}{\lambda_j} \mathbf{v}_j. \qquad (17)$$

The increments made to the prior estimate x_0^b by the assimilation process are thus given by a linear combination of the right singular vectors of $\hat{\mathbf{H}}$, weighted by the two factors

$$f_j = \frac{\lambda_j^2}{\mu^2 + \lambda_j^2} \ , \qquad c_j = \frac{\mathbf{u}_j^T \hat{\mathbf{d}}}{\lambda_j} \ . \qquad (18)$$

If there is no background (prior) estimate constraining the solution, so that $\mu^2 = 0$, then the 'filter factor' $f_j = 1, \ \forall j$, and the weight c_j given to each singular vector in the increment is proportional to the angle between the 'innovation' vector $\hat{\mathbf{d}}$ and the corresponding left singular vector $\mathbf{u}_j$. Large c_j indicates that the corresponding singular vector contributes significantly to the correct reconstruction of the system state.

If a background (prior) estimate is given, so that $\mu^2 > 0$, then the weighting on each singular vector is reduced by the corresponding filter factor f_j. For noisy observations, where $\mu^2 >> \lambda_j^2$, the contribution to the analysis from the singular vector is damped and the observational information is strongly filtered. This filtering is a vital aspect of the assimilation process, as both background and observations contain errors. The choice of the specified value of the variance ratio μ^2 is then crucial to enable the extraction of the maximum information available in the observations.
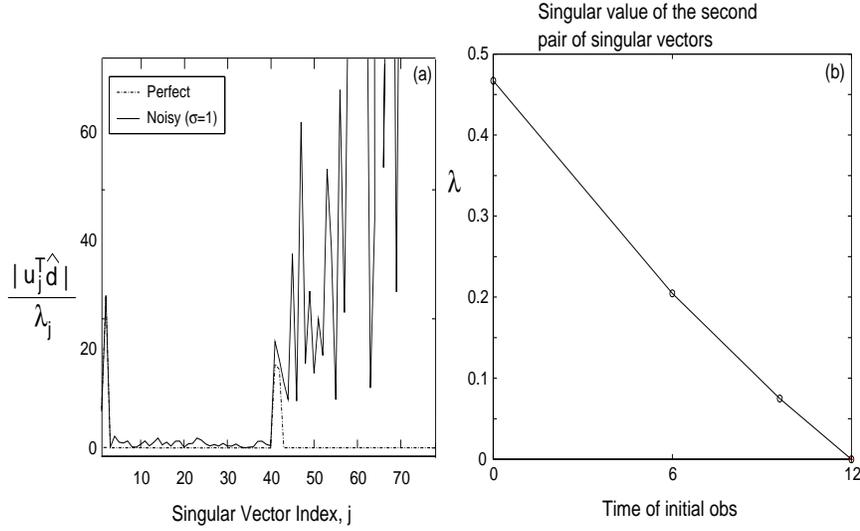
Figure 4: Figure (a) shows for the Eady model the values of the coefficients c_j (as defined in (18)) for perfect (dashed line) and noisy (thin solid line) observations taken at the beginning and end of the assimilation interval. Figure (b) shows the singular values of the second pair of significant singular values plotted as a function of the time when the first set of observations are made. In all cases, the final observations are made at the end of the assimilation window.
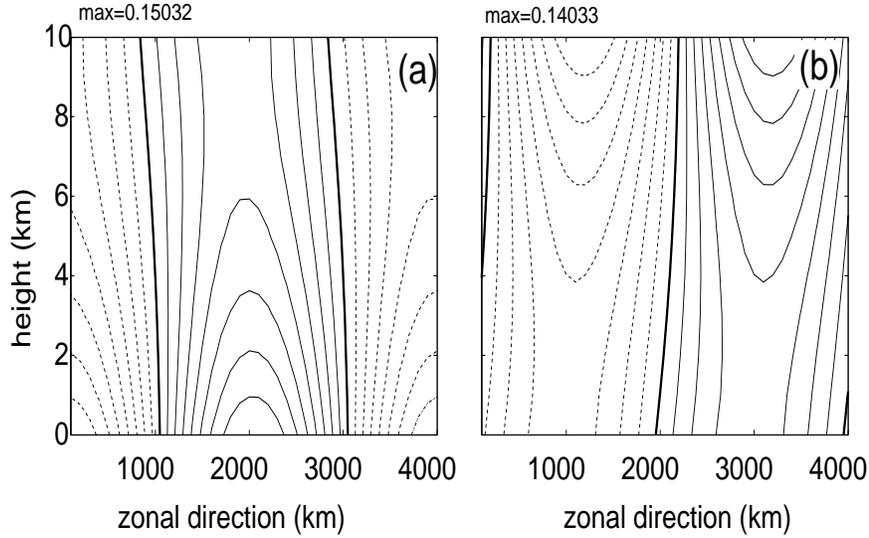

Figure 5: The streamfunction field for the right singular vectors corresponding to the two pairs of singular vectors that are significant in the reconstruction of the state of the Eady system. The corresponding singular values for the two cases are (a) $\lambda = 1.45$ and (b) $\lambda = 0.27$.

11

4.2   Interpretation of Experimental Results

We now apply the SVD theory to interpret the experimental results of Section 3. In Figure 4(a) we show the coefficients c_j for the singular value decomposition of the Eady model. For perfect observations there are two pairs of singular vectors that are significant for the reconstruction of the solution. The corresponding right singular vectors are shown in Figure 5. The singular vectors associated with the larger pair of singular values are seen to contribute primarily to the accurate estimation of the lower boundary wave, where the observations are given. The second set, associated with the smaller pair of singular values, contributes to the accurate reconstruction of the upper wave. For noisy observations, the weights c_j are seen to grow as the singular values decay. If the corresponding singular vectors are not sufficiently filtered, then the estimated states may be very inaccurate due to the noise. In this case, however, unless the variance ratio μ^2 is carefully specified, important information contained in the second set of significant singular vectors may be filtered out, causing the reconstructed upper wave to lose accuracy as seen in Figures 1 and 2.

In Figure 4(b) we show the magnitude of the second set of singular values as a function of the time when the first observations are made within an assimilation interval corresponding to 12 hours. In all cases the second (final) set of observations is made at the end of the interval. We see that the farther apart that the observations are taken, the larger is the second pair of significant singular values. Therefore, important information in the observations is less likely to be lost if the observations are taken as far apart as possible in time, as seen in Figure 3.

4.3   Summary

We may summarize the results of the theoretical analysis as follows.

 (i) The right singular vectors with small singular values contain information needed to reconstruct the state in unobserved regions.

 (ii) The background state is needed to filter the right singular vectors that correspond to noise, but may also filter significant information needed to reconstruct the states accurately unless the variance ratio μ^2 is specified carefully.

 (iii) The singular values associated with the singular vectors that contain significant information in unobserved regions increases as the observations are moved farther apart in time, thus increasing the useful information that can be extracted from the observations.

5 TIKHONOV REGULARIZATION

In Section 4 we have demonstrated the importance of the value of the variance ratio μ^2, between the variances of the background and observational errors, in maximizing the information that can be extracted from the observations. Good choices for μ^2 can be determined by using Tikhonov regularization theory [8].

We first reformulate the objective function (3) for the variational assimilation problem by making a change of variable. We let C_B and $\mathbf{C}_R$ be such that $B_0 = \sigma_b^2 C_B$, $\hat{\mathbf{R}} = \sigma_o^2 \mathbf{C}_R$, and define $\boldsymbol{\chi} = C_B^{1/2}(\mathbf{x}_0 - \mathbf{x}_0^b)$. For the linear model (11), minimizing the objective function (3) is then equivalent to minimizing the function

$$\tilde{J}(\boldsymbol{\chi}) = \mu^2 \|\boldsymbol{\chi}\|_2^2 + \|\mathbf{C}_R^{-1/2}\hat{\mathbf{d}} - \mathbf{C}_R^{-1/2}\hat{\mathbf{H}}C_B^{1/2}\boldsymbol{\chi}\|_2^2, \tag{19}$$

where $\mu^2 = \sigma_o^2/\sigma_b^2$.

We see that if $\mu^2 = 0$, that is, if there is no background constraint specified, then the problem is ill-posed in practice, since there are fewer observations than degrees of freedom in the initial states. The problem is then an under-determined least squares problem, which does not have a unique solution. To ensure that the problem is well-posed, an additional condition is added in order to *regularize* the problem and to constrain the solution in the null space. Generally this additional constraint requires the least-square (weighted) length of the solution also to be minimized. The complete problem is then written in the form (19).

For noisy data the regularized problem is still likely to be ill-conditioned, and therefore the solution may be very sensitive to the noise. In this case we aim to select a value for μ^2 to maximize the conditioning and minimize the sensitivity. We are thus able to extract the maximum information from the noisy data by making a good choice for the variance ratio. Several techniques for determining μ^2 are given in the literature. A simple (although still expensive) method is the L-curve technique, illustrated in Figure 6. Here the logs of the two separate least-square objective functions in (19) are plotted against each other for various values of μ^2. The point at which the curve has maximum curvature is known to give the optimal choice for μ^2 [9]. For the Eady problem examined here, we see that the optimal value should be in the region $\mu^2 \approx 0.08 - 1.0$, which is the range for the best values found by experiment. For more precise computation, the Generalized Cross Validation (GCV) technique provides an algorithm for determining the point of maximum curvature ([10]). An efficient implementation of the algorithm can be achieved using a Lanczos process ([11]).
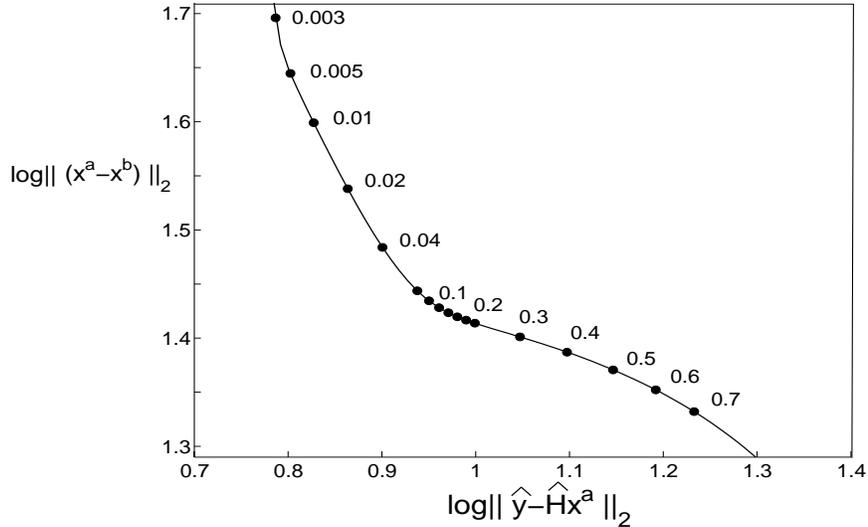
Figure 6: The L-curve: a parametric plot of the values of $\log \|\mathbf{x}^a(\mu) - \mathbf{x}^b\|_2$ and $\log \|\hat{\mathbf{y}} - \hat{\mathbf{H}}\mathbf{x}^a(\mu)\|_2$ for different values of μ^2, which are written beside each point.

.

6 CONCLUSIONS

We describe here the four dimensional variational data assimilation (4DVar) technique for combining numerical model predictions with noisy observations in order to generate the optimal estimate of the expected states of a system. This technique is currently used for numerical weather forecasting in various operational centres. We demonstrate, with a simple model of a baroclinic instability, that the 4DVar method is able to use a time sequence of observations to reconstruct the state of the system in unobserved regions. The reconstructed states are shown to be sensitive to noise, which is filtered by the constraint on the solution imposed by a 'prior', or background, estimate of the initial state. We demonstrate also that for the most accurate reconstruction of the states, the observations should be taken as far apart as possible in time, within the assimilation window.

An analysis of the 4DVar assimilation procedure using a singular value decomposition (SVD) technique is presented, and the results of the experiments are interpreted in terms of the singular values and singular vectors of the 'observability' matrix of the system. We show that the filtering effect of the background is controlled by a 'regularization' parameter, which may be considered as the variance ratio between the observation and background error variances. We also establish that a good choice of the regularization parameter can improve the reconstruction of the states in unobserved regions. Applicable techniques for selecting a good choice of this parameter based on

14

Tikhonov regularization theory are also described and demonstrated. More details of this work can be found in [12, 13].

References

[1] Lorenc AC. Analysis methods for numerical weather prediction. *Q. J. Roy. Meteorol. Soc.* 1986; **112**:1177–1194.

[2] Ghil M. Malanotte-Rizzoli P. In *Data Assimilation in Meteorology and Oceanography, Advances in Geophysics* Dmowska R. Saltzman B. (eds) 1991; **33**:141–266.

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

[4] Thépaut JN, Hoffman RN, Courtier P. Interactions of dynamics and observations in a 4D variational assimilation. *Mon. Weather Rev.* 1993; **121**:3393–3414.

[5] Eady ET. Long waves and cyclone waves. *Tellus* 1949; **1**:33–52.

[6] Barnett S, Cameron, RG. *Introduction to the Mathematical Theory of Control* (2nd edn). Clarendon: Oxford, 1985.

[7] Zou X, Navon IM, Le Dimet FX. Incomplete observations and control of gravity waves in variational data assimilation. *Tellus* 1992; **44A**:273–296.

[8] Hansen PC. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion.* SIAM: Philadelphia, 1998.

[9] Hansen PC, O'Leary DP. The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM J. Sci. Computing* 1993; **14**: 1487–1503.

[10] Golub GH, Heath M, Wahba G. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics* 1979; **21**: 215–223.

[11] Golub GH, von Matt U. *Generalized Cross-validation for Large Scale Problems* (Revised version). Stanford SCCM Report 96-08, Stanford University, 1996.

[12] Johnson C. *Information content of observations in variational data assimilation.* PhD Thesis, The University of Reading, 2003.

[13] Johnson C, Nichols, NK, Hoskins, BJ. *A Singular Vector Perspective of 4D-VAR: Filtering and Interpolating.* Numerical Analysis Rept NA 1/04, Department of Mathematics, The University of Reading, 2004.