4D-Var with the 2D Eady model

Dr Christine Johnson

cjohnson@ucar.edu

August 2005

The 2D Eady model (Eady, 1949) is used within four-dimensional variational data assimilation (4D-Var) identical twin experiments where a horizontal line of the temperature field is observed at two time levels. This program may be used to investigate how 4D-Var uses the model dynamics to interpolate through observations in both space and time, whilst filtering the observational noise. More details on this work can be found in Johnson (2003), Johnson *et al.* (2005a), Johnson *et al.* (2005b) and Johnson *et al.* (2005c).

Contents

1 Description of the code

1.1 Eady model

The Eady model is one of the most simple models of baroclinic instability, which is the dominant mechanism for the growth of storms at mid-latitudes. A schematic diagram of the growth of such storms is shown in Fig. 1. The diagram shows a basic state that consists of a meridional temperature gradient that is associated with a vertical zonal wind shear with height. Wave-like perturbations have been added to the basic state. The upper and lower level waves act like Rossby waves; the upper level wave propagates westwards relative to the flow and the lower level wave propagates eastwards relative to the flow. For a specified basic state shear, it is possible for the upper and lower level waves to become phase-locked. This occurs when the upper level wave is out of phase with the lower level wave so that the temperature field tilts eastwards with height. The circulation associated with the upper level wave extends down to the ground and, due to the tilt of the temperature field, intensifies the lower level wave. Similarly, the circulation associated with the lower level wave extends to the tropopause and intensifies the upper level wave. Thus, the entire wave grows with time. Conversely, a wave in which the temperature field tilts westwards with height decays with time.

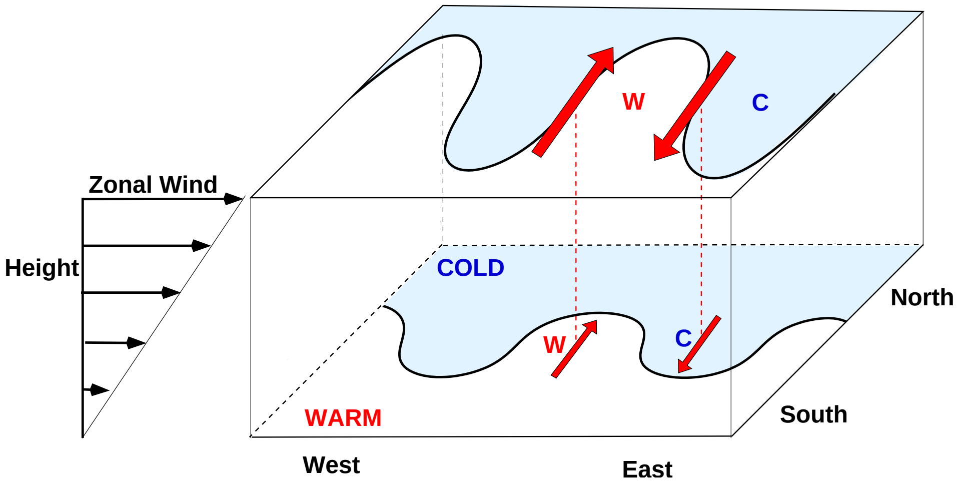

Figure 1: *Schematic diagram showing the vertical coupling between upper and lower level waves in the Eady model. The 4D-Var experiments observe the lower level wave and aim to retrieve the unobserved upper level wave and the basic state vertical shear.*

The non-dimensional equations for the 2D Eady model (Eady, 1949) are now described. The basic state is given by a linear zonal wind shear with height in a domain between two rigid horizontal boundaries. The domain is infinite in the meridional direction and the only dependence on this direction is the uniform meridional temperature gradient which is in thermal wind balance with the zonal wind shear. The density, static stability and Coriolis parameters are all taken to be constants.

The perturbation to the basic state is described by the non-dimensional buoyancy,

b, on the upper and lower boundaries and by the non-dimensional quasi-geostrophic potential vorticity (QGPV), q, in the interior. Equivalently, the perturbation may also be described by the non-dimensional geostrophic streamfunction, ψ, which satisfies:

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial z^2} = q, \qquad \text{in } z \in \left[-\frac{1}{2}, \frac{1}{2}\right], \qquad x \in [0, X], \tag{1a}$$

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

where x is the non-dimensional distance in the zonal direction, and z is the non-dimensional height. The non-dimensional time will be denoted by t. The perturbation to the basic state is advected zonally by the basic shear flow as described by the non-dimensional QG thermodynamic equation and QGPV equation:

$$\left(\frac{\partial}{\partial t} + z\frac{\partial}{\partial x}\right) b = \frac{\partial\psi}{\partial x}, \qquad \text{on } z = \pm\frac{1}{2}, \qquad x \in [0, X], \tag{2a}$$

$$\left(\frac{\partial}{\partial t} + z\frac{\partial}{\partial x}\right) q = 0, \qquad \text{in } z \in \left[-\frac{1}{2}, \frac{1}{2}\right], \qquad x \in [0, X]. \tag{2b}$$

The perturbation is periodic in the horizontal so that the lateral boundary conditions are: $b(0, z, t) = b(X, z, t)$ and $q(0, z, t) = q(X, z, t)$. The model is discretized using 11 vertical levels for QGPV with 20 grid points in one periodic interval in x. The advection equations are discretized using a leap-frog advection scheme.

1.2 4D-Var

In 4D-Var, the analysis at time t_0, $\mathbf{x}^a$, is given by the initial state, $\mathbf{x}_0$, which minimizes the cost function,

$$J(\mathbf{x}_0) = J^o + J^b \tag{3a}$$

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^T\mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}^b) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}_i - \mathbf{H}_i\mathbf{x}_i)^T\mathbf{R}^{-1}(\mathbf{y}_i - \mathbf{H}_i\mathbf{x}_i), \tag{3b}$$

subject to the linear model constraint,

$$\mathbf{x}_{i+1} = \mathbf{M}(t_{i+1}, t_i)\mathbf{x}_i \quad \text{for } i = 0, \ldots, N-1. \tag{3c}$$

Here, $\mathbf{x}_i$ is the state vector at time t_i, $\mathbf{x}^b$ is the background state, $\mathbf{y}_i$ is the vector of observations at time t_i, and $\mathbf{H}_i$ is the (linear) observation operator which converts from state space to observation space, and $\mathbf{M}(t_{i+1}, t_i)$ is the linear model operator from time t_i to t_{i+1}. $\mathbf{B}$ and $\mathbf{R}_i$ are the specified background state and observation error covariance matrices.

To compute the analysis we require the gradient of the unconstrained cost function with respect to the initial state $\mathbf{x}_0$. This is given by:

$$\nabla J^b = \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}^b) \tag{4}$$

$$\nabla J^o = -\left\{\mathbf{H}^T\mathbf{d}_0 + \mathbf{M}^T(\mathbf{H}^T\mathbf{d}_1 + \mathbf{M}^T(\mathbf{H}^T\mathbf{d}_2 + \ldots + \mathbf{M}^T\mathbf{H}^T\mathbf{d}_N)\ldots)\right\} \tag{5}$$

$$\nabla J = \nabla J^b + \nabla J^o \tag{6}$$

where $\mathbf{d}_i = \mathbf{R}^{-1}(\mathbf{y}_i - \mathbf{H}\mathbf{x}_i)$ is the normalized innovation vector at time t_i. $\mathbf{M}^T$ is the adjoint of the forecast model, which is derived by taking the adjoint of the fortran code of the forward model.

The cost function is define as:

$$J(\mathbf{x}_0) = \mathbf{x}_0^T \begin{pmatrix} \sigma_B^{-2}\mathbf{B}_{bb} & 0 \\ 0 & \sigma_q^{-2}\mathbf{B}_{qq} \end{pmatrix} \mathbf{x}_0 + \sigma_o^{-2} \left(\hat{\mathbf{y}} - \hat{\mathbf{H}}\mathbf{x}_0\right)^T \left(\hat{\mathbf{y}} - \hat{\mathbf{H}}\mathbf{x}_0\right) \tag{7}$$

σ_o^2 is the specified observation error variance, σ_B^{-2} and σ_q^{-2} are the specified background error variances for buoyancy and QGPV respectively. $\mathbf{B}_{bb}$ is block diagonal, with two horizontal correlation matrices, $\boldsymbol{\rho}$, on the diagonal, which correspond to the buoyancy on the upper and lower levels. $\mathbf{B}_{qq}$ is also block diagonal, but with 11 horizontal correlation matrices, $\boldsymbol{\rho}$, on the diagonal that correspond to the 11 vertical levels for QGPV.

The matrices $\boldsymbol{\rho}$, which define the horizontal correlation between variables on one vertical level are specified by defining the inverse correlation matrix,

$$\boldsymbol{\rho}^{-1} = \gamma \left(\mathbf{I} + \frac{l^4}{2} (\mathbf{L}_{xx})^2 \right). \tag{8}$$

$\mathbf{L}_{xx}$ is a finite difference second derivative matrix in the x direction which incorporates periodic boundary conditions, l is the horizontal correlation length-scale, and γ is a scalar parameter that is specified so that the diagonal elements of $\boldsymbol{\rho}$ have a value of one.

1.3 Testing the 4D-Var code

Adjoint Test: From the definition of the adjoint model, the value

$$\beta = (\mathbf{M}\mathbf{x})^T(\mathbf{M}\mathbf{x}) - \mathbf{x}^T(\mathbf{M}^T\mathbf{M}\mathbf{x}) \tag{9}$$

should be zero.

Gradient Test: From a Taylor series of the cost function, the value

$$\phi(\alpha) = \frac{J(\mathbf{x} + \alpha\boldsymbol{\delta x}) - J(\mathbf{x})}{\alpha\boldsymbol{\delta x}^T \nabla J(\mathbf{x})} \tag{10}$$

should tend to 1 as α tends to zero, but this does not hold when α is too close to machine zero.

1.4 SVD of the observability matrix

The 4D-Var analysis increments can be written as:

$$\mathbf{x}^a - \mathbf{x}^b = \sum_{j=1}^{r} f_j c_j \mathbf{B}^{1/2}\mathbf{v}_j, \tag{11a}$$

where

$$f_j = \frac{\lambda_j^2}{\mu^2 + \lambda_j^2}, \tag{11b}$$

$$c_j = \frac{\mathbf{u}_j^T \hat{\mathbf{d}}}{\lambda_j}. \tag{11c}$$

$\hat{\mathbf{d}} = \hat{\mathbf{R}}^{-1/2}(\hat{\mathbf{y}} - \hat{\mathbf{H}}\mathbf{x}^b)$ is the generalized normalized innovation vector, and λ_j, $\mathbf{u}_j$, $\mathbf{v}_j$ and r are the singular values, left singular vectors (LSVs), right singular vectors (RSVs) and rank of the observability matrix

$$\widetilde{\mathbf{H}} = \hat{\mathbf{R}}^{-1/2}\hat{\mathbf{H}}\mathbf{B}^{1/2} \tag{12a}$$

where

$$\hat{\mathbf{H}} = \left[\begin{array}{cccc} \mathbf{H}_0^T & (\mathbf{H}_1\mathbf{M}(t_1,t_0))^T & \ldots & (\mathbf{H}_N\mathbf{M}(t_N,t_0))^T \end{array}\right]. \tag{12b}$$

The particular combination of RSVs that are included in the analysis increment is determined by the coefficients, c_j. If the vector $\hat{\mathbf{d}}$ has a large projection onto the LSV $\mathbf{u}_j$, then the corresponding RSV $\mathbf{v}_j$ is given a large weight. Typically, for perfect observations $\mathbf{u}_j^T\hat{\mathbf{d}}$ has a similar rate of decay as λ_j. For noisy observations, $\mathbf{u}_j^T\hat{\mathbf{d}}$ is often extremely large for small singular values. In this case, the corresponding RSVs would dominate the analysis increment. However, they are filtered by the filter factors, f_j, which damp the RSVs with small singular values, $\lambda_j^2 << \mu^2$. Hence the algorithm selectively filters unrealistic structures and the analysis increment is dominated by the RSVs with large singular values.

The RSVs of the observability matrix are independent of the observed values and the background state. They reveal the possible spatial structures that can be analysed by 4D-Var for the specified linear model dynamics, linearization trajectory (in the case of incremental 4D-Var), error covariances and observation locations in both space and time. The corresponding singular values describe our confidence in that structure being part of the analysis increment. For example, a large singular value indicates that it is likely that the corresponding structure has large errors in the background state, the observations are in a good location to observe that structure, and the observations that are used to observe that structure have small errors. For this reason, the RSVs with large singular values are filtered the least by the 4D-Var algorithm.

2 Basic instructions for using the model

1. Download the files **Eady.f90, Eady.m, Eadyq.m, KE.m, cost.m, SVD.m** and **SVDq.m** and make a directory in which the data is to be written. Adjust the directory path name in each .f90 and .m file so that the data will be written to and read from the appropriate directory. The path name is currently called **directory='/home/cjohnson/4DEady/data/'**.

2. Set the model **parameters** in sections MODULE VARIABLES and DEFINE EXPERIMENT DETAIL PARAMETERS. **Compile** the model by typing in at the command line: `f90 -o run Eady.f90` **Run** the model by typing `./run` or `run` (depending on your system).
3. View the data using **matlab**. At the matlab command line, type `Eady` to view the 4D-Var analysis, type `KE` to view the corresponding Kinetic Energy, type cost to view the behaviour of the minimization algorithm and type `SVD` to view the corresponding SVD results. (Note: If interior QGPV is used, Eadyq.m and SVDq.m should be used instead of Eady.m and SVD.m).

2.1 Example 1

This first example illustrates the plots that can be produced by running the code with the current settings. Thus, the user does not need to make any changes to the code except for the directory path name.

For completeness, the current settings are now described. In MODULE VARIABLES, the size of the spatial grid and length of the assimilation window are defined. The number of grid points and size of the domain can also be varied. The current setting is NN =20 and dx=0.0625*pi, which gives a non-dimensional length of 3.9. By multiplying by the Rossby radius of deformation, $L_R = NH/f$, this is equivalent to a dimensional length of about 4000km. The number of vertical levels is specified as JJ =11. The length of the 4D-Var window is set by the number of time levels, given by the parameter TIME, which is defined in MODULE VARIABLES. The current setting is TIME=11 and the time step is DT=4320 sec,so this is equivalent to a 12 hour window. The parameters in DEFINE EXPERIMENT DETAIL PARAMETERS are:

```
length    =5.d0    !background correlation length scale. 5 gives 5*dx=
                   !1000km when dx=0.0312d0*pi*2. 5 gives 5*dx= 500km
                   !when dx=0.0312d0*pi
posn      =1       !The temporal position of the first set of observations
                   !(value between 1 and time - note for posn=time need
                   !to double the weight given to the observations).
height    =0       !The height of the observations (value between 0 and 11)
truth     =0       !True state: 0-Growing 1-Decaying 2-PV
phase     =5       !Size of the phase error
background=0       !Background state: 0-Amplitude, 1-Phase, 3-zero
no_pv     =1       !Is pv a control-variable? 0-yes, 1-no
sigmao2   =1.0d0   !Assumed observation error variance
sigmab2_B =1.0d0   !Assumed background error variance for buoyancy
sigmab2_q =1.0d0   !Assumed background error variance for QGPV
noise     =1       !Noise added to the observations? 0-yes 1-no
sigma_o   =1.0d0   !Standard deviation of the actual observational noise
EPS       =1.2d-10 !Convergence parameter
```

```
mxfun     =500     !Max number of simulations (function evaluations)
itermax   =500     !Max number of iterations  (gradient evaluations)
nmeth     =1       !Type of minimization: 0-Memoryless, 1-BFGS
```

These define the experiment where there is a horizontal line of observations of the buoyancy field at two time levels. One set is at the end of the window, and the other set is at the beginning of the window (defined by posn=1). Both sets are on the lower boundary (defined by height =0). The true state is given by a growing normal mode (defined by truth=0) and the background state has an amplitude error (defined by background =0). Only the upper and lower level buoyancy fields can be adjusted in the minimization (defined by no_pv=1). The assumed observation error variance is 1 and the assumed background error variance is 1, so the parameter $\mu^2 = \sigma_o^2/\sigma_b^2$ is 1. The observations are perfect (defined by noise=1). A quasi-Newton algorithm is used for the minimization (defined by nmeth=1).

The program is split into three sections: testing the code, computing the analysis and performing the SVD computations. The code will ask whether each section is to be computed. An example of the questions and the output from the adjoint and gradient tests is shown below.

```
Do you want to test the linear and adjoint models? 0- Yes, 1-No
Do you wish to perform the minimization? - Yes(0), No(1)
Do you wish to compute the SVD? - yes(0), no (1)

 Do you want to test the linear and adjoint models? 0- Yes, 1-No
0
 * Performing the Adjoint Test - to test the adjoint model code
 The difference in the norms should be zero to machine precision.
 The actual difference in norms is: -2.220446049250313E-16
 * Performing the Gradient Test - to test the cost function code
 The values of phi should be 1, the actual values are:
phi =      0.497176721110   alpha =  0.10E-11
phi =      1.026927089465   alpha =  0.10E-10
phi =      1.005239898009   alpha =  0.10E-09
phi =      1.000696731420   alpha =  0.10E-08
phi =      0.999987825992   alpha =  0.10E-07
phi =      1.000007970221   alpha =  0.10E-06
phi =      1.000027248677   alpha =  0.10E-05
phi =      1.000265093734   alpha =  0.10E-04
phi =      1.002650401575   alpha =  0.10E-03
phi =      1.026503981463   alpha =  0.10E-02
phi =      1.265039812364   alpha =  0.10E-01
phi =      3.650398123572   alpha =  0.10E+00
phi =     27.503981235719   alpha =  0.10E+01
phi =    266.039812357185   alpha =  0.10E+02
```

The output from computing the analysis is interpreted using the matlab programs Eady.m and KE.m. The figures are shown in Fig.2 and 3. For the minimization, the program prints the variables iter, sim and grad J on each iteration. iter is the iteration number (number of gradient evaluations), and sim is the total number of function evalutions. There are more function evaluations because a line search is performed on each iteration. grad J is the size of the gradient of J = $(\nabla J)^T(\nabla J)$. At the end of the minimization, the value of the cost function at the minimum is shown. In examples where the observations are perfect, we can not expect the value to be 1. However, when noise is added to the observations, the error variance parameters can be tuned so that the cost function has a value of 1 at the minimum. The output from the minimization algorithm is viewed using cost.m. The plot is shown in Fig. 4.

For the SVD computations, the code first computes the observability matrix and then computes the SVD. It then prints a list of RSV numbers, the values of $\mathbf{u}^T\mathbf{d}$ and the singular values. The code then asks:

```
Which RSV do you wish to output? - input number
```

You should enter the RSV number which you wish to view in detail using matlab. The values of $\mathbf{u}^T\mathbf{d}$ indicate which RSVs give a large contribution to the analysis increment, so these RSVs are useful to view. The output from computing the SVD is interpreted using the matlab program SVD.m. The output is shown in Figs. 5 to 7.

2.2 Example 2

The second example illustrates the range of experiments that can easily be produced using this code. The following parameters need to be set in the .f90 code. Set TIME=11, NN=80, DX=0.0312d0*pi and the parameters in the DEFINE EXPERIMENT DETAIL PARAMETERS section are:

```
length    =5.d0    !correlation length scale. 5 gives 5*dx= 1000km
                   !when dx=0.0312d0*pi*2. 5 gives 5*dx= 500km
                   !when dx=0.0312d0*pi
posn      =1       !The temporal position of the first set of observations
                   !(value between 1 and time - note for posn=time need to
                   !double the weight given to the observations).
height    =5       !The height of the observations
                   !value between 0 (lower) and 11 (upper boundary)
truth     =2       !True state: 0-Growing 1-Decaying 2-PV
phase     =5       !Size of the phase error
background=3       !Background state: 0-Amplitude, 1-Phase, 3-zero
no_pv     =0       !Is pv a control-variable? 0-yes, 1-no
sigmao2   =1.0d-5  !Assumed observation error variance
sigmab2_B =1.0d-5  !Assumed background error variance for buoyancy
```

```
sigmab2_q =1.0d0   !Assumed background error variance for QGPV
noise     =1       !Noise added to the observations? 0-yes 1-no
sigma_o   =1.0d0   !Standard deviation of the actual observational noise
EPS       =1.2d-10 !Convergence parameter
mxfun     =500     !Max number of simulations (function evaluations)
itermax   =500     !Max number of iterations  (gradient evaluations)
nmeth     =1       !Type of minimization: 0-Memoryless, 1-BFGS
```

The correlation length scale is 500km, there are observations at the beginning and the end of the 6 hour window. The observations are at a height of 4.5km. The true state has a PV dipole, and the background state is zero. PV is a control variable, so both the interior QGPV and the buoyancy on the boundaries can be adjusted in the assimilation. The specified variances are $\sigma_o^2 = 10^{-5}$, $\sigma_b^2 = 10^{-5}$ and $\sigma_q^2 = 1$. Thus the mu parameters are $\mu_q^2 = \sigma_o^2/\sigma_q^2 = 10^{-5}$ and $\mu_b^2 = \sigma_o^2/\sigma_b^2 = 1$. The observations are perfect. For this experiment, it is also necessary to remove the preconditioning. This is achieved by commenting out the line (add a !):

```
objgrd=matmul(P_matrix,objgrd)
```

in subroutine 'calculate_gradJ'. The result from the minimization is viewed using Eadyq.m and is shown in Fig. 8. The result from the SVD is viewed using SVDq.m and is shown in Fig. 10.

3 Simple investigations

Here are some suggested experiments that you might like to try using this model.

- **Adaptive observations.** There have been many field experiments where extra observations have been added to the observing system. For example, by dropping extra dropsondes from aircraft. An important question to address is where to place these extra observations. Use the Eady model to investigate where the optimal location is for the observations in both space and time. This may be achieved by adjusting the parameters 'height' and 'posn'.

- **Noisy observations.** The purpose of data assimilation is to extract useful information from noisy observations. Investigate the effect of noise by adding noise to the observations (parameter 'noise'), and adjusting the standard deviation of the observational noise (parameter 'sigma_o'). Also investigate the effect of the background error correlations by adjusting the length scale parameter 'length'.

- **Covariance specification.** There has been a great deal of research in providing good estimates of the background error covariances. These covariances are often estimated using previous data and analyses and are not flow-dependent. Investigate the effect of misspecifying the specified background error variances (sigmab2_B and sigmab2_q) and the observation error variance (sigmao2).

4 Figures

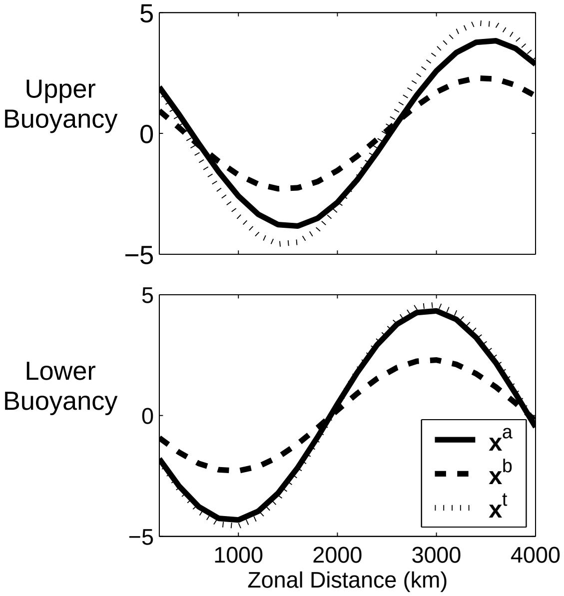

Figure 2: *4D-Var analysis (solid), background state (dashed) and true state (dotted) shown at the middle of the assimilation time window. Produced from Example 1 using Eady.m*

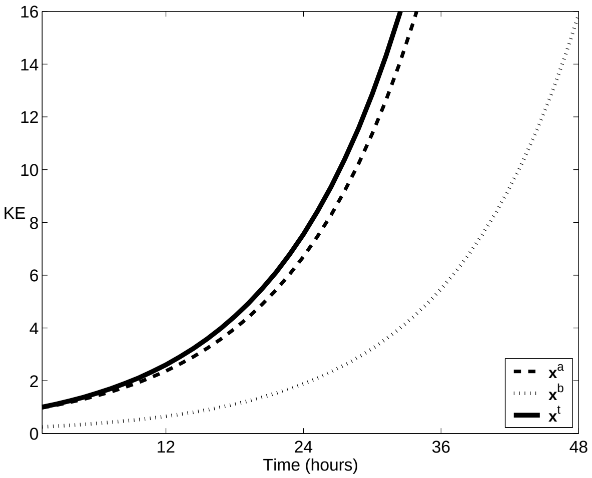

Figure 3: *Evolution of the Kinetic energy during the 12 hour assimilation window and following forecast. Produced from Example 1 using KE.m*

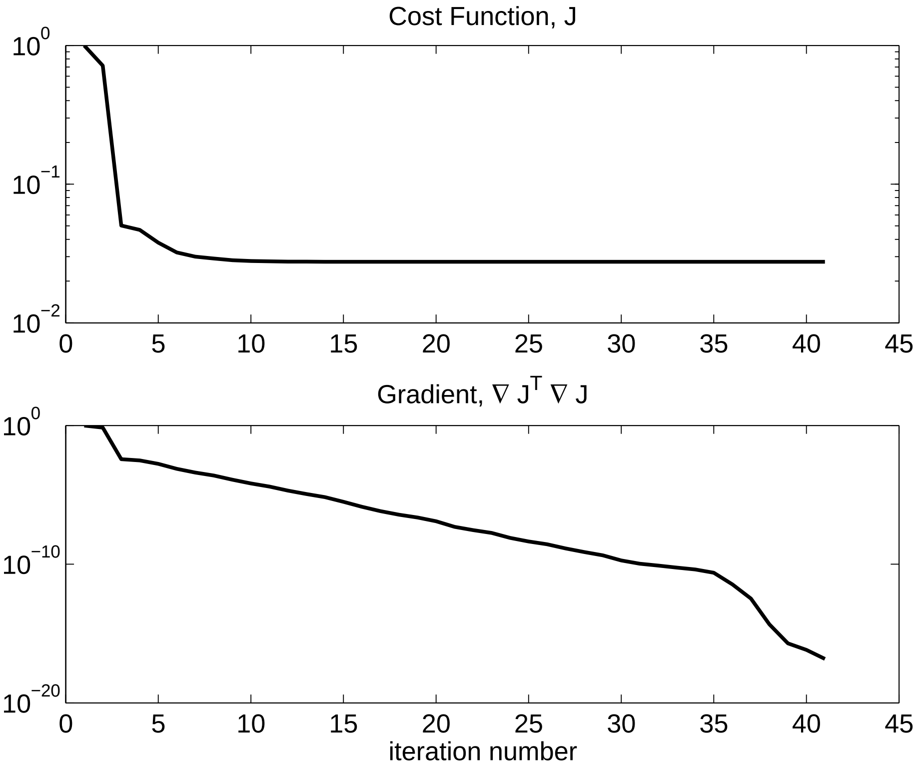

Figure 4: *The behaviour of the minimization algorithm. Produced from Example 1 using cost.m.*

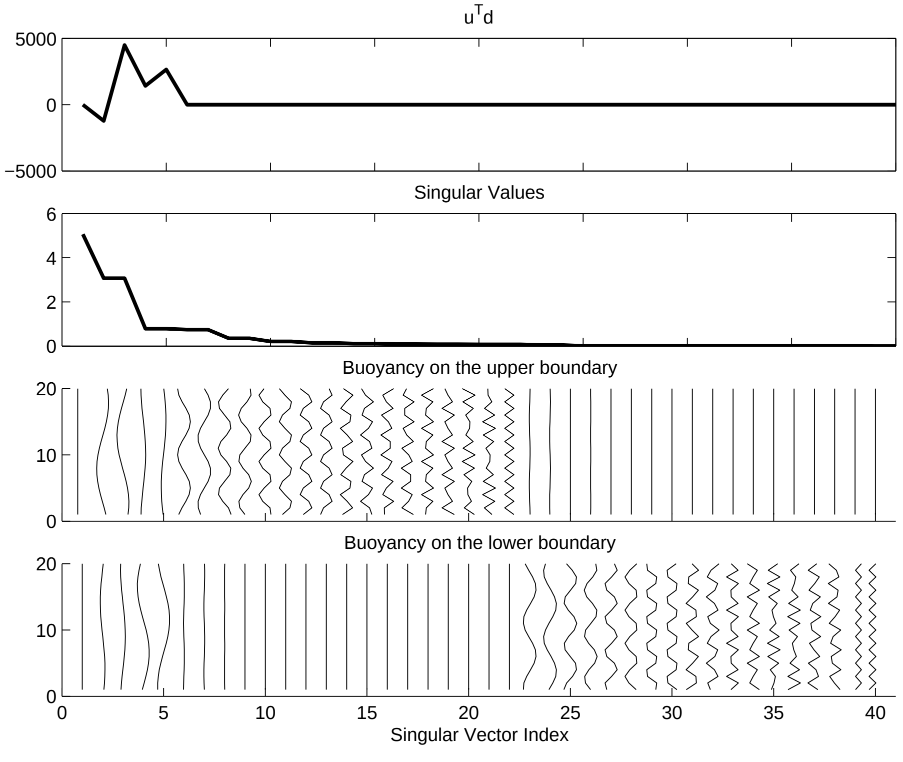

Figure 5: *The SVD result. The values of* $\mathbf{u}^T\mathbf{d}$, λ *and the RSVs. The RSVs are defined in state space (the upper and lower boundaries). Produced from Example 1 using SVD.m*

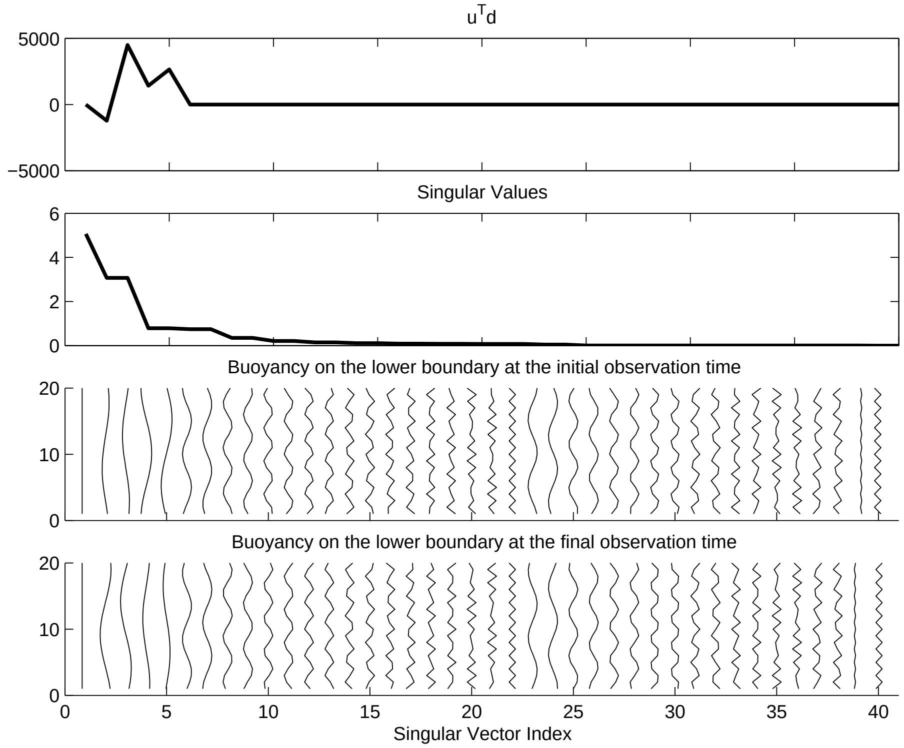

Figure 6: *The SVD result. The values of* $\mathbf{u}^T\mathbf{d}$*,* λ *and the LSVs. The LSVs are defined in observation space. Produced from Example 1 using SVD.m*

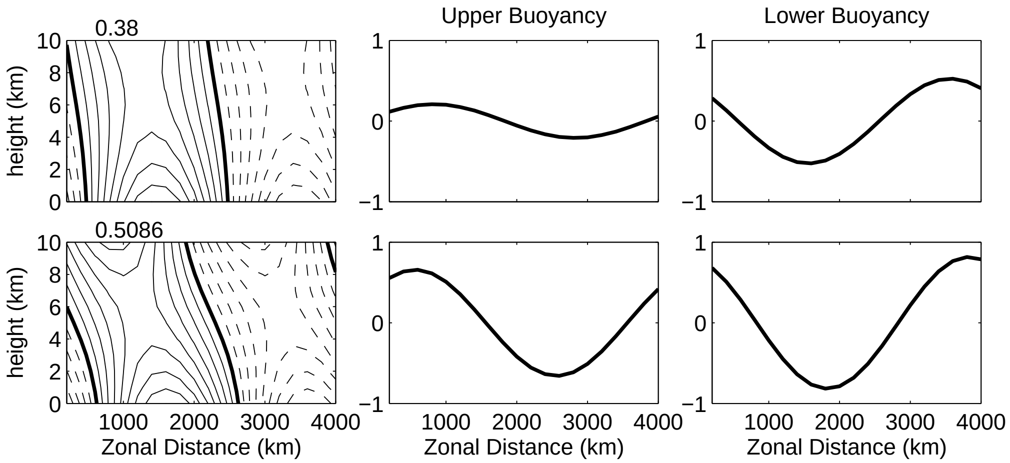

Figure 7: *The second RSV of the observability matrix, premultiplied byt he square root of the background error correlation matrix. The lower panels show the result of integrating the fields by the Eady model over the assimilation window. The panels on the right and the int he center show the lower and upper buoyancy respectively The panels on the left show the corresponding streamfunction field, where the numbers at the top indicate the maximum magntiude of the streamfunction field. Produced using from Example 1 using SVD.m*

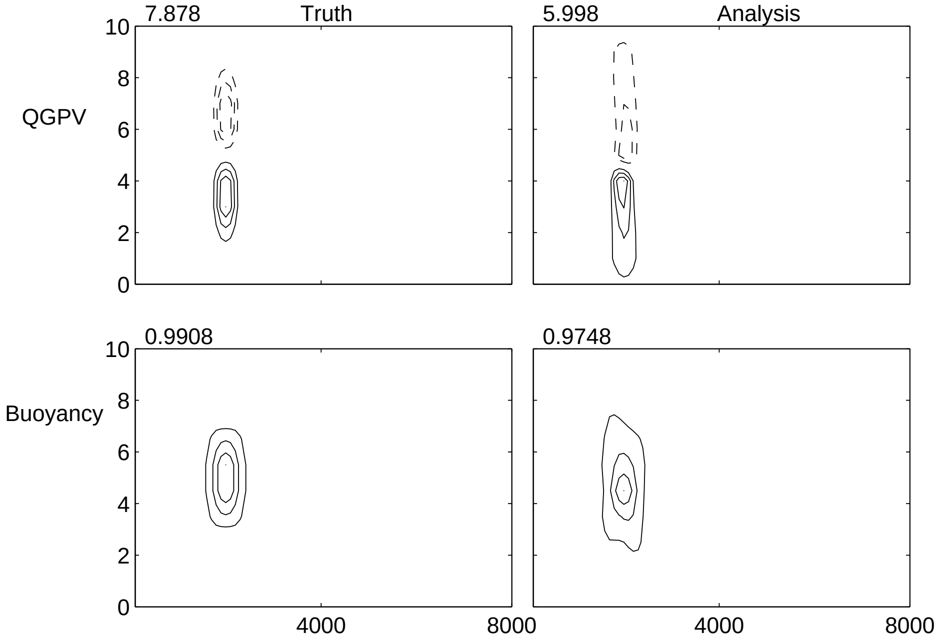

Figure 8: *The true state (a) and the 4D-Var analysis (b). The upper panels show the QGPV and the lower panels show the buoyancy at the beginning of the assimilation window. Produced from Example 2 using Eadyq.m*

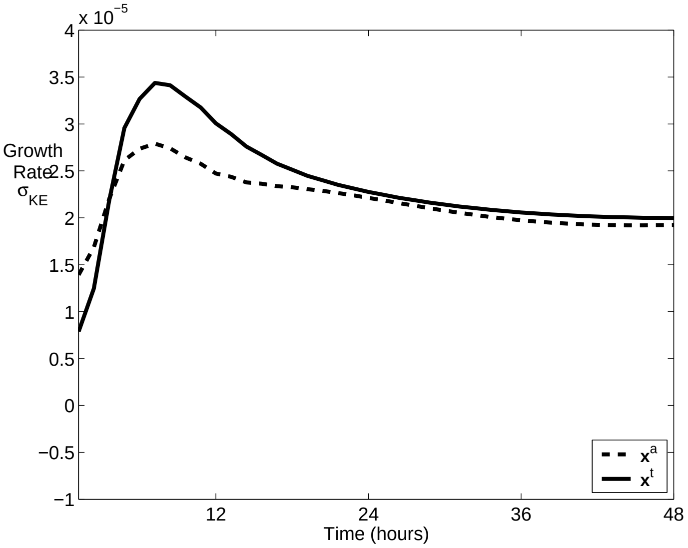

Figure 9: *The evolution of the KE growth rates. The true state is shown by the solid line and the analysis is shown by the dashed line. Produced from Example 2 using Eadyq.m*

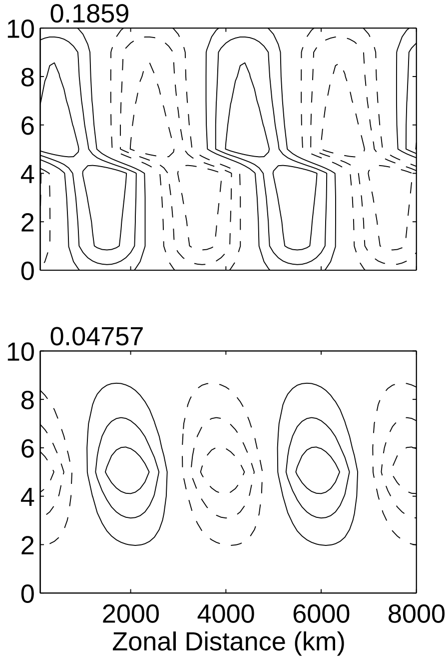

Figure 10: *RSV 3. The upper panel shows the QGPV and the lower panel shows the Buoyancy. Produced from Example 2 using SVDq.m.*

References

Eady, E. T. (1949). Long waves and cyclone waves. *Tellus*, **1**, 33–52.

Johnson, C. (2003). *Information Content of Observations in Variational Data Assimilation*. Ph.D. thesis, University of Reading.

Johnson, C., Hoskins, B. J., and Nichols, N. K. (2005a). A singular vector perspective of 4d-var: Filtering and interpolation. *Q. J. R. Meteorol. Soc.*, **131**, 1–20.

Johnson, C., Hoskins, B. J., and Nichols, N. K. (2005b). A singular vector perspective of 4d-var: The spatial structure and evolution of baroclinic weather systems. Submitted to Mon. Weather Rev.

Johnson, C., Nichols, N. K., and Hoskins, B. J. (2005c). Very large inverse problems in atmosphere and ocean modelling. *Int. J. Numer. Meth. Fluids*, **47**, 759–771.