FOUR DIMENSIONAL VARIATIONAL DATA ASSIMILATION IN THE PRESENCE OF IDEALISED RAPIDLY GROWING WEATHER SYSTEMS

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

S. P. Ballard [3] and A.S. Lawless [2]

NUMERICAL ANALYSIS REPORT 1/02

Department of Mathematics

University of Reading

Whiteknights, PO Box 220

Reading, RG6 2AX

This work was funded by the Natural Environment Research Council (NERC) and supported by the Met Office through a CASE studentship.

[1]Department of Meteorology, University of Reading, UK
[2]Department of Mathematics, University of Reading, UK
[3]Met Office, UK

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

List of Figures

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.

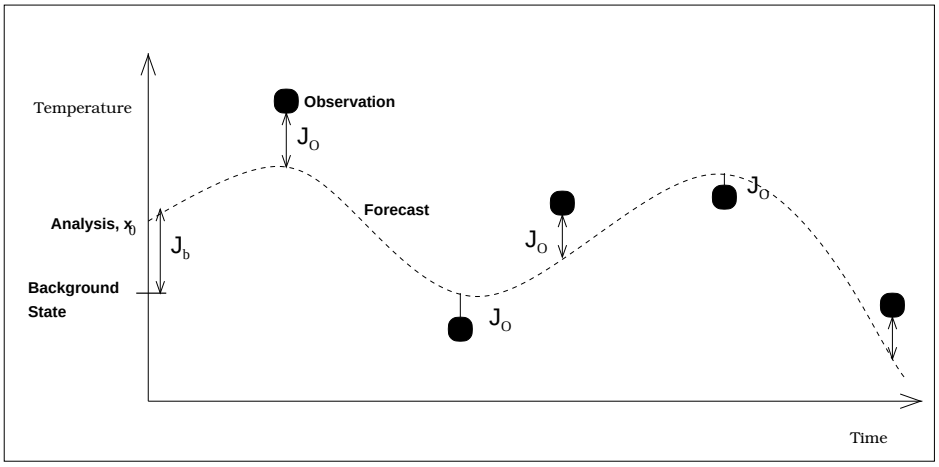


Figure 1: Schematic diagram of the Four Dimensional Variational data assimilation method: minimise the squared distance between the analysis and the background state at the beginning of the assimilation window, and the squared distance between the observations and the forecast state throughout the assimilation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

In general, the cost function J contains a background term J_b and an observational term J_o.

$$J(\mathbf{x}_0) = J_b + J_o \tag{1}$$

where the background term is

$$J_b = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) \tag{2}$$

where $\mathbf{x}$ is the state vector, $\mathbf{x}_0$ is the state vector at the initial time, used as the control variable (updated), $\mathbf{x}_0^b$ is the background state at the initial time, and $\mathbf{B}$ is the background error covariance matrix. The observational term is

$$J_o = \frac{1}{2}\sum_{i=1}^{N}(\mathbf{H}_i\mathbf{x}_i - \mathbf{y}_i)^T \mathbf{R}_i^{-1}(\mathbf{H}_i\mathbf{x}_i - \mathbf{y}_i) \tag{3}$$

where $\mathbf{y}_i$ is a vector of observations at time t_i, $\mathbf{H}_i$ is the observation operator, which converts from model space to observations space and $\mathbf{R}$ is the observational error covariance matrix.

The 4D Var problem is then to minimise $J(\mathbf{x}_0)$ subject to the strong constraint(assume a perfect model) that the sequence of model states must also be a solution of the nonlinear model equations $\dot{\mathbf{x}} = \mathbf{M}(\mathbf{x}, t)$. The 4D Var method is illustrated in Figure 1. In general this is a nonlinear constrained optimization problem. However, by linearizing the dynamical model and using an adjoint model, the problem can be transformed into an unconstrained quadratic minimisation through the use of lagrange multipliers.

The work in this report aims to investigate what the capabilities and weaknesses of 4D Var are in relation to 3D Var, by testing the the 4D Var method in a theoretical case of baroclinic instability which leads to the rapid development of mid-latitude cyclones.

1.2 Baroclinic Instability

The meridional (north-south) temperature gradient of the atmosphere is associated with a vertical shear of the zonal (west-east) wind. The vertical shear of the wind results in baroclinic instability, and hence mid-latitude cyclones (Holton, 1992). The baroclinic instability mechanism is characterised by a pressure field that has a westward tilt with increasing height and a temperature field that has an eastward tilt with increasing height. This mechanism is illustrated in Figure 2. It is vital that a data assimilation system should be able to perform well in these situations in order to provide good forecasts for severe storms.

The implementation of 4D Var in an operational system at ECMWF showed an improvement in extratropical forecast scores, compared with 3D Var (Rabier *et al.*, 1998). However, poor analyses of mid-latitude storms are still generated from 4D Var. For example, the analyses of the storm 'Lothar' on the 26th December 1999, were not as intense as the observed system, mainly due to the mishandling of observations (Ulbrich *et al.*, 2001).

2 Identical Twin Experiments

The 4D Var method using the Eady model is investigated using identical twin experiments, so that observational and model errors can be ignored.

Identical twin experiments are commonly used to investigate data assimilation methods. First, synthetic observations are generated by integrating the forward model to give the 'true' solution. It can then be assumed that the model and observations are perfect, or known errors can be added. All experiments, throughout this report, use perfect synthetic observations.

2.1 Mathematical Formalism: 4D Var

Suppose we have a set of N vectors of observations $\mathbf{y}_i$ $i = 1, \ldots, N$ at discrete times $0 \leq t_i \leq T$ within an assimilation window of length T. We also suppose that the observations are a linear combination of the true signal $\mathbf{x}_i^t$ and some random noise $\boldsymbol{\varepsilon}_i^o$ that is assumed to have a gaussian distribution

$$\mathbf{y}_i = \mathbf{H}\mathbf{x}_i^t + \boldsymbol{\varepsilon}_i^o \ (i = 1, \ldots, N) \tag{4}$$

where $\mathbf{H}$ is the linear observation operator which converts from model variables to 'observed' parameters. Suppose we also have a background state at the initial time, $\mathbf{x}_0^b$. We assume that this also has noise $\boldsymbol{\varepsilon}^b$ that has a gaussian distribution

$$\mathbf{x}_0^b = \mathbf{x}_0^t + \boldsymbol{\varepsilon}^b. \tag{5}$$

The aim is to find the state vector at the initial time $\mathbf{x}_0$ which minimises the variance of the analysis error and also satisfies the model equations throughout the assimilation time window Bouttier and Courtier (1999). The model is assumed to be linear as the Eady model (used in this work) is linear. Denoting the linear model integrated from time t_i to t_{i+1} by $\mathbf{M}_{i+1}$, then we require $\mathbf{x_i}$ to satisfy

$$\mathbf{x}_{i+1} = \mathbf{M}_{i+1}\mathbf{x}_i \ (i = 1, \ldots, N-1). \tag{6}$$

In summary, we are to find the initial state $\mathbf{x}_0$ such that the cost function:

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}\sum_{i=1}^{N}(\mathbf{y_i} - \mathbf{H}\mathbf{x}_i)^T \mathbf{R_i}^{-1}(\mathbf{y}_i - \mathbf{H}\mathbf{x}_i]) \tag{7}$$

is minimised and $\mathbf{x}_i$ also satisfies the model equations $\mathbf{x}_{i+1} = \mathbf{M}_{i+1}\mathbf{x}_i$.

To be able to calculate the gradient of the cost function at the beginning of the assimilation window, an adjoint model is required. Simple linear algebra can be used to illustrate how the method works.

Let

$$J_o = \sum_{i=1}^{N} J_i \tag{8}$$

where

$$\begin{aligned} J_i &= \tfrac{1}{2}(\mathbf{y}_i - \mathbf{H}_i\mathbf{x}_i)^T \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathbf{H}_i\mathbf{x}_i) \\ &= \tfrac{1}{2}(\mathbf{y}_i - \mathbf{H}_i\mathbf{M}_i \ldots \mathbf{M}_1\mathbf{x}_0)^T \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathbf{H}_i\mathbf{M}_i \ldots \mathbf{M}_1\mathbf{x}_0) \end{aligned} \tag{9}$$

Then the gradient of J is given by:

$$\begin{aligned} \nabla_{\mathbf{x_0}} J_i &= -(\mathbf{H}_i\mathbf{M}_i \ldots \mathbf{M}_1)^T \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathbf{H}_i\mathbf{M}_i \ldots \mathbf{M}_1\mathbf{x}_0) \\ &= -\mathbf{M}_1^T \ldots \mathbf{M}_i^T \mathbf{H}_i^T \mathbf{d}_i \end{aligned} \tag{10}$$

where $\mathbf{d}_i = \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathbf{H}_i\mathbf{x}_i)$ denotes the departure vector at time t_i. Hence the gradient of the observation term is given by:

$$\nabla_{\mathbf{x}_i} J_o = -\mathbf{H}^T\mathbf{d}_0 + \mathbf{M}_1^T(\mathbf{H}_1^T\mathbf{d}_1 + \mathbf{M}_2^T(\mathbf{H}_2^T\mathbf{d}_2 + \ldots + \mathbf{M}_N^T\mathbf{H}_N^T\mathbf{d}_N)\ldots). \tag{11}$$

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 $\int_0^T F(x,t)dt$, 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

$$\mathcal{L} = \int_0^T (F(x,t) + \lambda(t)[\dot{x} - f(x,t)])dt \tag{12}$$

where $\lambda(t) \in C_2[0,T]$ is a lagrange multiplier or adjoint variable. Letting $G(x,\lambda,\dot{x},\dot{\lambda},t) = F(x,t) + \lambda(t)(\dot{x} - f(x,t))$ then this can be written

$$\mathcal{L} = \int_0^T (G(x,\lambda,\dot{x},\dot{\lambda},t))dt \tag{13}$$

Taking the first variation of $\mathcal{L}$ and using a Taylor series expansion, then

$$\delta\mathcal{L} = \int_0^T \left[\delta x G_x + \delta\dot{x}G_{\dot{x}} + \delta\lambda G_\lambda + \delta\dot{\lambda}G_{\dot{\lambda}}\right] dt \tag{14}$$

Note that $G_{\dot{\lambda}} = 0$. Applying integration by parts then

$$\delta\mathcal{L} = \int_0^T \left[\delta x\left(G_x - \frac{d}{dt}G_{\dot{x}}\right) + \delta\lambda\left(G_\lambda - \frac{d}{dt}G_{\dot{\lambda}}\right)\right] dt + \left[\delta x G_{\dot{x}}\right]_0^T + \left[\delta\lambda G_{\dot{\lambda}}\right]_0^T \tag{15}$$

Necessary conditions for $\delta\mathcal{L} = 0$ are then given by Euler's Equations:

$$\begin{array}{lll} G_x - \frac{d}{dt}G_{\dot{x}} = 0 & \Rightarrow & \dot{\lambda} = F_x - \lambda f_x \quad \text{(adjoint equation)} \\ G_\lambda - \frac{d}{dt}G_{\dot{\lambda}} = 0 & \Rightarrow & \dot{x} = f \quad\quad\quad\quad\quad \text{(state (forward) equation)} \end{array} \tag{16}$$

The transversality condition must also be applied. From

$$\delta\mathcal{L} = \left[\delta x G_{\dot{x}}\right]_0^T = \delta x(T)\lambda(T) - \delta x(0)\lambda(0) \tag{17}$$

then set $\lambda(T) = 0$. The gradient of $\mathcal{L}$ at the beginning of the assimilation window is found from the adjoint variable at the beginning of the window

$$\frac{\delta\mathcal{L}}{\delta x(0)} = -\lambda(0). \tag{18}$$

This theory can be extended to the multivariate discrete case: The lagrangian functional can be constructed to give:

$$\mathcal{L} = J + \sum_{i=1}^{N-1} \boldsymbol{\lambda}_{i+1}^T(\mathbf{x}_{i+1} - \mathbf{M}_{i+1}\mathbf{x}_i) \tag{19}$$

$$\begin{array}{ll} \boldsymbol{\lambda}_i = 0 & i = N \\ \boldsymbol{\lambda}_i = \mathbf{M}_i^T\boldsymbol{\lambda}_{i+1} - \nabla_{\mathbf{x}_i}J & i = N-1,\ldots,0. \end{array} \tag{20}$$

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

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

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

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

where $z \in [-\frac{1}{2}, \frac{1}{2}]$ is the height, $x \in [0, N]$ is the horizontal distance in the zonal direction, $t \in [0, T]$ is the time, $\psi = \psi(x, z, t)$ is the streamfunction (similar to pressure), $q = q(x, z, t)$ is the quasi-geostrophic potential vorticity (which combines both dynamical and thermodynamical information). There are also periodic boundary conditions, the mean of the streamfunction field is zero, and the initial state is prescribed by the interior QGPV and the buoyancy on the upper and lower boundaries:

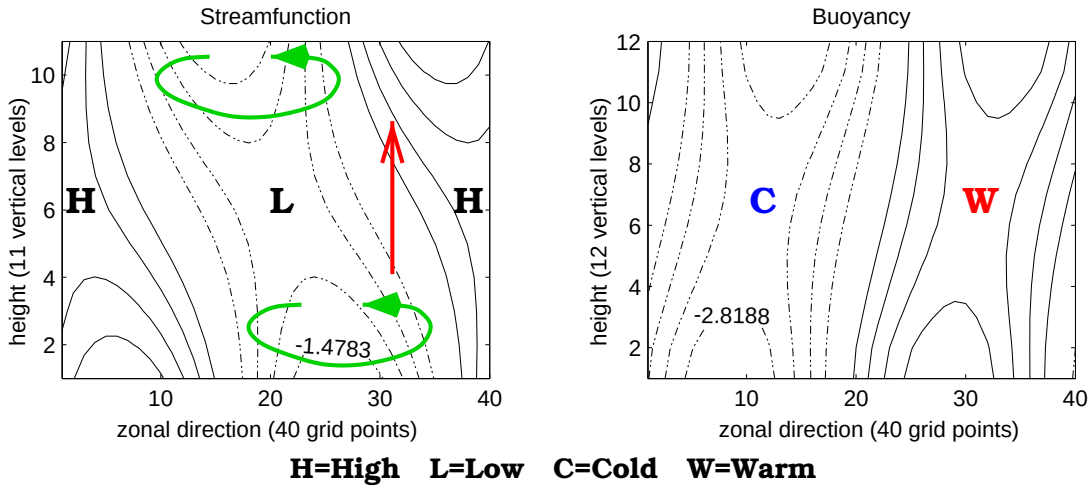
$$\begin{array}{ll} \psi(0, z, t) = \psi(N, z, t) & \text{in } z \in [-\frac{1}{2}, \frac{1}{2}], \\ \iint \psi \, dx \, dz = 0, & \text{in } x \in [0, N],\ z \in [-\frac{1}{2}, \frac{1}{2}], \\ q(x, z, 0) = q_0(x, z), & \text{in } x \in [0, N],\ z \in (-\frac{1}{2}, \frac{1}{2}), \\ \frac{\partial \psi}{\partial z}(x, 0) = \frac{\partial \psi_0}{\partial z}(x) & \text{on z} = \pm\frac{1}{2}. \end{array} \tag{24}$$

From the non-dimensional hydrostatic equation, $\frac{\partial \psi}{\partial z} = b$, where b is the buoyancy which is a measure of temperature and from geostrophic balance, $\frac{\partial \psi}{\partial x} = v$ where v is the meridional wind. Note that although there is no y-dependence in the model, it is the wind in the y-direction that provides the crucial forcing for baroclinic instability. The buoyancy and streamfunction

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.



The streamfunction field tilts to the west with increasing height. From the omega equation, there is ascent ahead of the cyclonic anomaly. This vertical motion gives vortex stretching and so intensifies the lower level cyclonic anomaly.

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

2.3 Numerical Models

The numerical Eady model has previously been used by Badger and Hoskins (2001) and Fletcher (1999). The advection equations are discretised using a leapfrog centred time centred space (CTCS) scheme (with a Forward Time Centred Space (FTCS) scheme for the first timestep). The matrix representation of the discrete laplace equation is defined using natural ordering. So that the mean of the streamfunction field is also set to zero, a small positive number is added to every element of the matrix. A NAG routine is used to perform an LU factorisation to solve the laplace equation.

An adjoint model of the Eady model has been created directly from the forward model code, in a similar way to Navon *et al.* (1992), and tested using the norm test and the gradient test as described in Li *et al.* (1994). The adjoint model equations are derived in appendix A.

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, giving slow convergence when the condition number of $\mathbf{H}$ is large.

The numerical steepest descent algorithm used in this report uses the gradients only to provide the search direction, and uses an inexact search to calculate the step length.

2.4.3 Conjugate Gradient Method

It is possible to use the gradient vectors to construct conjugate directions [1], that take into account the hessian matrix. A good description of conjugate directions is given in Shewchuk (1994). Conjugate directions can be constructed using the conjugate Gram-Schmidt process

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

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

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 on the unobserved top boundary. This experiment has been run until 200 iterations, with no further change in the rate of descent.

The other methods use the hessian information to speed up the rate of descent of the minimisation (Lea, 2001). The conjugate gradient method calculates conjugate directions, that take into account the shape of isocontours of J, that is given by the hessian matrix, whilst the quasi-Newton method evaluates an approximation to the hessian. Both methods give a dramatic increase in the rate of convergence, and the conjugate gradient method is slightly faster than the quasi-Newton method in this experiment. However, the conjugate-gradient performs more evaluations of the cost function (simulations), and so in this case is computationally more demanding. The limited-memory method combines the quasi-Newton method and the conjugate gradient method, but in this case, does not perform as well.

Therefore, the quasi-Newton method is used in all further experiments.

Termination criteria are required to be able to stop the minimisation algorithm. An experiment assimilating a horizontal line of the streamfunction field, with a zero background state is used to investigate the behaviour of the quasi-Newton method. The decrease of the cost function using the quasi-Newton method is shown in Figure 4(a). There are 44 evaluations of the cost function on the 7th iteration, showing that machine precision has been reached. It is important that the minimisation is terminated just before this point, so that computer time is not wasted. However, it is important that the minimum has been found before the algorithm is stopped.

The magnitude of a vector can be described using different norms. The 2-norm or Euclidean

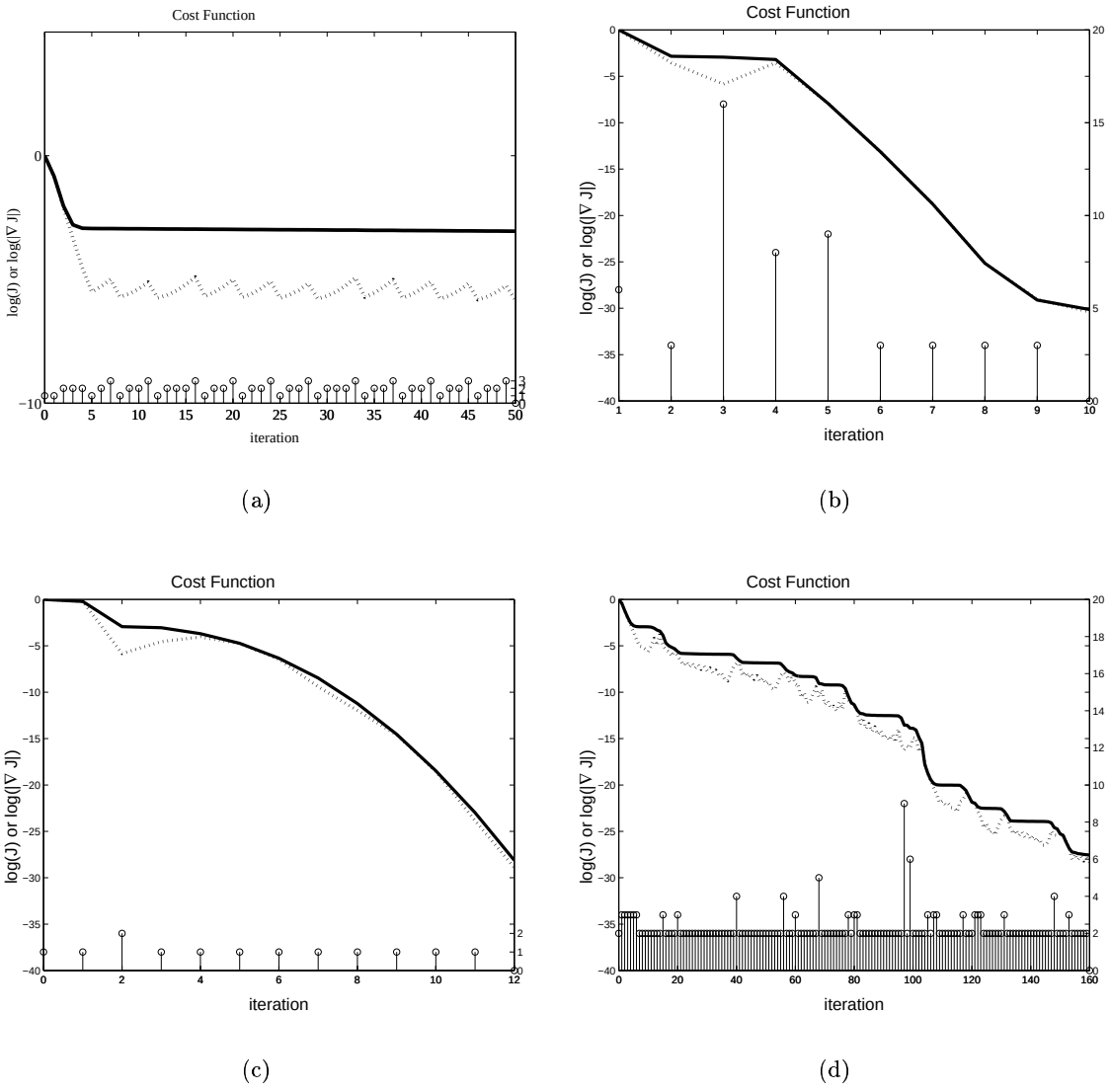


Figure 3: Behaviour of the cost function using (a) Steepest Descent Method and (b) Conjugate Gradient Method (c) Quasi-Newton Method (d) Limited Memory Method, with increasing 'iterations' or gradient evaluations. The solid line corresponds to the cost function J, and the dotted line corresponds to the 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.

norm is defined by

$$\|\mathbf{x}\|_2 = \left(\sum_i x_i^2 \right)^{\frac{1}{2}}, \tag{34}$$

and the infinity norm or maximum norm is defined by

$$\|\mathbf{x}\|_\infty = \max_i |x_i|. \tag{35}$$

The absolute error of the gradient of J is defined as:

$$abs(\nabla J) = \|\nabla J\|, \tag{36}$$

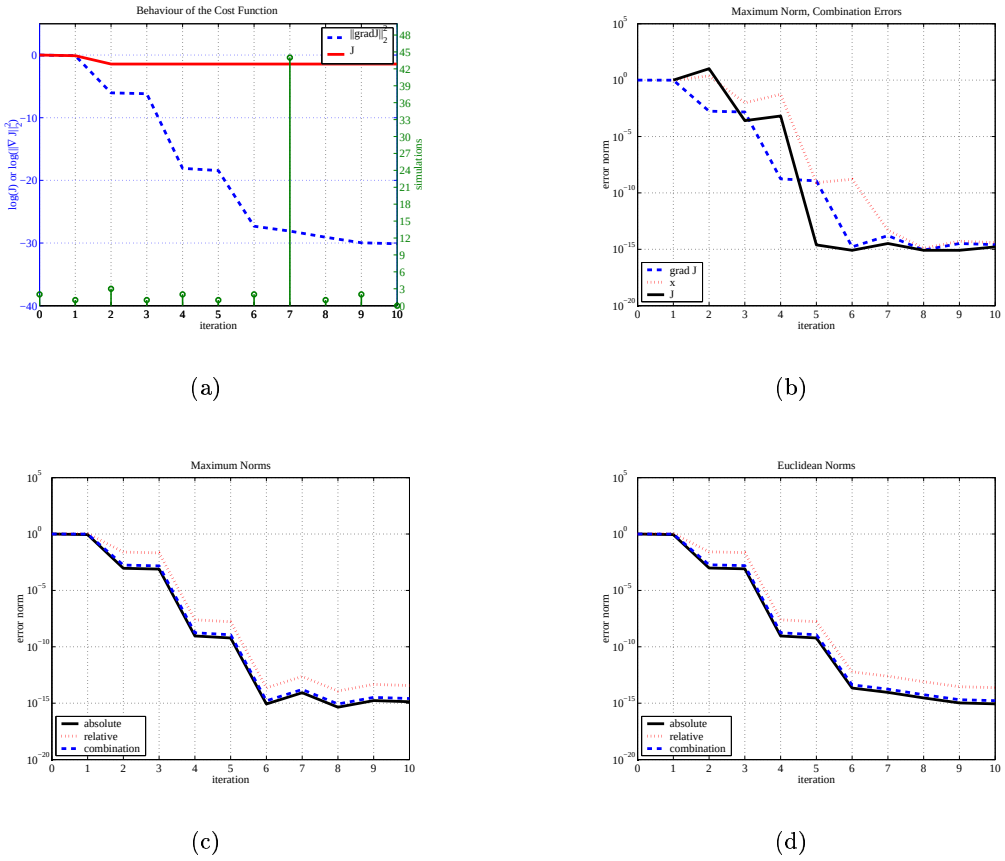


Figure 4: The behaviour of the cost function with increasing iterations (a) cost function (solid line), squared 2-norm of the gradient vector (dashed line), simulations or function evalutions (circles) (b) The set of termination criteria (equations 39, 40, 41) (c)Comparison of Maximum Norms of ∇J (equation 35) (d)Comparison of Euclidean Norms of ∇J (equation 34).

the relative error is defined as:

$$rel(\nabla J) = \frac{\|\nabla J\|}{J}, \tag{37}$$

and the combination error (Gill *et al.*, 1981) is defined as:

$$comb(\nabla J) = \frac{\|\nabla J\|}{1 + J}. \tag{38}$$

Figure 4 (c) shows a comparison of the absolute, relative and combination errors defined using the maximum norm. It is harder to specify a tolerance for the absolute error as it does not take into account the value of J. Although the relative error does take the size of J into account, there are problems when J is zero at the minimum. The combination error is a combination of both errors, and has a value in between the absolute error and the relative error in Figure 4 (c).

Figure 4 (d) shows a similar comparison, but using the Euclidean norm. The graphs (c) and (d) are very similar until the 6th iteration, where the maximum norm reaches a lower value.

From these comparisons a maximum norm, combination error has been chosen for the set of termination criteria, which are defined as:

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

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

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

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

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

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

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

$$\hat{\mathbf{H}} = \begin{pmatrix}\mathbf{H}\\ \mathbf{HM}\end{pmatrix} \tag{45}$$

to have rank n. This is equivalent to saying that $\hat{\mathbf{H}}^{-1}$ exists and that $\hat{\mathbf{H}}\mathbf{x} = \mathbf{y}$ has a unique solution $\mathbf{x}$ for every $\mathbf{y}$. If this is the case, then the system is said to be *observable*. If the matrix is not of full rank, then it means that the same information is given at least twice.

[1] The null space of a matrix $\mathbf{A}$ is the set of solutions to $\mathbf{Ax} = \mathbf{0}$. The dimension of the null space is equal to the number of zero singular values. The null space corresponds to the undetermined part of state space Strang (1986)

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 identical twin experiments using the 4D Var method with the Eady model and the quasi-Newton minimisation method, are now presented. All the experiments in this report use perfect observations.

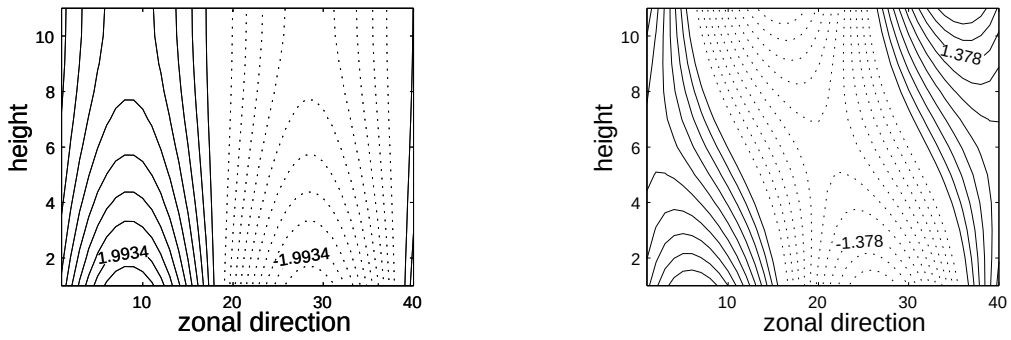


Figure 5: Streamfunction Analysis from (a) 3D Var and (b) 4D Var where there are no observations given on the upper boundary. 4D Var is able to reconstruct the upper wave, through the use of the model dynamics.

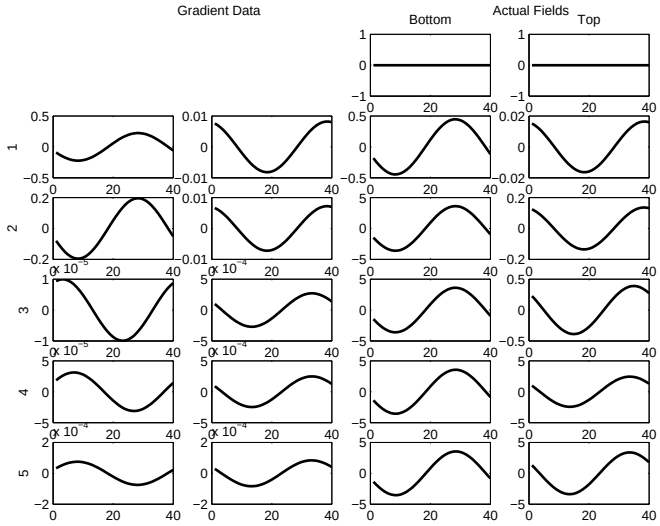


Figure 6: The behaviour of the solution with increasing iterations for the Quasi-Newton Method. The columns on the left show $-\nabla_{\mathbf{b}_0} J$,$-\nabla_{\mathbf{B}_1} J$ and the columns on the right show b_0 and b_1.

4.1 No Background Term

In the following numerical experiment, the buoyancy on the lower boundary and the interior zero QGPV are 'observed'. However, the upper boundary buoyancy is not observed. With 2 time levels for the assimilation window, this system is observable (as proved in section 3).

The results are shown in Figure 5. The 3D Var uses only the bottom boundary and interior potential vorticity as the control variable, and the top boundary is left as zero. This is a trivial case, so the minimisation reaches a minimum in one iteration. However, the 4D Var system is able to propagate information from the observations on the lower level through the model dynamics to reconstruct the upper level wave that is needed to give the growth of the observed lower level wave.

It is also of interest to investigate how the information is propagated to the top boundary.

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 growing, the amplitude of the bottom wave is slightly too large at this point. However, on the next iteration, the phase of the upper boundary wave is shifted to the correct position. This wave is then in the optimal position to give the growth in amplitude of the analysed wave. Hence on further iterations, the amplitude of the lower level wave is reduced. Thus, the gradient field on the lower boundary on the fourth iteration is the complete opposite of that on the first iteration.

4.2 With the Background Term

Figure (7) shows the results from an experiment where observations are given again on the lower boundary and interior at two time levels. A zero background state $\mathbf{x}_0^b = 0$ is also included, with identity covariance matrices, so that the cost function is of the form:

$$J(\mathbf{x}_0) = \sigma_b^{-2}(\mathbf{x}_0 - \mathbf{x}_0^b)^2 + \sum_{i=0}^{1} \sigma_o^{-2}(\mathbf{y}_i - \mathbf{H}\mathbf{x}_i)^2 \tag{51}$$

In Figure 7(a) the respective weights to the J_o and J_b terms are such that equal weight is given to the background and the observations. The analysis of the buoyancy on the bottom boundary is halfway between the background state and the observations as expected with equal weights. However, the wave on the upper boundary has not been reconstructed. The weighting to the observations is increased in Figure 7(b). Although the analysis on the lower boundary is very close to the truth, there is still very little information on the top boundary. Increasing the weighting so that the observed field is given 1000 times more weighting than the background field (and note that the observed field is also given on two time levels, which effectively doubles the weighting), means that there is a wave on the top boundary. However, the wave still has an amplitude that is too small.

From these results, we see that the background state has a significant impact on the regions which are not observed, but obtain information through the model dynamics.

4.3 Assimilating Temperature Observations

The assimilation of temperature (buoyancy) observations is now considered. The control variable is still given by buoyancy on the upper and lower boundaries and quasi-geostrophic potential vorticity in the interior. However the forward observation operator is modified to convert from buoyancy on the boundaries and QGPV in the interior to streamfunction using the laplace solver. Then, the vertical derivative is taken to give the buoyancy field throughout the whole domain. The adjoint of this operator is also required, and is derived from the discrete equations, given in appendix A.

A background state is required so that the problem is well-determined. Therefore in the following experiments a zero background state is specified, with a 20 time level assimilation

window.

4.3.1 Growing Eady Wave

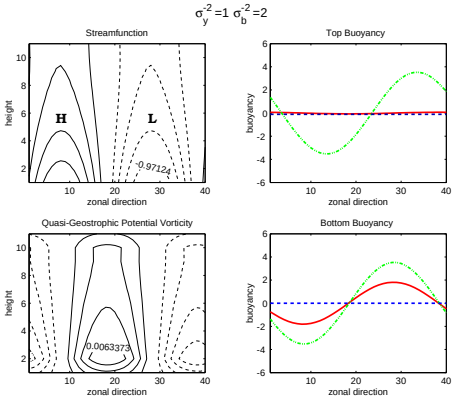
The true solution is given by the most unstable Eady wave. A horizontal line of buoyancy is observed at the end of the assimilation window. The analysis, shown in Figure 8, is close to zero at the beginning of the assimilation window, but has a wave structure at the end. This fast growing solution is characterised by the streamfunction field that tilts to the west with increasing height.

4.3.2 Decaying Eady Wave

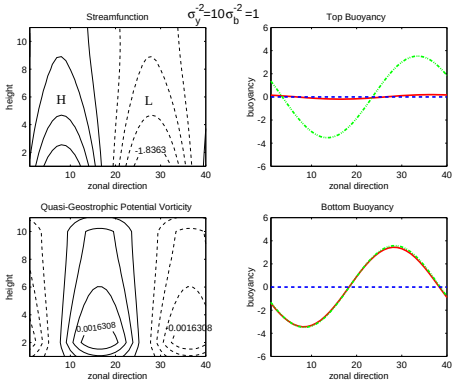
Now the true solution is given by the most unstable decaying Eady wave. The streamfunction fields tilts to the east with increasing height, and the temperature field tilts to the west. With the horizontal line of observations at the end only, we find (trivially) that the results are exactly the same as for the growing Eady wave. Instead of a decaying solution, as partly observed, a growing solution is obtained, due to the zero background state.

The experiment is repeated, but a horizontal line of buoyancy is observed at every time level the assimilation window. The analysis is shown in Figure 9. The 4D Var scheme is provided with conflicting information. The zero background state means that the solution should grow, however, a decaying solution is observed. The analysis decays for the first part of the time window and grows for the second part.

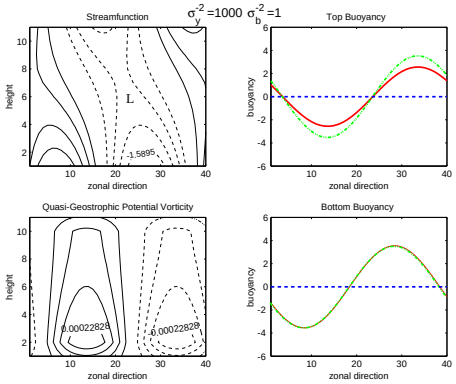
The growth rate throughout the time window depends on the relative weight given to the background state, as shown in Figure 10. With a large weight given to the observations, the analysis also decays throughout the analysis window. With a large weight given to the background state, the analysis grows throughout the window. With equal weights we see that the analysis decays through the first part and grows through the second part of the window.



(a)



(b)



(c)

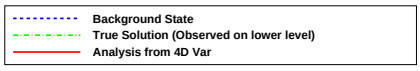
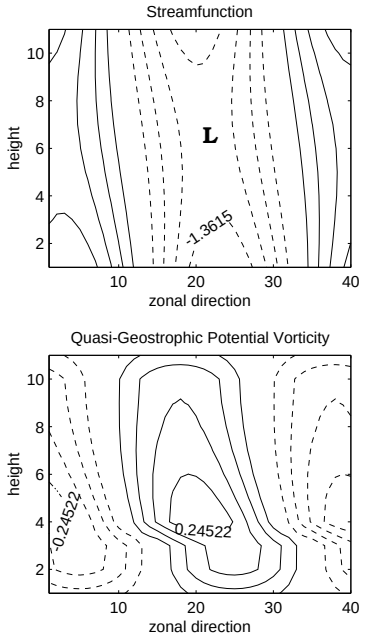
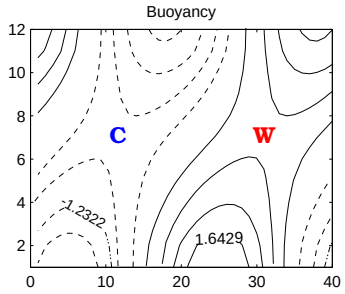


Figure 7: Analyses from a two timelevel assimilation with no observations on the top level and a zero background state, with (a) Effectively equal weights, (b) Weight given to the observations is 10 times greater than the background state weight, (c) Weight given to the observations is 1000 times greater 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

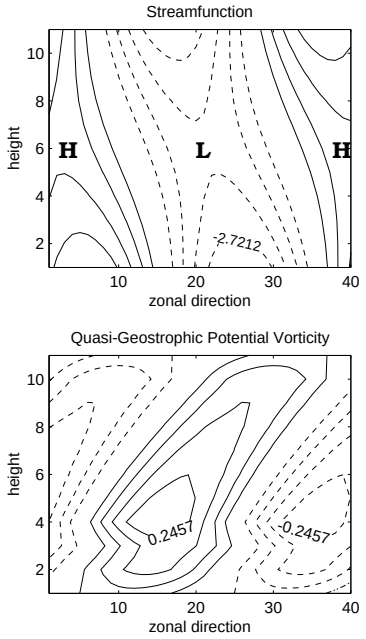


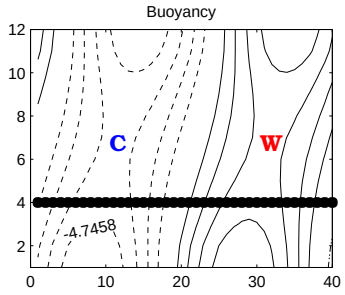


Analysis at the beginning of the assimilation window

Weights: Jb=1 Jo=1

(a)



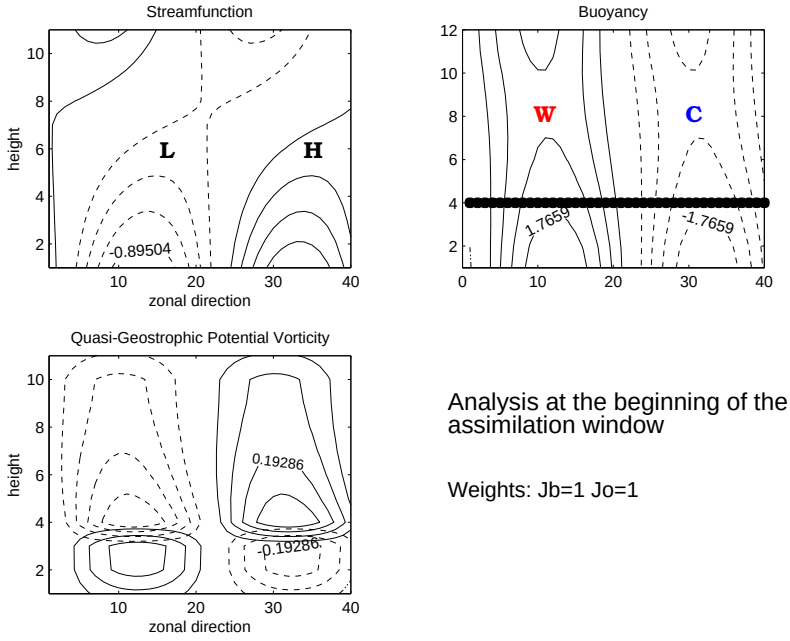


Analysis at the end of the assimilation window

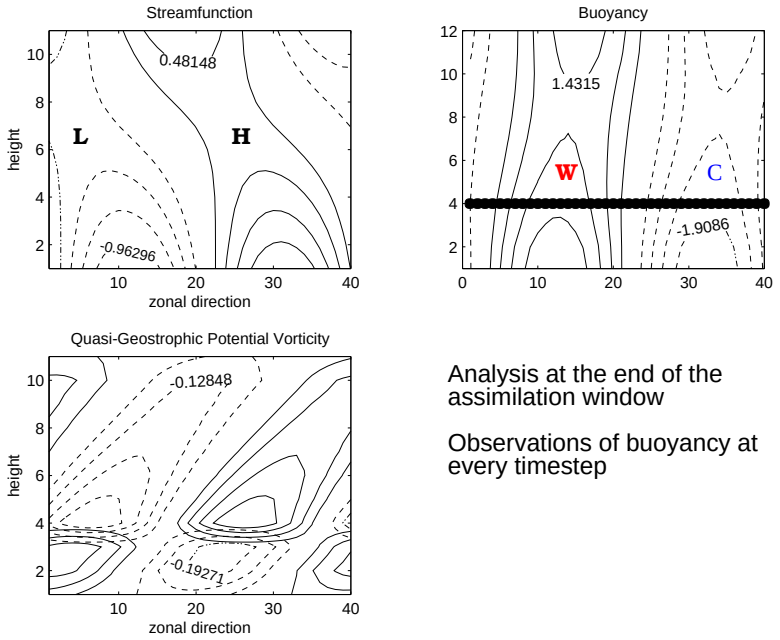
Horizontal line of buoyancy is observed at this time - given by the black dots

(b)

Figure 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.

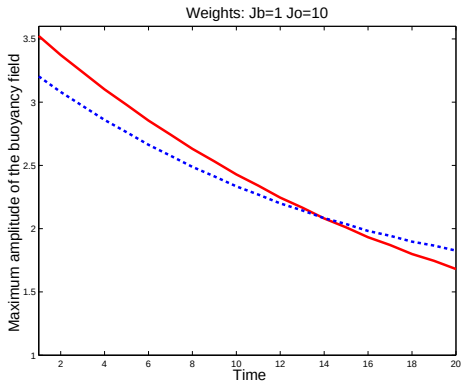


(a)

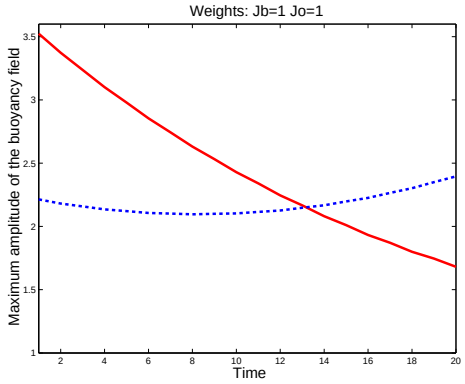


(b)

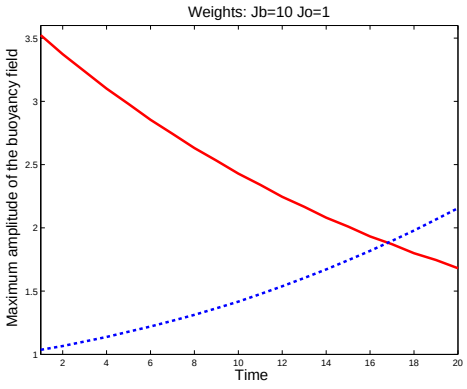
Figure 9: Decaying Eady Wave, Analysis from 4D Var. A horizontal line of the buoyancy field is observed at every time level, with a zero background state, and identity covariance matrices.



(a)



(b)



(c)

Figure 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.

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.

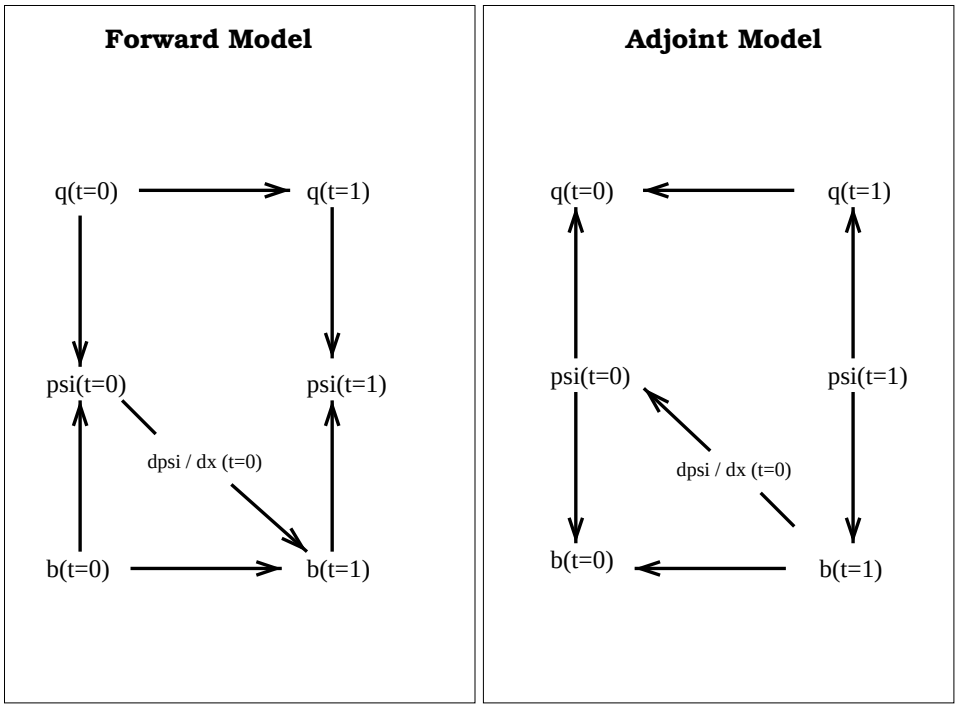


Figure 11: The structure of the linear models. In the forward model, the initial conditions are given by potential vorticity,q and buoyancy, b at time t=0. b(t=0) and q(t=0) are then used to calculate the streamfunction, psi at the same time. The QGPV and buoyancy are then advected to the next time level t=1, with the buoyancy also forced by the streamfunction field. The arrows show the direction of propagation of information.

A.1 General Advection Equation

Consider the general structure of an advection equation:

$$A = B + \alpha(C - D) + \beta E \tag{52}$$

In matrix form this can be written as:

$$\begin{pmatrix} B \\ C \\ D \\ E \\ A \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & \alpha & -\alpha & \beta & 0 \end{pmatrix} \begin{pmatrix} B \\ C \\ D \\ E \\ A \end{pmatrix} \tag{53}$$

The adjoint of this matrix is:

$$\begin{pmatrix} \hat{B} \\ \hat{C} \\ \hat{D} \\ \hat{E} \\ \hat{A} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & \alpha \\ 0 & 0 & 1 & 0 & -\alpha \\ 0 & 0 & 0 & 1 & \beta \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{B} \\ \hat{C} \\ \hat{D} \\ \hat{E} \\ \hat{A} \end{pmatrix} \tag{54}$$

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:

$$\begin{array}{l} \hat{\psi}_{j+1} = \hat{\psi}_{j+1} + \frac{1}{2\Delta x}\hat{F}_j \\ \hat{\psi}_{j-1} = \hat{\psi}_{j-1} - \frac{1}{2\Delta x}\hat{F}_j \\ \hat{F}_j = 0 \end{array} \tag{65}$$

Note that this is equivalent to the adjoint equation:

$$\hat{\psi}_j = -\frac{F_{j+1} - F_{j-1}}{2\Delta x} \tag{66}$$

A.5 Solving the Laplace Equation

To solve the equations

$$\nabla^2 \psi = q \tag{67}$$

subject to $b = \frac{\partial \psi}{\partial z}$ on the top and bottom boundaries and periodic boundaries, we solve the matrix equation:

$$\mathbf{A}\psi = \mathbf{B} \tag{68}$$

for ψ, where $\mathbf{B} = \begin{bmatrix} \mathbf{q}_{\text{bottom}} + \mathbf{b_0} \\ \mathbf{q}_{\text{interior}} \\ \mathbf{q}_{\text{top}} + \mathbf{b_1} \end{bmatrix}$.

The adjoint of this is to solve

$$\mathbf{A}^T \hat{\mathbf{B}} = \hat{\psi} \tag{69}$$

for $\mathbf{B}$. In this case $\mathbf{A}$ is symmetric, so $\mathbf{A} = \mathbf{A}^T$.

B Derivation of the Eady model equations

The non-dimensional equations and normal mode solutions used in this report are now derived.

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-Geostrophic Potential Vorticity (QGPV) equation, and the QG Thermodynamic equation which are given in Holton (1992).

The QG Thermodynamic equation states that potential vorticity is conserved following the flow (assuming dry, adiabatic motion). More specifically, it states that as air rises, it cools via adiabatic expansion.

$$\frac{D\theta}{Dt} = 0 \tag{70}$$

where θ is potential temperature. We consider the perturbations from a hydrostatically balance reference state that is only a function of height, z (Boussinesq approximation) to remove the large static state of the atmosphere. Let $\theta = \bar{\theta}(z) + \theta'(x, y, t)$, then

$$\frac{D}{Dt}(\bar{\theta} + \theta') = 0. \tag{71}$$

Expand to give

$$\left(\frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y}\right)\theta' = -w\frac{\partial}{\partial z}\bar{\theta} \tag{72}$$

Assuming that the winds are close to geostrophic and defining $D_g = \frac{\partial}{\partial t} + u_g\frac{\partial}{\partial x} + v_g\frac{\partial}{\partial y}$ where g denotes geostrophic then

$$D_g\theta' + w\frac{\partial\bar{\theta}}{\partial z} = 0 \tag{73}$$

If we now define buoyancy b, by

$$b' = \frac{g}{\bar{\theta}}\theta' \tag{74}$$

and static stability N^2 or Brunt-Vaisala frequency N as

$$N^2 = \frac{g}{\bar{\theta}}\frac{\partial\bar{\theta}}{\partial z} \tag{75}$$

then the QG thermodynamic equation can be rearranged to give

$$D_g b' + wN^2 = 0. \tag{76}$$

The horizontal momentum and continuity equations can be used to derive the QG Vorticity Equation:

$$D_g(f + \xi_g) = f_0\frac{\partial w}{\partial z}. \tag{77}$$

This equation states that if a column of air is stretched vertically, then its absolute vorticity $(f + \xi_g)$ will increase.

The QG vorticity equations (equation (77)) and the QG thermodynamic equation (equation (76)) can be combined by eliminating the vertical velocity w, to give the Quasi-Geostrophic Potential Vorticity (QGPV) equation:

$$D_g q = 0 \tag{78}$$

where

$$q = f + \xi_g + \frac{\partial}{\partial z}\left(\frac{f_0 b'}{N^2}\right). \tag{79}$$

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, $\bar{u} = Az$. The vertical wind shear is associated with a meridional temperature gradient, as given by the thermal wind equation

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

The thermal wind relation is shown in Figure 12. Although the (2D) Eady model equations only describe the flow in the zonal and height directions, they do incorporate the meridional temperature gradient, which is vital for the baroclinic instability mechanism.

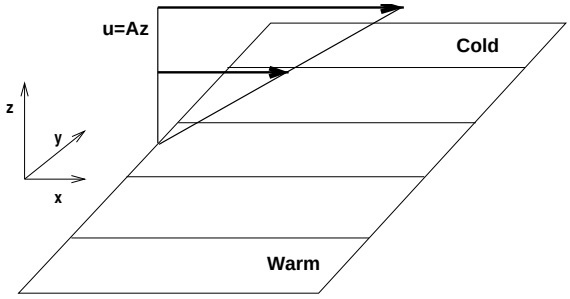


Figure 12: Basic state of the Eady model:The meridional temperature gradient is proportional to the zonal wind shear, through thermal wind balance

The geostrophic streamfunction ψ is defined by:

$$\psi = \frac{\phi'}{f_0} \tag{81}$$

where ϕ is the geopotential height. The equations of geostrophic balance (from the horizontal momentum equations) can then be written in the form:

$$u_g = -\frac{\partial \psi}{\partial y} \quad v_g = \frac{\partial \psi}{\partial x}. \tag{82}$$

Similarly, hydrostatic balance (from the vertical momentum equation) can be written as:

$$b' = f_0 \frac{\partial \psi}{\partial z}. \tag{83}$$

Assuming that the perturbation is independent of y and we then linearise the equations using,

$$\begin{array}{lcll} u_g & = & \bar{u}(z) & +u'_g(x,z,t) \\ v_g & = & v'_g(x,z,t) & \\ q & = & \bar{q}(y) & +q'(x,z,t) \\ b' & = & \bar{b}(y) & +b''(x,z,t) \end{array}$$

Note that $\bar{q} = f$, the coriolis parameter.

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:

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

Now, $\bar{u} = Az$ so

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

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

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

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:

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

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.

B.5 Non-dimensionalising and Co-ordinate Change

To eliminate the constants within the equations, we now apply a coordinate change to non-dimensionalise the problem, and also so that the x-coord is in the middle of the domain, and the origin of z is at mid-levels.

$$\tilde{z} = \frac{z - \frac{H}{2}}{H} \tag{92}$$

$$\tilde{x} = \frac{x - \frac{AH}{2}t}{L_R} \tag{93}$$

$$\tilde{t} = \frac{fA}{N}t \tag{94}$$

where $L_R = \frac{NH}{f}$ is the Rossby radius of deformation. Then the equations can be rewritten as:

$$\left(\frac{\partial}{\partial\tilde{t}} + \tilde{z}\frac{\partial}{\partial\tilde{x}}\right)\frac{\partial\tilde{\psi}'}{\partial\tilde{z}} = \frac{\partial\tilde{\psi}'}{\partial\tilde{x}}. \tag{95}$$

$$\left(\frac{\partial}{\partial\tilde{t}} + \tilde{z}\frac{\partial}{\partial\tilde{x}}\right)\tilde{q}' = 0 \tag{96}$$

where

$$\tilde{q}' = \frac{\partial^2\tilde{\psi}'}{\partial\tilde{x}^2} + \frac{\partial^2\tilde{\psi}'}{\partial\tilde{z}^2} = 0. \tag{97}$$

and

$$\tilde{\psi}' = \frac{\psi'}{\psi_0} \tag{98}$$

$$\tilde{q}' = \frac{L_R^2}{\psi_0}q' \tag{99}$$

and ψ_0 is the ampliude of ψ'.

An extra equation is needed to ensure that the model is well-posed. We impose the condition that

$$\int\int \tilde{\psi}' dx dz = 0 \tag{100}$$

so that the mean value of the streamfunction in the domain is zero.

B.6 Solutions of the equations

A growing and decaying wave-like solutions of the form

$$\tilde{\psi}'(x,z,t) = \hat{\psi}(z)e^{-i\tilde{k}(\tilde{x}-\tilde{c}\tilde{t})} \tag{101}$$

where $\tilde{k}$ is the wave number, and $\tilde{c}$ is the phase speed, is substituted into equation (97) too give

$$0 = -\tilde{k}^2\hat{\psi} + \frac{\partial^2\hat{\psi}}{\partial z^2} \tag{102}$$

which has the solution

$$\hat{\psi} = C_1 sinh(\tilde{k}\tilde{z}) + C_2 cosh(\tilde{k}\tilde{z}). \tag{103}$$

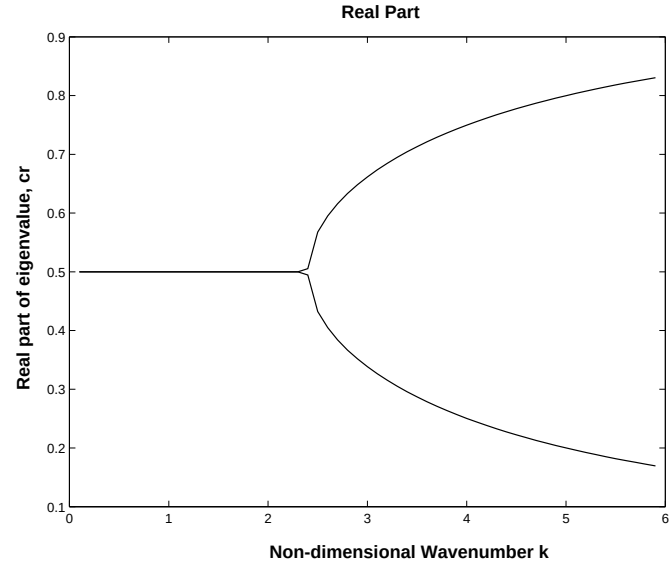


Figure 13: Real Part: Neutral Modes corresponding to boundary waves for $\tilde{k} > 2.4$

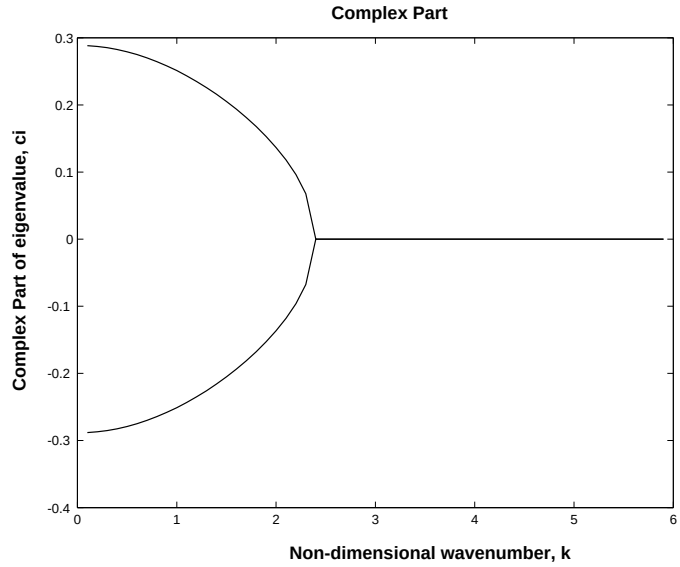


Figure 14: Complex Part: Exponentially growing and decaying solutions for $\tilde{K} < 2.4$

Therefore the normal mode solutions can be written as:

$$\tilde{\psi} = \psi_0 e^{\tilde{\sigma}\tilde{t}}[cosh(\tilde{k}\tilde{z})cos(\tilde{k}\tilde{x}) - \alpha sinh(\tilde{k}\tilde{z})sin(kx)] \tag{104}$$

where the non-dimensional growth rate is $\tilde{\sigma} = i\tilde{k}\tilde{c}$.

By substituting this into equation (95), then we find that:

$$\alpha = \sqrt{\frac{1 - \frac{k}{2}tanh\frac{k}{2}}{\frac{k}{2}coth\frac{k}{2} - 1}} \tag{105}$$

and that

$$\tilde{c} = \frac{1}{\tilde{k}}\sqrt{\left(tanh\left(\frac{\tilde{k}}{2}\right) - \frac{\tilde{k}}{2}\right)\left(\frac{\tilde{k}}{2} - coth\left(\frac{\tilde{k}}{2}\right)\right)} \tag{106}$$

There are 2 modes of behaviour, depending on whether c is real or complex. If $\frac{\tilde{k}}{2} \geq coth\left(\frac{\tilde{k}}{2}\right)$ (Short Wavelength) then both roots are real corresponding to two trapped neutral boundary waves. If $\frac{\tilde{k}}{2} < coth\left(\frac{\tilde{k}}{2}\right)$ (Long Wavelength) then the roots are complex conjugates giving 2 normal modes. One root corresponds to the growing mode, whilst the other corresponds to the decaying mode. Thus there are 2 normal modes for each wave number, $\tilde{k}$. The critical value of $\tilde{k}$ is therefore at $\frac{\tilde{k}}{2} = coth\left(\frac{\tilde{k}}{2}\right)$, which gives $\tilde{k} = 2.3994$. These normal mode solutions are shown in Figures 13 and 14.

Note that the non-dimensional wavenumber is related to the dimensional wavenumber by

$$\tilde{k} = \frac{NH}{f}k = L_R k \tag{107}$$

so the wavelength is also dependent on the height of the domain, as well as the static stability and coriolis parameter.

The exponential growth rate $\sigma = ikc = kc_i$ is shown in Figure 15. The maximum growth is at $\hat{k} = 1.6$, corresponding to a growth rate of $\hat{\sigma} = 0.31$. This maximum growth rate corresponds to the most unstable Eady Wave.

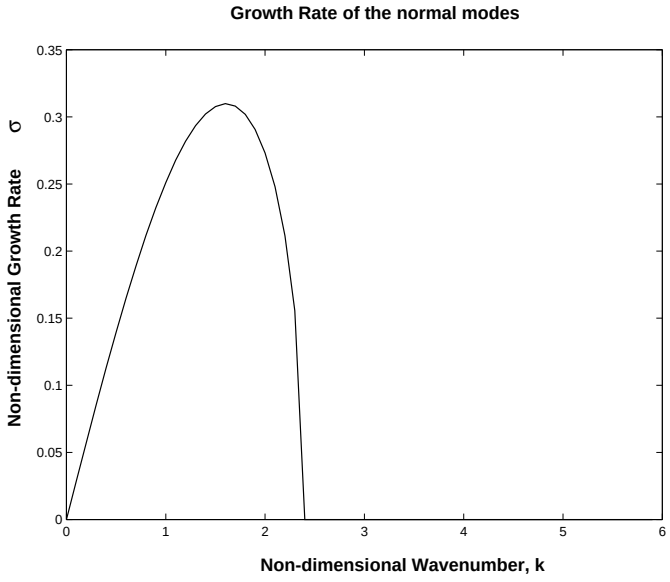


Figure 15: Growth Rate of Normal Modes

B.7 Vertical Structure of the Normal Mode Solutions

By rewriting the wave 104 in the form:

$$\tilde{\psi} = A(\tilde{z})sin[\tilde{k}(\tilde{x} - \tilde{x}_0)] = A(\tilde{z})[sin(\tilde{k}\tilde{x})cos(\tilde{k}\tilde{x}_0) - cos(\tilde{k}\tilde{x})sin(\tilde{k}\tilde{x}_0)] \tag{108}$$

where $\tilde{x}_0 = \tilde{x}_0(\tilde{z})$, then we find that, for the most unstable growing Eady wave,

$$tanh\tilde{x}_0 = \frac{1}{\alpha} coth(\tilde{k}\tilde{z}) \approx 1 \tag{109}$$

and hence $\tilde{k}\tilde{x}_0 \approx 45^o$. The wave tilts 45^o between mid levels and the tropopause and hence, tilts to the west by $\frac{\pi}{2}$ or $\frac{\lambda}{4}$ between the ground and the tropopause. Similarly, this can be repeated for the temperature field, by first finding the vertical derivative of the streamfunction. We find that the temperature field tilts eastward with height, and in this case $\tilde{k}\tilde{x}_0 = 23.7^o$.

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 Royal Meteorological Society*, **124**, 1861–1887.

Sasaki, Y. (1970). Some basic formalisms in numerical variational analysis. *Monthly Weather Review*, **98**(12), 875–883.

Shanno, D. F. and Phua, K. H. (1976). Algorithm 500:minimization of unconstrained multivariate functions. *ACM Transactions On Mathematical Software*, **2**, 87–94.

Shewchuk, J. R. (1994). An introduction to the conjugate gradient method without the agonizing pain, edition 1.25. Technical report, School of Computer Science, Carnegie Mellon University, Pittsburgh.

Strang, G. (1986). *Linear Algebra and its applications.* Harcourt Brace Jovanovich College Publishers.

Thepaut, J., Courtier, P., Belaud, G., and Lemaitre, G. (1996). Dynamical structure functions in 4d variational assimilation: A case study. *Quarterly Journal of the Royal Meteorological Society*, **122**, 535–561.

Ulbrich, U., Fink, A. H., Klawa, M., and Pinto, J. G. (2001). Three extreme storms over europe in december 1999. *weather*, **56**, 70–80.