Abstract

Four Dimensional Variational Data Assimilation (4D Var) is needed to find the present state of the atmosphere, to use as the initial conditions for numerical weather prediction models. In this report, the 4D Var method is applied to a theoretical idealised case of baroclinic instability, using the 2D Eady model. The Eady model describes the vertical coupling between upper and lower boundary waves in the atmosphere.

The case where only part of the flow is observed is considered, to investigate how information is propagated to the unobserved regions by the model dynamics. This is then extended to include an a priori constraint, or background term. The effect of the background term on the growth rate is also examined by assimilating horizontal lines of buoyancy observations.

These simple theoretical case studies allow us to develop a greater understanding of the processes within 4D Var and the limitations of the method when used in the presence of baroclinic instability.

Contents

1 Introduction and Aims **5**

1.1 F

A.3 Centred Time Centred Space (CTCS), Leapfrog 27
A.4 Calculating $\frac{\partial \psi}{\partial x}$ 28
A.5 Solving the Laplace Equation 28

B Derivation of the Eady model equations **28**

B.1 Quasi-Geostrophic Equations 29
B.2 Basic State 30
B.3 Equations on the Boundaries 31
B.4 Equations in the Interior 31
B.5 Non-dimensionalising and Co-ordinate Change 32
B.6 Solutions of the equations 32
B.7 Vertical Structure of the Normal Mode Solutions 34

List of Figures

1 Schematic diagram of the Four Dimensional Variational data assimilation method: minimise the squared distance between the analysis and the background state at the beginning of the assimilation window, and the squared distance between the observations and the forecast state throughout the assimilation window. 5

2 The most unstable growing Eady wave used as the true solution in the identical twin experiments, illustrating the baroclinic instability mechanism 10

3 Behaviour of the cost function using (a) Steepest Descent Method and (b) Conjugate Gradient Method (c) Quasi-Newton Method (d) Limited Memory Method, with increasing 'iterations' or gradient evaluations. The solid line corresponds to the cost function J, and the dotted line corresponds to the squared euclidean norm of the gradient, $\|\nabla \mathbf{J}\|^2$ with the magnitude on the left hand axes. The circles show the number of 'simulations' or function evaluations to calculate the next step, with the magnitude on the right hand axes This minimisation is for the case with no observations on the top boundary. 14

4 The behaviour of the cost function with increasing iterations (a) cost function (solid line), squared 2-norm of the gradient vector (dashed line), simulations or function evalutions (circles) (b) The set of termination criteria (equations 3 , 40, 41) (c)Comparison of Maximum Norms of ∇J (equation 35) (d)Comparison of Euclidean Norms of ∇J (equation 34). 15

5 Streamfunction Analysis from (a) 3D Var and (b) 4D Var where there are no observations given on the upper boundary. 4D Var is able to reconstruct the upper wave, through the use of the model dynamics. 18

6 The behaviour of the solution with increasing iterations for the Quasi-Newton Method. The columns on the left show $-\nabla_{\mathbf{b}_0} J$,$-\nabla_{\mathbf{B}_1} J$ and the columns on the right show b_0 and b_1. 18

7 Analyses from a two timelevel assimilation with no observations on the top level and a zero background state, with (a) Effectively equal weights, (b) Weight given to the observations is 10 times greater than the background state weight, (c) Weight given to the observations is 1000 times greater than the background state weight. Streamfunction and QGPV fields: Solid - positive contours, dashed - negative contours. Buoyancy on upper and lower boundaries: Green Dotted -True Solution (Observed), Blue Dashed - Background State, Red Solid - Analysis 21

8 Growing Eady Wave, Analysis from 4D Var. A horizontal line of the buoyancy field is observed at the end of the assimilation window, with a zero background state, and identity covariance matrices. 22

9 Decaying Eady Wave, Analysis from 4D Var. A horizontal line of the buoyancy field is observed at every time level, with a zero background state, and identity covariance matrices. 23

10 Growth rates of 4D Var analyses for the decaying Eady wave, using different weightings for the J_b and J_o terms of equation 51. Solid Red - Truth (observed), Blue dashed - analysed. 24

11 The structure of the linear models. In the forward model, the initial conditions are given by potential vorticity,q and buoyancy, b at time t=0. b(t=0) and q(t=0) are then

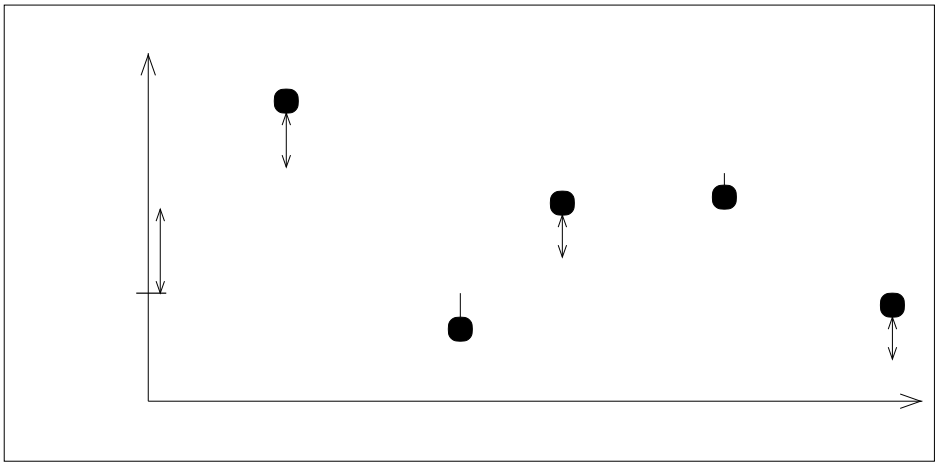
1 Introduction and Aims

The work presented in this report aims to isolate the baroclinic instability mechanism within a 4D Var framework. By performing identical twin experiments, we examine whether a 4D Var system is capable of reconstructing the correct vertical structures of the atmosphere necessary for the growth of mid-latitude cyclones.

1.1 Four Dimensional Variational Data Assimilation (4D Var)

Data assimilation is needed to find the best estimate of the present state of the atmosphere by using observations, a forecast model and climatology. Recent data assimilation methods assume that the analysis is given by an optimal blend of the observed values and a background state (first guess given by a previous forecast). The weights that prescribe the blend are determined by error covariance matrices (Lorenc, 1986). Variational assimilation was introduced to dynamic meteorology by Sasaki in 1958. Variational assimilation determines the analysis by minimising a cost function. Three dimensional variational assimilation (3D Var) includes observations at only one time level, however, four dimensional variational assimilation (4D Var) extends this to include observations that are distributed in time (Sasaki, 1970).

Le Dimet and Talagrand (1986) applied optimal control theory to 'reduce the control variable'. This means that instead of updating the state variables at every time level, only the initial conditions are used as the control variables. This is achieved by constraining the state variables to fit the model equations, through the use of an adjoint model (Errico, 1997). Note that there are two uses of the word control variable. From control theory, the control variable refers to the variable that is updated and it is this definition that is used in this report. However, in other literature, control variables can also refer to the transformed variables that are used when defining the background error covariance matrix. The 4D Var method can be used operationally by linearising the dynamical model to give an 'incremental' formulation (Courtier *et al.*, 1994). Despite these simplifications, 4D Var is still much more computationally expensive than 3D Var.



These equations can also be derived by using optimal control techniques to transform the constrained optimization problem to an unconstrained problem (Griffith (1997), Le Dimet and Talagrand (1986)). This technique is illustrated below by considering the continous scalar case.

Minimize the functional $\int_0^{'1'} F(x,t)dt$, defined over an assimilation window $[0,T]$ $T>0$, subject to the (strong) model constraint $\frac{\partial x}{\partial t} = f(x,t)$, where F and f are scalar real valued functions that are continuous with respect to x and t, and continuously differentiable with respect to x, $x(t) \in \mathbb{R}$ is the state vector and $t \in [0,T]$ is the time,

The lagrangian functional $\mathcal{L}$ is constructed by using the method of Lagrange to give

$$\mathcal{L} = \int_0^{'1'} (F(x,t) + \lambda(t)[\dot{x} - f(x,t)])dt \tag{12}$$

where $\lambda(t) \in C$

Then the gradient of J_o at the initial time is given by

$$\nabla_{\mathbf{x}_0} = -\boldsymbol{\lambda}_0 \tag{21}$$

If $\mathbf{M}$ is the forward linear model, then $\mathbf{M}^T$ is the *adjoint model*, $\boldsymbol{\lambda}$ is the vector of *adjoint variable* and $\nabla_{\mathbf{x}_i} J_o = -\mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y}_i - \mathbf{H}_i \mathbf{x}_i)$ is known as the *adjoint forcing*. Hence the adjoint model is integrated backwards in time, using zero as the final state, and adding an adjoint forcing at each step.

The adjoint equations find an equation for ∇J_0 which can then be used by a descent algorithm to update the control vector $\mathbf{x}_{t=0}$.

2.2 The Eady Model

The 2D Eady model (Eady, 1949) is one of the most simple theoretical model used to study baroclinic instability in the x-z plane. It is a simple linear Quasi-Geostrophic (QG) model which can be used to describe the vertical coupling between waves at the tropopause and the ground. The Eady model equations support two types of normal mode solutions; neutral modes correponding to boundary waves with short wavelengths and unstable long waves, that grow (or decay) exponentially (Farrell, 1982).

The Eady model basic state is given by a zonal wind with a linear shear with height. Perturbations to this are described by the following non-dimensional equations, that are derived in appendix B.

The quasi-geostrophic thermodynamic equation on the top and bottom boundaries

$$\left(\frac{\partial}{\partial t} + \text{▪}\frac{\partial}{\partial x}\right)\frac{\partial \psi}{\partial \text{▪}} = \frac{\partial \psi}{\partial x} \text{ on z } = \pm\frac{1}{2} \tag{22}$$

and the quasi-geostrophic potential vorticity (QGPV) equation in the interior,

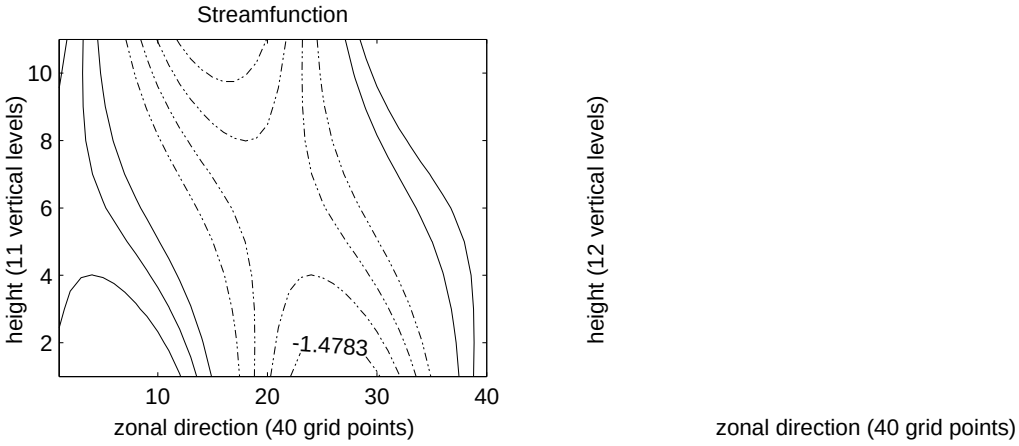
$$q = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial \text{▪}^2} = 0 \text{ in } -\frac{1}{2} < \text{▪} < \frac{1}{2} \tag{23}$$

where $\text{▪} \in [-\frac{1}{2}, \frac{1}{2}]$ is the height, $x \in [0, N]$ is the horizontal distance in the

fields are defined on staggered grids in the vertical, giving 11 vertical levels for streamfunction and 12 vertical levels for buoyancy.

The 4D Var method is applied, using the Eady model as the strong constraint. The control variable, that is updated, is given by the buoyancy and the quasi-geostrophic potential vorticity.

The initial conditions are given by the most unstable Eady Wave (normal mode for the non-dimensional wave number k=1.6, equation (104)) , as shown in Figure 2.



2.4 Minimisation

The data assimilation optimization problem requires a numerical minimisation algorithm. This must be suitable for large-scale unconstrained problems. By making the forward model linear, then the cost function becomes quadratic. This ensures that there exists a unique minimum. Therfore, we need only ensure that the algorithm is capable of finding a local minimum rather than the global minimum of a non-quadratic function.

The two most common methods for such a problem are conjugate gradient methods and quasi-Newton methods. They both approximate Newton's Method, which is first examined.

2.4.1 Newton's Method

Consider the minimisation of an objective function $F(\mathbf{x})$ with respect to a vector $\mathbf{x}$. It is possible to derive an iterative algorithm that can be used to find the value of $\mathbf{x}$ at the minimum. By considering a second order Taylor series expansion of $F(\mathbf{x})$, and using the fact that the gradient of F is zero at a minimum, then the Newton's method algorithm is obtained (Beale, 1988):

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \mathbf{H}^{-1} \mathbf{g}_k \tag{25}$$

where k is the iteration, $\mathbf{H} = \nabla\nabla F$ is the hessian matrix, and $\mathbf{g} = \nabla F$ is the gradient vector. The parameter α_k has been inserted into the equation so that a line minimisation of the function F is performed in the direction $\mathbf{H}^{-1}\mathbf{g}$.

In a large scale problem such as data assimilation the hessian matrix $\mathbf{H}$ is too expensive to calculate and store explicitly, and so approximations must be made. Four such suitable approximations have been tested within the 4D Var framework applied to the Eady model. These are the methods of steepest descent, conjugate gradient, quasi-Newton and limited-memory; which are now described.

2.4.2 Steepest Descent Method

The steepest descent method approximates the hessian matrix by the identity matrix. The algorithm therefore becomes:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \mathbf{g}_k \tag{26}$$

The search direction is given by the 'downhill' direction. However, this method can 'zig zag' into the minimum, givi argl.

using a set of linearly independent vectors $\mathbf{v}_k$. The formula for constructing such directions is given by:

$$\mathbf{d}_k = \mathbf{v}_k + \sum_{j=0}^{k-1} \beta_{kj} \mathbf{d}_j \tag{27}$$

where the β_{kj} are chosen so that $\mathbf{d}_k^T \mathbf{H} \mathbf{d}_j = 0$ for $k \neq j$. This gives:

$$\beta_{kj} = -\frac{\mathbf{v}_k^T \mathbf{H} \mathbf{d}_j}{\mathbf{d}_j^T \mathbf{H} \mathbf{d}_j}. \tag{28}$$

The problem is that $\mathbf{H}$ is unknown. However, it is still possible to construct conjugate directions by using the gradient vectors g_k at each point x_k (Beale, 1972, 1988). That is, we choose $v_k = -g_{k-1} = \nabla F_{k-1}$, to obtain

$$\mathbf{d}_k = -\mathbf{g}_{k-1} + \beta_{k,k} \mathbf{d}_{k-1} \tag{29}$$

$$\beta_{k,j} = \frac{\mathbf{g}_{k-1}^T (\mathbf{g}_j - \mathbf{g}_{j-1})}{\mathbf{d}_j^T (\mathbf{g}_j - \mathbf{g}_{j-1})} \tag{30}$$

The final conjugate direction is conjugate to all previous directions, however, through the way they have been constructed, only the previous search direction needs to be stored.

The conjugate gradient method used in this work is known as A22GCM (Nash, 1990) and is a conjugate gradient method that uses a linear search to bracket a minimum.

The conjugate gradient method uses the hessian matrix information implicitly. However, the quasi-Newton method approximates the hessian explicitly.

2.4.4 Quasi-Newton Method

The quasi-newton method (variable metric method) approximates the hessian matrix using first derivatives. This information is built up during successive iterations.

The approximation to $\mathbf{H}$ should satisfy the quasi-Newton condition (Press *et al.*, 1992):

$$\mathbf{H}(\mathbf{x_{k+1}} - \mathbf{x_k}) = \mathbf{g_{k+1}} - \mathbf{g_k} \tag{32}$$

which is derived by performing a taylor expansion of the gradient of F.

Suppose the hessian is approximated by $\mathbf{B}_k$. Then $\mathbf{B}_k$ can be updated on each iteration using rank 1 matrices. For example, Broydon's update is given by:

$$\mathbf{B}_k = \mathbf{B}_{k-1} + c\mathbf{z}\mathbf{z}^T \tag{33}$$

By substituting this into the quasi-Newton condition (equation (32)), then formulae for $\mathbf{z}$ and c can be found.

The quasi-Newton method allows the hessian to be approximated. However, it is not always possible to store the approximation as it is too large (it is not possible for operational data assimilation), and hence limited memory methods are required.

[1] Two vectors $\mathbf{u}$ and $\mathbf{v}$ are conjugate (or A-orthogonal) with respect to the matrix $\mathbf{A}$ if

$$\mathbf{u^T A v} = 0 \quad \mathbf{u} \neq \mathbf{v} \tag{31}$$

2.4.5 Limited Memory Methods

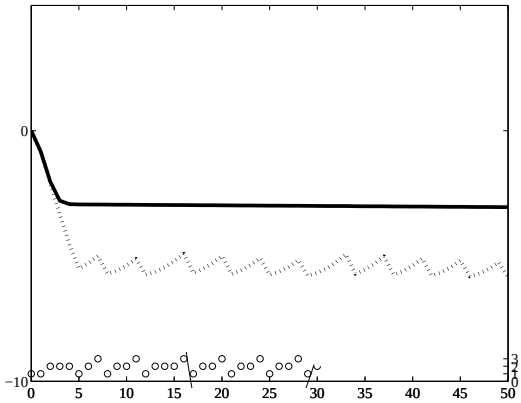
The properties of the quasi-newton method are combined with conjugate gradient method to give limited memory methods. Instead of $\mathbf{v_k} = -\mathbf{g_k}$ in the conjugate gradient method, choose $\mathbf{v_k} = -\mathbf{H^{-1}g_k}$ to encorporate the curvature information aswell. The inverse hessian does not need to be computed explicitly, so only the vector updates to the approximate $\mathbf{H}^{-1}$ are stored, (Navon and Legler, 1987).

The quasi-Newton and limited memory quasi-Newton methods that are used in this work, are known as CONMIN or algorithm 500 from TOMS. The quasi-Newton method is based on a (Broyden-Fletcher-Goldfarb-Shanno update) quasi-Newton method suggested by Shanno and Phua (Shanno and Phua, 1976), and the limited memory method is a memoryless quasi-Newton conjugate gradient algorithm with Beale restarts (Beale,1972,1988). The algorithms have previously been used by Chao and Chang (1992).

2.4.6 Testing Minimisation Methods

An experiment using no background state and observations of buoyancy on the lower boundary, and zero QGPV in the interior over 2 time levels is used to compare minimisation methods.

A comparison of the methods is given in Figure 3. The steepest descent algorithm takes a long time to converge - the algorithm was terminated at 50 iterations, before the algorithm converged. There is a sharp decrease initially, where the buoyancy values on the observed lower boundary are obtained. The rest of the iterations are needed to obtain the information


0
−10
0
5
10
15
20
25
30
35
40
45
50

1. Gradient of J

$$\frac{\|\nabla J\|_\infty}{1+J} \leq \tau_1 \tag{39}$$

2. Change in $\mathbf{x}$

$$\frac{\|\mathbf{x}_k - \mathbf{x}_{k-1}\|_\infty}{1+\|\mathbf{x}_k\|_\infty} \leq \tau_2 \tag{40}$$

3. Change in J

$$\frac{\sqrt{J_k} - \sqrt{J_{k-1}}}{1+\sqrt{J_k}} \leq \tau_3 \tag{41}$$

where τ_i is a specified tolerance.

The behaviour of these criteria are shown in Figure 4 (b), and from this graph, the tolerances are specified. If any of these stopping criteria are satisfied, then the minimisation algorithm is stopped.

3 No Background Term: Observability

In the data assimilation problem the null space[1] needs to be eliminated by adding constraints to the cost function. In 4D Var, we add the model equations as a strong constraint. That is, the solution must also satisfy the model equations. The null space of the 4D Var problem with no background term is now investigated.

For ease of notation, define $\mathbf{v}^2 = \mathbf{v}^T\mathbf{v}$. Consider an assimilation window length of 1 timestep (2 time levels), and identity error covariance matrices. The cost function is

$$J = \frac{1}{2}\left[(\mathbf{y_0} - \mathbf{H x_0})^2 + (\mathbf{y_1} - \mathbf{H x_1})^2\right] \tag{42}$$

where $\mathbf{H}$ is the observation operator. This can be written in the form

$$J(\mathbf{x}_0) = \frac{1}{2}\left[(\mathbf{y}_0 - \mathbf{H x}_0)^2 + (\mathbf{y}_1 - \mathbf{HM x}_0)^2\right] = \frac{1}{2}\left(\begin{bmatrix}\mathbf{y}_0\\ \mathbf{y}_1\end{bmatrix} - \begin{bmatrix}\mathbf{H}\\ \mathbf{HM}\end{bmatrix}\mathbf{x}_0\right)^2. \tag{43}$$

where $\mathbf{M}$ is the linear model. Note that from this, we can write the gradient of J as:

$$\nabla J(\mathbf{x}_0) = -\mathbf{H}^T(\mathbf{y}_0 - \mathbf{H x}_0) - \mathbf{M}^T\mathbf{H}^T(\mathbf{y}_1 - \mathbf{HM x}_0). \tag{44}$$

If $\mathbf{x}_0$ is of length n, then for a unique minimum, we req

We therefore know that with no background term, we need at least as many observations as unknowns (the observations may be distributed in time). Further, the rank of the observation operator $\hat{\mathbf{H}}$ must be equal to the number of unknowns equal to the size of the control vector, n.

3.1 Reconstructing the buoyancy wave on the top boundary

Consider the case where the lower boundary buoyancy and interior quasi-geostrophic potential vorticity are observed over two time levels, but the upper boundary is not observed.

Let the state vector be

$$\mathbf{x} = \begin{bmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \\ \mathbf{q} \end{bmatrix} \tag{46}$$

where the buoyancy on the lower level $\mathbf{b}_0$ and the buoyancy on the upper level $\mathbf{b}_1$ are vectors of length n and the interior QGPV $\mathbf{q}$ is a vector of length m. Suppose that only $\mathbf{b}_0$ and $\mathbf{q}$ are observed, so that the observation operator is

$$\mathbf{H} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}. \tag{47}$$

The potential vorticity is advected with no forcing, the buoyancy is also advected and also forced by the streamfunction, which is determined from the buoyancy on both boundaries and the interior potential vorticity. So, the model can be approximated through scaling and data-flow arguments by the linear operator:

$$\mathbf{M} = \begin{bmatrix} \mathbf{I}_n & a\mathbf{I}_n & b\mathbf{I}_m \\ a\mathbf{I}_n & \mathbf{I}_n & b\mathbf{I}_m \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix} \tag{48}$$

where a, b are scalar constants such that $|a| < 1$ and $|b| < 1$. Then, we can write

$$\mathbf{HM} = \begin{bmatrix} \mathbf{I}_n & a\mathbf{I}_n & b\mathbf{I}_m \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix} \tag{49}$$

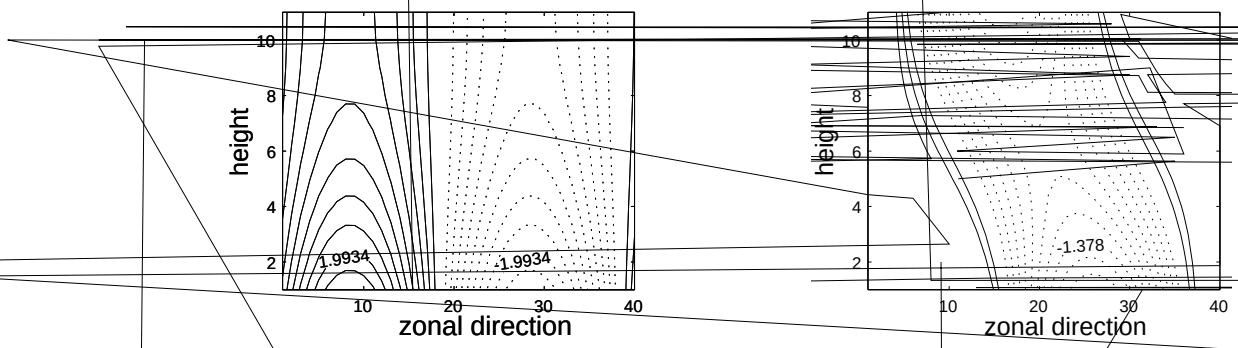
and hence

$$\hat{\mathbf{H}} = \begin{bmatrix} \mathbf{I}_n & a\mathbf{I}_n & b\mathbf{I}_m \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \\ \mathbf{I}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix} \tag{50}$$

In this case $\hat{\mathbf{H}}$ has rank $m + 2n$ and so the problem is well-determined.

4 Results

Results from the idealized twin experi



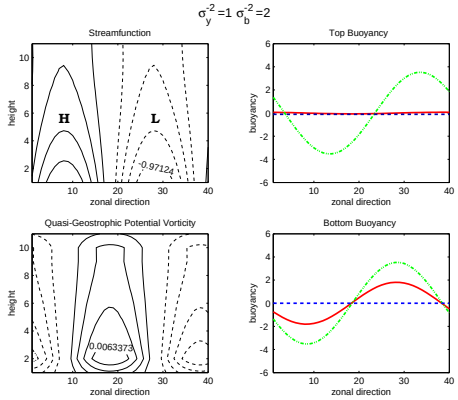
(a)

The gradient fields (produced on output from the adjoint model) and actual fields (produced on output from the minimisation) with increasing iterations are shown in Figure 6. The adjoint model produces gradient fields with buoyancy waves on both boundaries on the first iteration. The wave on the lower boundary (observed) has the correct phase, but the wave on the upper boundary (unobserved) has the wrong phase and is of a smaller magnitude. The amplitudes of both waves increases on the second iteration. To be able to compensate for the fact that the observed wave is growi

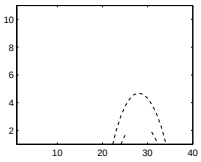
window.

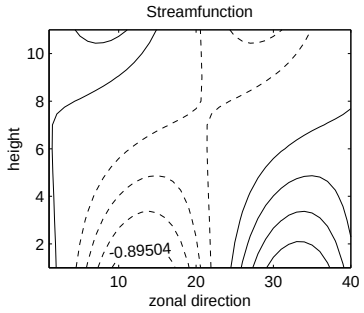
4.3.1 Growing Eady Wave

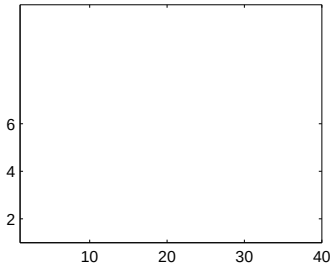
The true solution is given by the most unstable Eady wave. A horizontal line of buoyancy is observed at the end of the assimilation window. The analysis, shown in Figure 8, is close to zero at the beginning of the assimilation window, but has a wave structure at the end. This


σ-2y =1 σ-2b =2
Streamfunction
Top Buoyancy
Quasi-Geostrophic Potential Vorticity
Bottom Buoyancy
H
L
0.971241
0.0063373
height
buoyancy
zonal direction

(a)




Streamfunction
height
zonal direction
-0.89504
10
8
6
4
2
10
20
30
40


6
4
2
10
20
30
40

5 Conclusions and Future Work

This report has examined the 4D Var method by applying the method to a theoretical case of baroclinic instability using the Eady model. This work has shown that 4D Var (with no J_b term) is able to propagate information from observations to the unobserved regions using the model dynamics, giving clear benefits over 3D Var.

It is necessary to add a background constraint to the cost function so that the problem is well determined. However, this work has shown that this can be detrimental to the analysis in the case of a poor background state and poor error statistics. In particular, the J_b Term has a large impact on the regions that are unobserved but obtain information through the model dynamics. The benefits of 4D Var over 3D Var are now lost by adding the background constraint.

With observations at the end of the assimilation window and a zero background state, the analysis increments are projected onto a fast growing solution. If the amplitude of the background state is in conflict with the observations, then the growth rate of the system may be incorrect.

We conclude that the 4D Var method is able to extract much more information from the observations than 3D Var. However, the background constraint must be applied with care.

Future work will continue experiments with a background state using theoretical case studies such as singular vector type structures. In particular, we aim to identify which parts of the atmosphere should be observed to correct a phase error, and to give the correct vertical structure. This work will be able to compare the effect of different observing systems by simulating satellite data, radiosondes and aeroplane data within the Eady model.

The specification of the background error covariance matrix is extremely important as it determines where they information from the observations should be spread. It is known that in a 4D Var system, the background error covariance matrix evolves implicitly through the assimilation window, giving some flow dependency and better vertical structures than 3D Var analyses (Thepaut *et al.*, 1996). This property of 4D Var will be investigated within the simple Eady model context.

A The Adjoint Model

The general structure of the forward Eady model and the corresponding Adjoint model are shown in Figure 11. The equations for the models are given below.

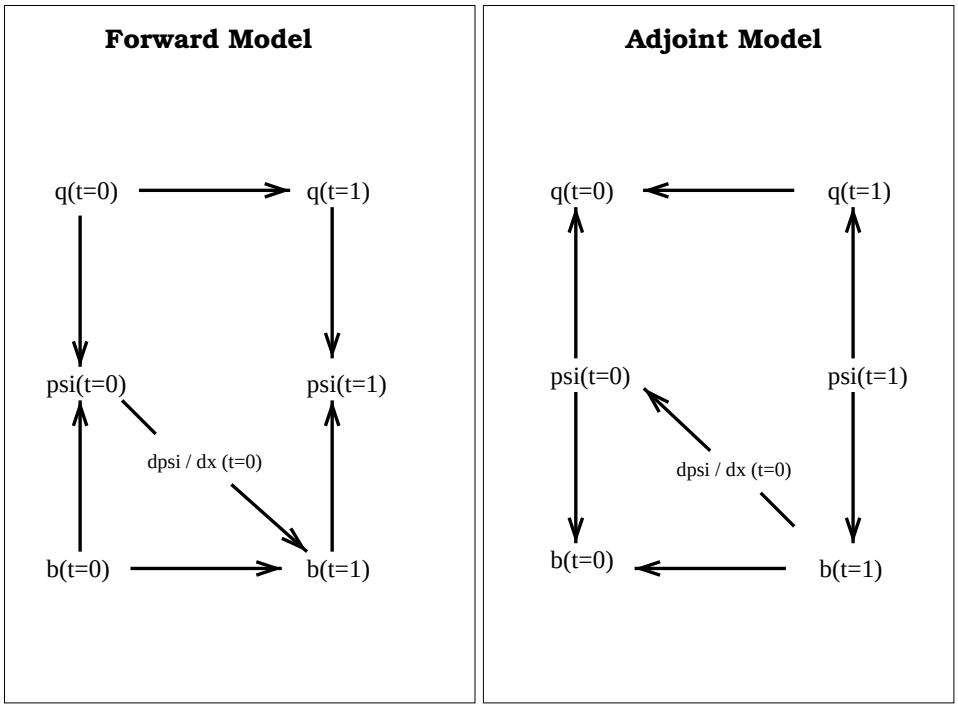


Figure 11:

where the hat denotes adjoint variable. For the minimisation problem, the adjoint variable for A, for example, is $\hat{A} = \frac{\partial J}{\partial A}$. And therefore the discrete adjoint model corresponding to equation 52 is:

$$\begin{aligned} \hat{B} &= \hat{B} + \hat{A} \\ \hat{C} &= \hat{C} + \alpha\hat{A} \\ \hat{D} &= \hat{D} - \alpha\hat{A} \\ \hat{E} &= \hat{E} + \beta\hat{A} \\ \hat{A} &= 0 \end{aligned} \tag{55}$$

The adjoint equations for the numerical schemes used in the Eady model are now given.

A.2 Forward Time Centred Space (FTCS)

For the equation

$$u_j^{n+1} = u_j^n + R(u_{j+1}^n - u_{j-1}^n) + SF_j^n \tag{56}$$

the corresponding discrete adjoint model is:

$$\begin{aligned} \hat{u}_j^n &= \hat{u}_j^n + \hat{u}_j^{n+1} \\ \hat{u}_{j+1}^n &= \hat{u}_{j+1}^n + R\hat{u}_j^{n+1} \\ \hat{u}_{j-1}^n &= \hat{u}_{j-1}^n - R\hat{u}_j^{n+1} \\ \hat{F}_j^n &= \hat{F}_j^n + S\hat{u}_j^{n+1} \\ \hat{u}_j^{n+1} &= 0 \end{aligned} \tag{57}$$

Note that this is equivalent to the adjoint equation:

$$\hat{u}_j^n = \hat{u}_j^{n+1} - R(\hat{u}_{j+1}^{n+1} - \hat{u}_{j-1}^{n+1}) \tag{58}$$

A.3 Centred Time Centred Space (CTCS), Leapfrog

For the equation

$$u_j^{n+1} = u_j^{n-1} + R(u_{j+1}^n - u_{j-1}^n) + SF_j^n \tag{59}$$

the corresponding discrete adjoint model is:

$$\begin{aligned} \hat{u}_j^{n-1} &= \hat{u}_j^{n-1} + \hat{u}_j^{n+1} \\ \hat{u}_{j+1}^n &= \hat{u}_{j+1}^n + R\hat{u}_j^{n+1} \\ \hat{u}_{j-1}^n &= \hat{u}_{j-1}^n - R\hat{u}_j^{n+1} \\ \hat{F}_j^n &= \hat{F}_j^n + S\hat{u}_j^{n+1} \\ \hat{u}_j^{n+1} &= 0 \end{aligned} \tag{60}$$

Note that this is equivalent to the adjoint equation:

$$\hat{u}_j^n = \hat{u}_j^{n+2} - R(\hat{u}_{j+1}^{n+1} - \hat{u}_{j-1}^{n+1}) \tag{61}$$

A.4 Calculating $\frac{\partial \psi}{\partial x}$

The forward equation for calculating $\frac{\partial \psi}{\partial x}$ is:

$$F_j^n = \frac{\psi_{j+1}^n - \psi_{j-1}^n}{2\Delta x} \tag{62}$$

In matrix form this can be written as:

$$\begin{pmatrix} \psi_{j+1} \\ \psi_{j-1} \\ F_j \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2\Delta x} & -\frac{1}{2\Delta x} & 0 \end{pmatrix} \begin{pmatrix} \psi_{j+1} \\ \psi_{j-1} \\ F_j \end{pmatrix} \tag{63}$$

The adjoint of this is:

$$\begin{pmatrix} \hat{\psi}_{j+1} \\ \hat{\psi}_{j-1} \\ \hat{F}_j \end{pmatrix} = \begin{pmatrix} 1 & 0 & \frac{1}{2\Delta x} \\ 0 & 1 & -\frac{1}{2\Delta x} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{\psi}_{j+1} \\ \hat{\psi}_{j-1} \\ \hat{F}_j \end{pmatrix} \tag{64}$$

and therefore the discrete adjoint model is:

$$\hat{\psi}$$

B.1 Quasi-Geostrophic Equations

The Eady model consists of a basic state and a small amplitude perturbation which represents the baroclinic wave or cyclone. The Eady Model equations are derived from the Quasi-

The quasi-geostrophic potential vorticity q combines both dynamical and thermodynamical information and is conserved following adiabatic motion.

These equations can now be used to derive the Eady model equations, by linearizing about a basic state.

B.2 Basic State

Assume that the basic state has a linear vertical wind shear, $\bar{u} = A$■. The vertical wind shear is associated with a meridional temperature gradient, as given by the thermal wind equation

$$f_0 \frac{\partial \bar{u}}{\partial \blacksquare} = -\frac{\partial \bar{b}}{\partial y}. \tag{80}$$

The thermal wind relation is shown in Figure 12. Although the (2D) Eady model equations only describe the flow in the z

B.3 Equations on the Boundaries

Assume that the vertical velocity w is zero on both the upper and lower boundaries. Then, the QG thermodynamic equation becomes:

$$D_g b' = 0 \Rightarrow \left(\frac{\partial}{\partial t} + (\bar{u} + u')\frac{\partial}{\partial x} + v'\frac{\partial}{\partial y}\right)(\bar{b}(y) + b'') = 0 \tag{84}$$

Expanding out, and neglecting small terms then,

$$\frac{\partial}{\partial t} + \bar{u}\frac{\partial}{\partial x}b'' + v'\frac{\partial \bar{b}}{\partial y} \tag{85}$$

Using hydrostatic balance, geostrophic balance and thermal wind balance, then this can be written in terms of the geostrophic streamfunction:

$$\left(\frac{\partial}{\partial t} + \bar{u}\frac{\partial}{\partial x}\right)\frac{\partial \psi}{\partial \blacksquare} + \frac{\partial \psi}{\partial x}\frac{\partial^2 \psi}{\partial \blacksquare \partial y} = 0 \tag{86}$$

Now, $\bar{u} = A\blacksquare$ so

$$\frac{\partial}{\partial \blacksquare}\left(\frac{\partial \psi}{\partial y}\right) = -\frac{\partial u}{\partial \blacksquare} = -A \tag{87}$$

and therefore, the equations on the top and bottom boundaries are:

$$\boxed{\left(\frac{\partial}{\partial t} + \bar{u}\frac{\partial}{\partial x}\right)\frac{\partial \psi}{\partial \blacksquare} = A\frac{\partial \psi}{\partial x}}. \tag{88}$$

Thus, the temperature field is determined by linear advection, and (on the RHS of the equation), by the meridional wind forcing.

B.4 Equations in the Interior

The QGPV equation (78) can be expanded (and small terms neglected):

$$D_g q = 0 \Rightarrow \left(\frac{\partial}{\partial t} + (\bar{u} + u')\frac{\partial}{\partial x} + v'\frac{\partial}{\partial y}\right)(\bar{q} + q') = 0 \tag{89}$$

to give linear advection of q in the interior:

$$\boxed{\left(\frac{\partial}{\partial t} + \bar{u}\frac{\partial}{\partial x}\right)q' = 0} \tag{90}$$

where the equation for QGPV is given by:

$$\boxed{q' = \frac{\partial^2 \psi'}{\partial x^2} + \frac{f_0^2}{N^2}\frac{\partial^2 \psi'}{\partial \blacksquare^2}}. \tag{91}$$

Equations (88, 89 and 90) give the equations for the Eady Model. The meridional wind propagates information between the edge waves to give vertical coupling. In the 2D Eady model, the meridional wind is not defined explicitly, but as a forcing to the boundary equations, and is in fact the vital part for the growth and decay of the waves.

where $\tilde{x}_0 = \tilde{x}_0(\tilde{\blacksquare})$, then we fi

References

Badger, J. and Hoskins, B. J. (2001). Simple initial value problems and mechanisms for baroclinic growth. *Journal of the Atmospheric Sciences*, **58**(1), 38–49.

Beale, E. M. L. (1972). A derivation of conjugate gradients. In F. A. Lootsma, editor, *Numerical Methods for Nonlinear Optimization*, pages 39–43. Academic Press.

Beale, E. M. L. (1988). *Introduction to Optimization*, chapter 4. Wiley.

Bouttier, F. and Courtier, P. (1999). Data assimilation concepts and methods. Meteorological training course lecture series, ECMWF.

Chao, W. and Chang, L. (1992). Development of a four-dimensional variational analysis system using the adjoint method at GLA. part 1: Dynamics. *Monthly Weather Review*, **120**(8), 1661–1673.

Courtier, P., Thepaut, J.-N., and Hollingsworth, A. (1994). A strategy for operational implementation of 4d-var, using an incremental approach. **120**, 1367–1387.

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

Errico, F. M. (1997). What is an adjoint model? *Bulletin of the American Meteorological Society.*

Farrell, B. (1982). The initial growth of disturbances in a baroclinic flow. *Journal of the Atmospheric Sciences*, **39**(8), 1663–1686.

Fletcher, S. (1999). *Numerical Approximations to Buoyancy Advection in the Eady Model.* Master's thesis, Department of Mathematics, University of Reading.

Gill, P. E., Murray, W., and Wright, M. H. (1981). *Practical Optimization.* Academic Press.

Griffith, A. (1997). *Data Assimilation for Numerical Weather Predicition using Control Theory.* Ph.D. thesis, University of Reading.

Holton, J. (1992). *An Introduction to Dynamic Meteorology.* Academic Press.

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

Lea, D. J. (2001). *Joint Assimilation of Sea Surface Temperature and Sea Surface Height.* Ph.D. thesis, University of Oxford.

Li, Y., Navon, I., Yang, W., Zou, X., Bates, J., Moorthi, S., and Higgins, R. (1994). Four-dimensional variational data assimilation experiments with a multilevel semi-lagrangian semi-implicit general circulation model. *Monthly Weather Review*, **122**, 966–983.

Lorenc, A. (1986). Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, **112**, 1177–1194.

Nash, J. C. (1990). *Compact Numerical Methods for Computers:Linear Algebra and Function Minimization*, pages 186–206. Adam Hilger, IOP Publishing, 2 edition.

Navon, I., Zou, X., Derber, J., and Sela, J. (1992). Variational data assimilation with an adiabatic version of the NMC spectral model. *Monthly Weather Review*, **120**, 1433–1446.

Navon, I. M. and Legler, D. M. (1987). Conjugate-gradient methods for large-scale minimization in meteorology. *Monthly Weather Review*, **115**, 1479–1502.

Press, W., Teukolsky, S., Vetterling, W., and Flannery, B. (1992). *Numerical Recipes in Fortran 77*, volume 1. Academic Press.

Rabier, F., Thepaut, J., and Courtier, P. (1998). Extended assimilation and forecast experiments with a 4d var assimilationssystem. *Quarterly Journal of the Roy*