Scale Analysis of Short Term Forecast Errors

Iain Davison

MSc MNMAO

University of Reading

August 2006

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Iain Davison

Acknowledgements

Completing this dissertation would not have been possible without the help and support of my supervisors Amos Lawless and Nancy Nichols. I would also to thank Sarah Dance for her help and input over the summer.

Thanks are also due to all the sta at the Met Oce who allowed me to work on my dissertation there. In particular I wish to thank Gordon Inverarity whose guidence was invaluable.

I would like to acknowledge the NERC for funding this work, without which I would have been unable to complete the course. I would also like to thank the Met Oce for funding my time in Exeter, while working on this project.

I would like to say thanks to my parents for their support during my time at university and to my girlfriend Lucy for keeping me sane over the year.

Abstract

We aim to investigate the convergence of the difference between the forecast and analysis fields during the incremental 4D-Var method used in numerical weather prediction at the UK Met Office. We transform the analysis increment into Fourier space in order to look at how the power at each mode number, which is inversely proportional to the wavelength, evolves through the 4D-Var minimisation iterations. We investigate appropriate ways to display results, including power spectra and periodograms and we look at how to use windowing and binning in order to improve the accuracy of the periodogram estimation. Then we compare the results obtained to show that they correspond to those of previous work using the simple barotropic vorticity equation model. We conclude that the largest and smallest scales converge the fastest, after only 10 iterations, while the intermediate scales appeared still to be converging after 30 iterations of the inner loop.

Contents

1 Introduction 3

2 Useful Tools 7

2.1 Discrete Fourier Transforms 7

2.2 The Power Spectrum 11

2.3 The Periodogram 15

2.4 Data Windowing 20

3 The Model and Methodology 24

3.1 Full 4D-Var Assimilation 24

3.2 Incremental 4D-Var Assimilation 26

4 Previous Work on Forecast Error 29

4.1 Four-dimensional Data Assimilation with a Wide Range of Scales 29

4.2 A Validation of the Incremental Formulation of 4D Variational Data Assimilation in a Nonlinear Barotropic Flow 33

5.3 The binned 1D periodogram 46

6 Conclusion 51

Chapter 1

Introduction

Four-dimensional variational data assimilation (4D-Var) is a technique used to incorporate observational data into a dynamical model over a given period of time in order to predict all the future and current states of a system [11]. The method aims to minimise the square errors between the observational system data and the model predictions. This can be thought of as minimising a *cost function*, which is a measure of the distance between the model state and the observations, and the model state and the initial state estimate, [10].

The minimisation of the cost function comes at great computational expense. In order to reduce this expense the full minimisation can be approximated by a series of minimisations of approximate convex quadratic cost functions. This approximation to the full 4D-Var method is known as the *incremental* 4D-Var method, [10].

We begin the incremental 4D-Var process with a initial estimate of the field states, known as a *background* field and run the nonlinear model to calculate the field states at each time step. We then calculate the value of

an incremental field, around the background, which minimises an approximately quadratic *incremental cost function*. The minimisation process uses an iterative procedure such as quasi-Neuadratic

where $\underline{x}$ and $\underline{b}$ are vectors and A is a matrix. This will mean that by describing the gradient of the incremental cost function in this form, we can potentially solve the problem using the multigrid method.

The solution to the multigrid problem is found by eliminating the different components of the error on different resolution grids. High frequency components of the error are eliminated on a high resolution discretisation of the problem, whilst low frequency components are resolved on a coarser discretisation. Solving a problem using multigrid method is done in stages, [13]:

- Presmoothing stage. The high frequency error (relative to the grid being used) are effectively eliminated.
- Restriction stage. The remaining error is transferred to a coarser computational grid, on which we continue to solve the problem.
- Prolongation stage. The correction is then interpolated back onto the finer grid.
- Post-smoothing stage. Any remaining high frequency components are again eliminated.

One iteration by this method, in general, involves descending through a hierarchy of computational grids from finest to coarsest and back up again - a process known as a V-cycle. It is possible to link more than one V-cycle together to introduce further accuracy in the solution.

In order for the multigrid method to be used on the inner loop, we would require the small scale components of the forecast error to be smoothed out quickly so the inner loop can continue on the lower resolution grid. If small

Chapter 2

Useful Tools

In this chapter we deal with some tools, used in this project to analyse forecast errors, associated with data manipulation and massaging, such as data windowing and detrending. We also describe ways in which we are able to show the periodicity of a set of sample data by transforming it into Fourier space.

2.1 Discrete Fourier Transforms

Any function defined over a finite domain or any periodic function, $f(x$

where the coe□cients a_0, a_n and b_n are related to the given function by definite integrals,

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(t)\cos(nt)dt,$$
$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(t)\sin(nt)dt, \quad n = 0, 1, 2, \ldots.$$

This is subject to the existence of these integrals, which is true if $f(t)$ is piecewise continuous, [1].

If we now express $\cos nx$ and $\sin nx$, from (2.1), in exponential form, we may rewrite (2.1) as, [1]

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}, \tag{2.2}$$

where,

$$c_n = \frac{1}{2}($$

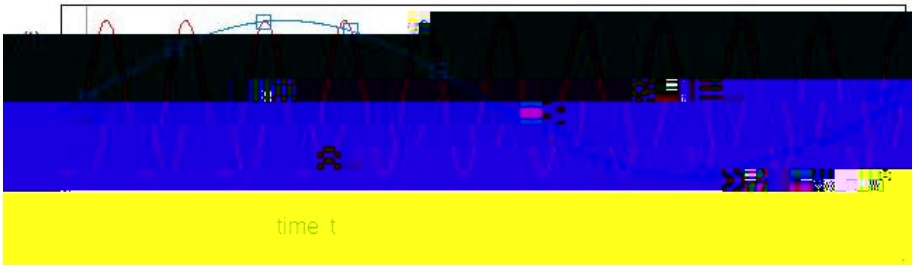


Figure 2.1: Diagram to illustrate the problems with the Nyquist frequency, from [6]

with the inverse discrete Fourier transform defined as,

$$f$$

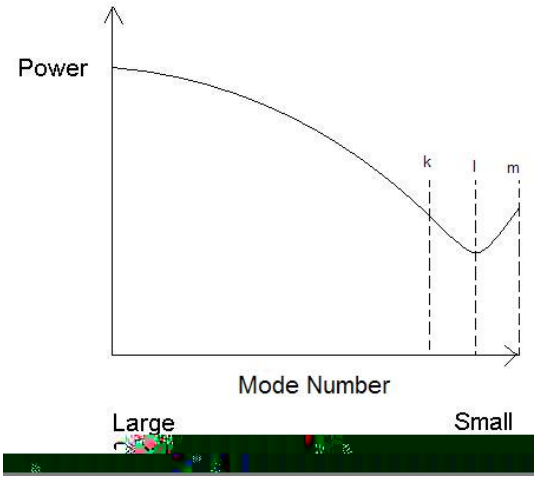


Figure 2.2: Diagram showing the folding of power between mode numbers k and l to between mode numbers l and m, adapted from [6]

that $|f| > \frac{1}{2\Delta}$, it would be impossible to distinguish between the two data sets using just the sampled points.

If a high frequency physical signal is not discretised frequently enough in order to resolve it correctly, the frequency function *folds* or *overlaps* on itself, in a process called *aliasing*, thereby forming an incorrect Fourier transform.

Figure 2.2 shows a simple example of what the aliasing effect, i.e. sampling a data set above the Nyquist frequency, would have on the power spectrum of a data set. As can be seen, there is a smooth decline in the power from large scales (low mode number). However as we reach the smallest scales, the power 'kicks' back up again. This is due to the aliasing effect on the data and is not accurate. In this example we can see that the power in between mode numbers k and l are folded over into the mode number range l to m.

For large values of N, the discrete Fourier transform (DFT) becomes computationally expensive, with the number of operations required to perform

the calculation approximately N^2. In order to reduce the expense of the DFT, we may use a method called the *Fast Fourier Transform* (FFT) which reduces the number of operations to $O(N log_2(N))$, [1].

When transforming a 2D set of $N \times M$ data into Fourier space we must use the corresponding 2D discrete Fourier transform, which is described from [6] as

$$\tilde{f}_{k,l} = \frac{1}{NM} \sum_{h=0}^{N-1} \sum_{j=0}^{M-1} f_{h,j} e^{-2\pi i(hk/N + jl/M)}, \tag{2.4}$$

for the discrete function, $f_{h,j}$, $h, k = 0, \ldots, N-1$ and $j, l = 0, \ldots, M-1$. Here k and l are defined as the mode numbers.

2.2 The Power Spectrum

The complex modulus of the DFT coefficients can be taken and then appropriately scaled in order to generate a plot which can be used to approximate the 'power spectrum' which gives the power associated with each wave mode present in the discrete set. These mode numbers correspond to different physical scales present in the discrete sample data, [5].

From Parseval's theorem, [1], we can show that the relation between the discrete Fourier transform and its inverse is as follows,

$$\frac{1}{N} \sum_{j=0}^{N-1} |f_j|^2 = \sum_{k=0}^{N-1} |\tilde{f}_k|^2, \tag{2.5}$$

which in 2D generalises to,

$$\frac{1}{NM} \sum_{h=0}^{N-1} \sum_{l=0}^{M-1} |f_{h,l}|^2 = \sum_{j=0}^{N-1} \sum_{k=0}^{M-1} |\tilde{f}_{j,k}|^2. \tag{2.6}$$

If we consider a function in one dimension, such that $f_j = \sin(\frac{2a\pi j}{N})$, with $a = 1, 2, 5, 7$ and $N = 32$, then calculate the power in this function from

(2.3), and using the fact that $|e^{-\frac{2\pi i j k}{N}}|^2 = 1$,

$$\sum_{j=0}^{N-1} |sin\left(\frac{2a\pi j}{N}\right)|^2 = \frac{N}{2}$$

Then from Parsevals theorem, (2.5),

$$\sum_{k=0}^{N-1} |\tilde{f}_k|^2 = \left(\frac{N}{2}\right)/N = 0.5$$

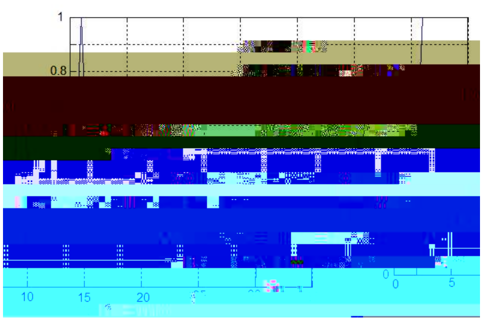
Therefore, in order for the power spectrum to have an amplitude of 1, rather than 0.5, thereby making the power spectra easily comparable, we introduce a scaling factor of 4. Figure 2.3 shows the power spectrum of the function $f_j = sin(\frac{2a\pi j}{N})$ with $a = 1, 2, 5, 7$ and $N = 32$.

The power spectrum of the sine function produces two distinct peaks of equal amplitude. The location and periodicity of these peaks can be found in the following way, [5].

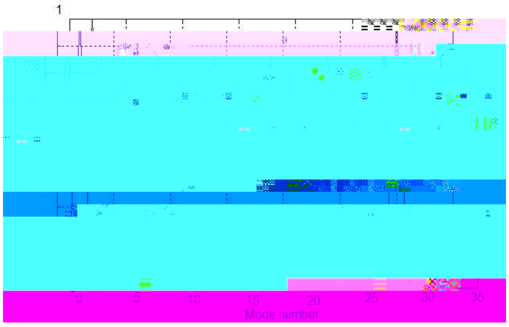
When the DFT is applied to a real sequence, f_j with $j = 0, \ldots, N-1$, the resultant will be, in general, a sequence of N complex numbers such that $\tilde{f}_k$ and $\tilde{f}_{N-k}$ are related, from [5], by

$$\tilde{f}_{N-k} = \tilde{f}_k^*$$

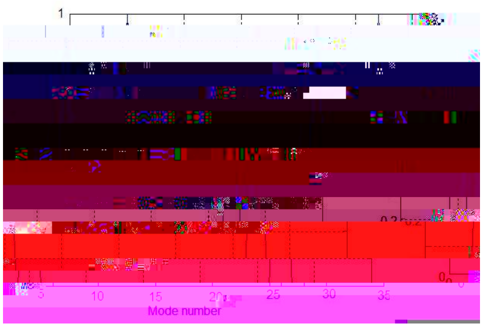
for $k = 0, \ldots, N-1$, where * denotes the complex conjugate. This relation causes a periodic function to display the two distinct transform peaks, as seen in Figure 2.3. It also explains why we are seeing results in mode numbers corresponding to frequencies higher than the Nyquist frequency. The highest accessible mode number, corresponding to the Nyquist frequency, is $\frac{N}{2} = 16$. The power series does not supply us with any new information past



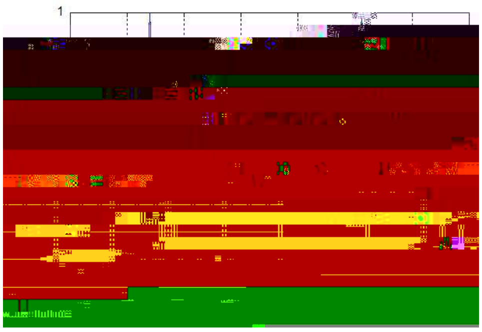
(a) Power spectrum of $sin(\frac{2\pi j}{N})$



(b) Power spectrum of $sin(\frac{4\pi j}{N})$



(c) Power spectrum of $sin(\frac{10\pi j}{N})$

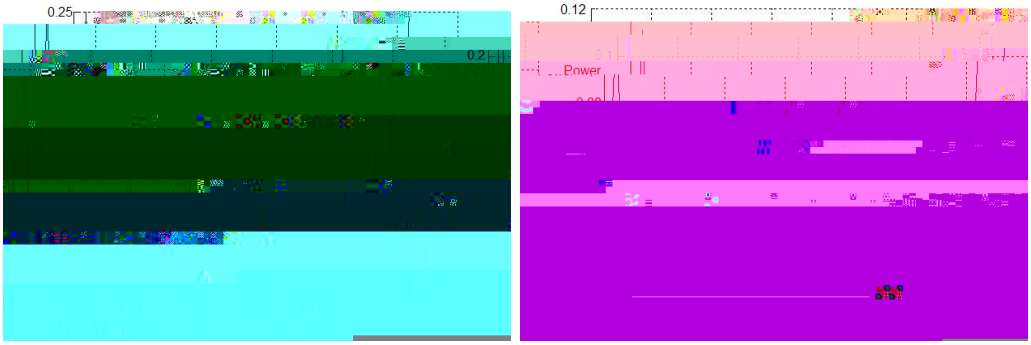


(d) Power spectrum of $sin(\frac{14\pi j}{N})$

Fomguterus m)

$\sin(\frac{2\pi a(N-j)}{N})$, so it can be seen that they appear at mode numbers a and $N - a$.

It is possible to use the power spectrum as an indicator of the wave pattern associated with a function. If we study the power spectrum of a function containing multiple waves, for example, $sin(\frac{2\pi j}{N}) + sin(\frac{6\pi j}{N})$ or $sin(\frac{4\pi j}{N}) + sin(\frac{14\pi j}{N})$ as shown in Figure 2.4, we can see that there is a peak in the power spectrum associated with each corresponding wavelength. Due to the Nyquist frequency, this will be the case for all mode numbers, $k < \frac{N}{2}$, since this is the largest mode number (corresponding to the smallest wavelength) repre-



(a) Power spectrum of $\frac{1}{2}sin(\frac{2\pi j}{N})$

$$\Delta \sum_{j=0}^{N} |f_j|^2 \approx \int_0^T |f(t)|^2 dt \equiv \text{time-integral squared amplitude,}$$

where T is the total sampling time, i.e. $T = N\Delta$, with Δ being the interval and N the total number of points sampled. The triangular rule is then applied in order to remove the final sampled point. In other words, we use $N + 1$ points to discretise the integral but only N points to define the periodogram.

There are even more possible ways in which to calculate the power spectrum, all of which have the characteristic of never integrating the spectrum outside the Nyquist interval, i.e. the sampling rate is below the Nyquist frequency. According to sampling theory, any power lying outside of this interval will be folded back into the region below the Nyquist frequency, [12].

One method used in estimating the power spectrum of a sampled function, is called a *periodogram*. If we apply an (N)-point fast Fourier transform to a discrete function, f_j to calculate its discrete Fourier transform,

$$\tfrac{1}{N} F_k = \sum_{j=0}^{N-1} f_j e^{-2\pi i j k/N} \quad k = 1, \ldots, N-1,$$

the periodogram used to estimate the power series of this discrete Fourier transform is defined at $\frac{N}{2} + 1$ frequencies as,

$$P(f_k) = [|F_k|^2 + |F_{N-k}|^2] \quad k = 1, 2, \ldots, (\tfrac{N}{2} - 1)$$

$$P(f_c) = P(f_{N/2}) = |F_{N/2}|^2$$

where f_c is the Nyquist frequency, f_k is defined for zero and positive frequencies and $P(f)$ denotes the power spectrum calculated. From (2.5ok

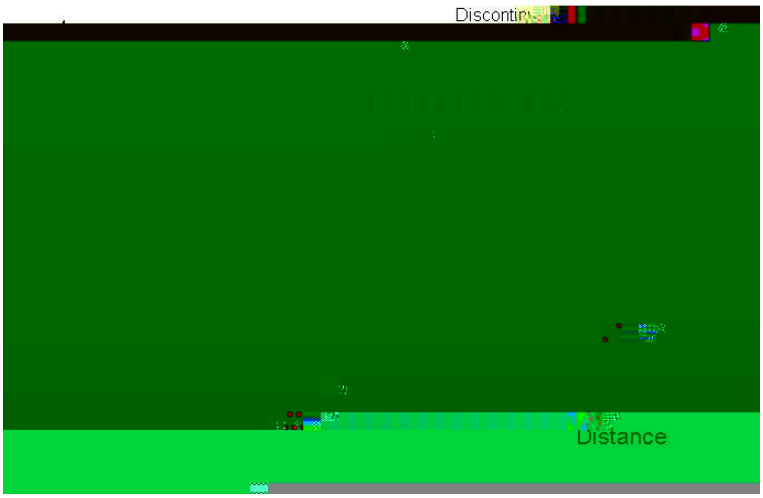
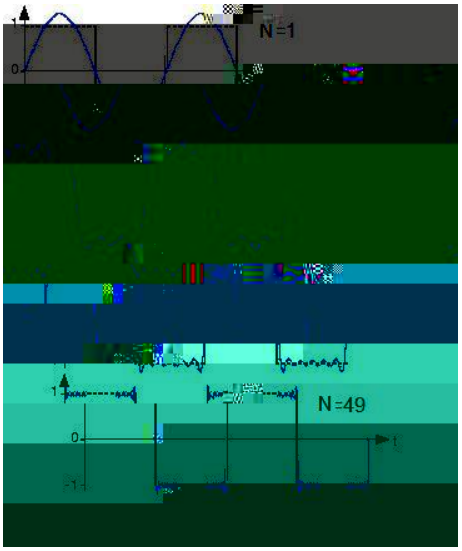


Figure 2.6: Diagram showing a discontinuity forming due to the FFT


N=1
1
0
N=49
1
0
-1
t

nomena, [2].

The Gibbs' phenomena is illustrated in Figure 2.7 by showing how a Fourier series approximates a square wave discontinuity. We can see that the fewer sampling points used by the Fourier series, the worse the approximation is. This same effect occurs on the discontinuity shown in Figure 2.6, and leads to aliasing effects shown in Figure 2.2.

One possible method and indeed the method used in this project, for reducing the Gibbs' phenomena from a set of non-periodic sample data is known as *detrending*. The most commonly used technique for detrending data is to fit a first order polynomial, or at least a low order one, to the data, then subtract it from the sample. This has the effect of reducing the distance in height between the two end points, so that the discontinuity created when the Fourier transform creates a false periodicity is reduced. The effect of this is to reduce the aliased power in the high frequency range of the periodogram.

Detrending also effects the low frequency end of the spectrum, which can no longer be regarded as reliable. This is not a major problem as the trends cannot be described by the lowest frequency therefore removing them does not change the overall result.

When forming the power spectrum of a two dimensional sample of data, we need to use a plane surface to detrend the data, [2].

In Figure 2.8 we show the detrending method used to reduce Gibbs' phenomena. As shown in the diagram, we begin with a set of data then remove its trend. In doing so we re-align the data set about zero and when we now use this set as in Figure 2.6, to transform it into Fourier space, we reduce the likelihood of creating a large discontinuity between the data and its copy.

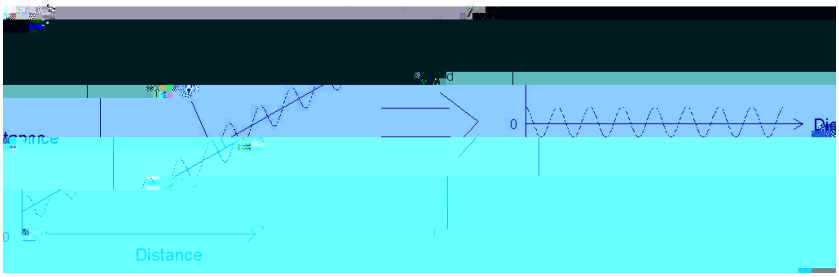


Figure 2.8: Diagram showing detrending of a sample data set

We recall that in order to use the FFT on a discrete set of data, the FFT assumes periodicity in the data and so copies the data and adds it to the end of the domain. In order to reduce the sample domain to the original size, the data set is multiplied by a rectangular domain, which is equal to 1 over the domain of the data and zero everywhere else, as shown in Figure 2.9.

These domain truncations can cause unacceptable approximations to the Fourier transform, [6]. In the case where we use a discrete set of non-periodic data, we must employ data windowing in order to produce a better estimate of the Fourier transform. This is discussed in the next section.

2.4 Data Windowing

As previously discussed in section 2.3, while looking at non-periodic discrete data, we may employ windowing techniques in order to reduce the undesired e□ects of domain truncation by the rectangular window.

The basic idea behind data windowing is to massage the sample data in order to approximate the power spectrum by the periodogram more closely. The multiplication of a discrete data set by a rectangular window function,

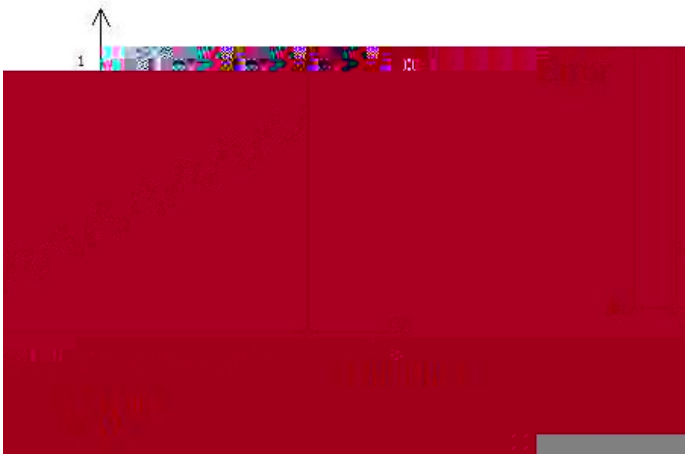
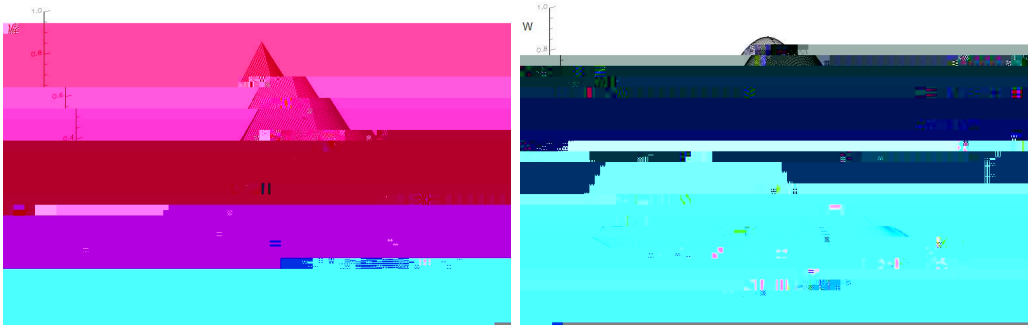


Figure 2.9: Diagram showing the square window multiplied by the sample data

creates discontinuities at the edges of the domain. When these discontinuities are resolved by the FFT, Gibbs' phenomena occur, causing false high frequency oscillations in the results. This e□ect is called *leakage*, [12].

Figure 2.9 shows a rectangular spatial window being applied to a set of data. As can be seen, at the edges of the domain, the data suddenly changes



(a) The 2D Bartlett Window (b) The 2D Hann Window

Figure 2.10: 2D windowing function

The technique of data windowing looks to modify the function, w_j, so that discontinuities no longer form at the edges of the spatial domain.

There are many variations of the window function, all of which involve starting at zero at the domain edges and gradually increasing to a maximum value of 1 at the middle of the domain. Two of the more popular window

size ($M \times N$) with N even and M odd:

$$w_{i,j} = \left(1 - \left(\frac{2i}{N} - 1\right)^2\right)\left(1 - \left(\frac{2j}{(M-1)} - 1\right)^2\right) \equiv \text{"2D Bartlett Window"}, \tag{2.8}$$

$$w_{i,j} = \left(1 - \cos(\pi i/N)^2\right)\left(1 - \cos(\pi j/M)^2\right) \equiv \text{"2D Hann Window"}. \tag{2.9}$$

These window functions are displayed in Figure 2.10. As can be seen, the window function starts at zero and gradually increases to a maximum value of 1 and then returns back to zero.

The corresponding normalising factor for the 2D periodogram is now

$$\frac{NM}{\sum_{i=0}^{N}\sum_{j=0}^{M}(w_{i,j})^2}.$$

The next Chapter looks at the full and incremental 4D-Var assimilation methods in detail. It is the incremental version which we shall be using in this project to study the evolution of the scales present in the forecast error.

Chapter 3

The Model and Methodology

In Chapter 3 we discuss the full 4D-Var data assimilation method and the incremental 4D-Var assimilation method. We shall use the incremental method in this project as it is this method which is used by the UK Met O□ce in numerical weather prediction.

We look at the full 4D-Var method because one of the few pieces of work on the evolution of forecast errors was done using this method on a simplified model. We shall be looking at the incremental method on the UK Met O□ce numerical weather prediction model, which has not previously been studied.

3.1 Full 4D-Var Assimilation

The data assimilation process aims to incorporate observed measurements into a dynamical system model in order to produce accurate estimates of all the state variables over a period of time, known as the assimilation window, [11]. We aim to minimise the square errors between the model predictions and the observed system states. In other words 4D-Var data assimilation

aims to find the model state, x_0, which minimises the cost function, in the time window $[t_0, t_n]$

$$J[x_0] = \frac{1}{2}(x_0 - x^b)^T B_0^{-1}(x_0 - x^b) + \frac{1}{2}\sum_{j=0}^{n}(H_j[x_j] - y_j^0)^T R_j^{-1}(H_j[x_j] - y_j^0), \tag{3.1}$$

subject to the discrete nonlinear model

$$x_j = S(t_j, t_0, x_0).$$

Here x^b is a background field, y_j^0 are the observations, H_j is the observation operator used to map the fields from model space onto observational space and $S(t_j, t_0, x_0)$ is the solution operator of the nonlinear model. The background and observational errors are stored in the covariance matrices, B_0 and R_j respectively, [10].

The idea behind the full 4D-Var assimilation method is to solve the full nonlinear cost function using an iterative process where k is the iteration number, [3],

1. Set $x_0^{(k)} = x^b$, where $k = 0$ and x^b is the background field.
2. Run the nonlinear model and calculate the cost function, $J(x_0^{(k)})$.
3. Find the gradient of the cost function, $J(x_0^{(k)}$

steps had been taken, or the solution had converged sufficiently.

Previous work by Tanguay et al., [15], has looked at how the error between the analysis field created by the assimilation process and the initial background field (known as the forecast error) evolves over one iteration of this process, and is discussed in chapter 4.

3.2 Incremental 4D-Var Assimilation

As in the full 4D-Var assimilation method, we define the incremental method as finding the model state, x_0, which minimses the cost function (3.1). × This minimisation procedure can carry extremely high computational cost due to the nonlinear nature of the observation operator and numerical model as before, which cause the cost function, J, to be a nonlinear least squares problem, [10].

The reduction in computational cost is achieved by approximating the full problem by a series of minimisations of approximate quadratic cost functions. We can view the 4D-Var process as the following iterative procedure, [10], where k is the iteration number:

1. We begin by defining a guess field, $x_0^{(k)}$. For $k = 0$, the first iteration, we set the background field, x^b, equal to the guess field, $x_0^{(0)}$(

4. An incremental field is then defined as $\delta \mathbf{x}_0^{(k)} = \mathbf{x}_0^{(k+1)} - \mathbf{x}_0^{(k)}$.

5. The value of $\delta \mathbf{x}_0^{(k)}$ is found, which minimises the cost function,

$$\begin{aligned} \tilde{J}^{(k)}[\delta \mathbf{x}_0^{(k)}] &= \tfrac{1}{2}(\delta \mathbf{x}_0^{(k)} - [\mathbf{x}^b - \mathbf{x}_0^{(k)}])^T \mathbf{B}_0^{-1}(\delta \mathbf{x}_0^{(k)} - [\mathbf{x}^b - \mathbf{x}_0^{(k)}]) \\ &\quad + \tfrac{1}{2} \sum_{j=0}^{n} (\mathbf{H}_j \delta \mathbf{x}_j^{(k)} - \mathbf{d}_j^{(k)})^T \mathbf{R}_j^{-1} (\mathbf{H}_j \delta \mathbf{x}_j^{(k)} - \mathbf{d}_j^{(k)}). \end{aligned} \tag{3.2}$$

   subject to

$$\delta \mathbf{x}_j^{(k)} = \tilde{\mathbf{L}}(t_j, t_0, \mathbf{x}^{(k)}) \delta \mathbf{x}_0^{(k)}$$

   where $\mathbf{H}_j$ is the linearisation of the observational operator H_j around the state $\mathbf{x}_j^{(k)}$ and $\tilde{\mathbf{L}}(t_j, t_0, \mathbf{x}^{(k)})$ is the solution operator of the linear model linearised around the nonlinear model trajectory.

6. We update the guess field using,

$$\mathbf{x}_0^{(k+1)} = \mathbf{x}_0^{(k)} + \delta \mathbf{x}_0^{(k)}.$$

   .

7. The procedure is repeated until a convergence threshold has been reached or after a set number of iterations has passed. We can now define an analysis field as the updated guess field, i.e. $\mathbf{x}^a = \mathbf{x}^{(M)}$, where M is the number of iterations performed.

Each iteration of this procedure is known as an *outer loop* and with each outer loop, the minimisation of the cost function (3.2) must be solved in another iterative procedure known as an *inner loop*, [10]. The solving of each inner loop may be thought of as a convex quadratic minimisation problem and as a consequence can be solved using methods such as the *quasi-Newton method* or the *conjugate gradient method*, [10].

It is possible to think of the incremental field, after it has converged sufficiently, as the difference between the background guess and the converged analysis field, i.e. the error between our initial guess and the analysis we have converged to,

$$\sum_{k=0}^{M} \delta x^{(k)} = x^a - x^b. \tag{3.3}$$

where M is the total number of iterations of the inner loop.

There has been little work on how the error, between the analysis and background field states, evolves as the number of iterations in the inner loop increase. Moreover, this work has only been conducted using simple models, such as the *barotropic vorticity equation*. In section 4 we discuss past work on this analysis of forecast errors.

Chapter 4

Previous Work on Forecast Error

As part of this chapter we discuss previous work on the evolution of the error between the analysis field and the initial background field through the full and incremental 4D-Var assimilation process.

All previous work has only been done using the simple barotropic vorticity equation model, whereas we shall be studying the nonlinear numerical weather prediction model used at the UK Met O□ce, which has never been used to investigated this question.

4.1 Four-dimensional Data Assimilation with a Wide Range of Scales

The paper written by Tanguay, Bartello and Gauthier entitled 'Four-dimensional data assimilation with a wide range of scales' looks at

whether it is possible for the adjoint method, used in the full 4D-Var process, to improve initial conditions in data-sparse regions of observational field data. They examine the method's ability to 'fill in' small scale detail determined dynamically from large scale data, [15].

The numerical model chosen for this piece of work was the barotropic β-plane flow, due to its simple geophysical settings and the wide range of scales involved. The barotropic vorticity equation on the β-plane approximation is defined as

$$\frac{\partial \zeta}{\partial t} + \mathrm{J}(\Psi, \zeta) + \beta v = f - D(\zeta), \tag{4.1}$$

where $\Psi = -U_0 y + \psi$, U_0 is the mean zonal wind (with dimensionless units) taken to be 0.3, y is the distance in the north-south direction, ψ is the stream function, $\zeta = \ ^2\psi$, (the relative vorticity), $u = -\partial\psi/\partial y$, (the wind field in the east-west direction), $v = \partial\psi/\partial x$, (the wind field in the north-south direction), J is the Jacobian, f is a forcing term, β is the coriolis parameter and D is a linear dissipation operator, [15].

The tangent linear model was introduced as an intermediate step in order to derive the adjoint model of the problem. This was done by considering a reference solution to the nonlinear equation, ζ and a perturbation about that solution, $\delta\zeta$. Then writing (4.1) for the total field $\zeta + \delta\zeta$ and then subtracting the equation for ζ to get, [15],

$$\left(\frac{\partial}{\partial t} + U_0 \frac{\partial}{\partial x} \right) \delta\zeta + \mathrm{J}(\psi, \delta\zeta) + \mathrm{J}(\delta\psi, \zeta) + \mathrm{J}(\delta\psi, \delta\zeta) + \beta\delta v = -D(\delta\zeta). \tag{4.2}$$

This becomes the *tangent linear model* by neglecting the term, $\mathrm{J}(\delta\psi, \delta\zeta)$. In one of the experiments undertaken by Tanguay et al. (1994) while investigating the adjoint method's ability to fill in small scale detail from larger scales, they looked at the convergence of the forecast error using the full 4D-

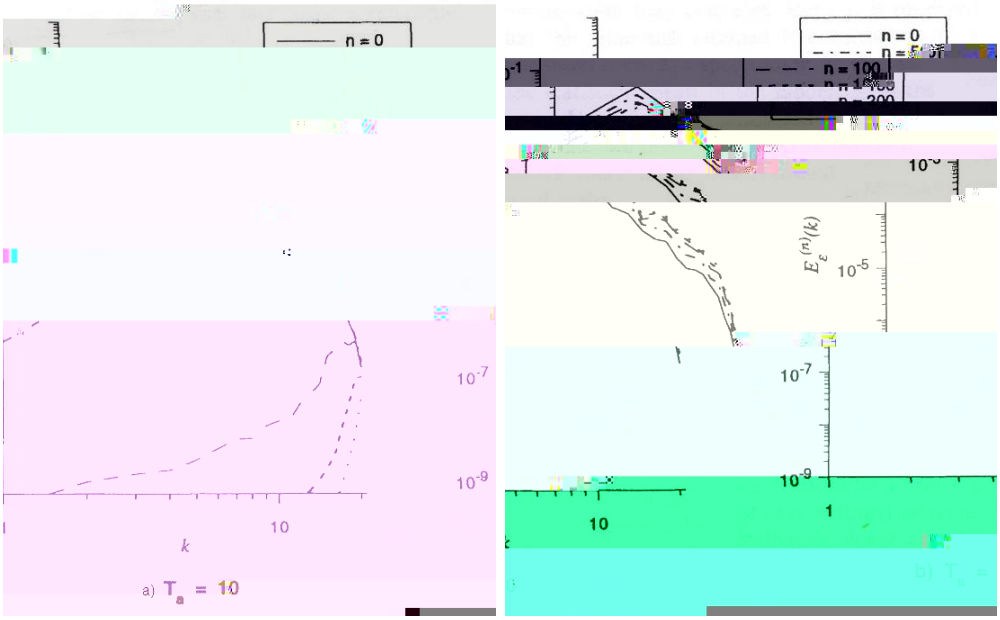


Figure 4.1: Evolution of forecast error through the adjoint method of the full 4D-Var process, using assimilation window of size a) T_a = 10 and b) T_a = 40

Var method applied to (4.2). The error field was defined as $^{(n)} = \zeta^{(t)} - \zeta^{(n)}$, where $\zeta^{(t)}$ is the true analysis field of relative vorticity created from by an integration of the nonlinear model, $\zeta^{(0)}$ is the initial background estimate field and $\zeta^{(n)}$ is the relative vorticity at the nth iteration of the adjoint method, [15]. Using one iteration of the full 4D-Var outer loop, they observed how the difference between the true analysis field and the updated background field, converge through the iterative adjoint method, described in section 3.1.

Figure 4.1 shows the results taken from [15], showing the evolution of the forecast error over the process of the adjoint method, using one outer

The assimilation window is defined as the period of time over which the assimilation process takes place. Figure 4.1(a) uses an assimilation window of size T_a = 10 model time units (≈ 3 days). Tanguay et al. concluded that the results in figure 4.1(a) showed that right from the start of the assimilation process, the adjoint method converged and the largest scales were adjusted to fit the observations. As the iterations proceeded, the minimisation focused its attention on increasingly smaller scales until, after about two hundred iterations, virtually all of the error was in the smallest scales,[15]. It can be seen that after a very short time, almost all of the error disappears. Tanguay et al., [15] concluded that since the errors at t = 0 (the initial time) are confined to the smallest scales, the initial small-scale field is only weakly dynamically coupled to the rest of the observational data, [15].

Figure 4.1(b) uses an assimilation window of size, T_a = 40 model time units (≈ 13 days). This was done in order to investigate the convergence when the assimilation period is long compared to the nonlinear timescale of the model. We see, in figure 4.1(b), that the convergence of the larger scales is very slow and more importantly the method appears to be diverging in the smaller scales. This shows that the consequence of having an assimilation period that exceeds beyond the validity time scale of the tangent linear model is that the solution at the end of the time period has essentially lost its memory of the initial conditions, which constitutes the limit of the method, [15].

The reason given in the paper for the divergence at smaller scales is the nonlinearity of the problem causing the cost function, instead of being quadratic as assumed, forming a second minimum. Therefore, an incorrect value of the

model state, which minimises the cost function, may be chosen.

We will now discuss a paper which uses the same model but uses the incremental 4D-Var method instead of the full method.

4.2 A Validation of the Incremental Formulation of 4D Variational Data Assimilation in a Nonlinear Barotropic Flow

This paper, written by Laroche and Gauthier, [9], investigated the ability of the incremental approach to 4D-Var to reduce computational cost, in order to meet operational limitations of the full 4D-Var method.

As with the paper written by Tanguay et al. ,[15], the model used in this paper was the physically simple 2D barotropic vorticity equation on a β-plane approximation (4.1).

The results in [9] which are most pertinent to our work are those obtained without updating the outer loop trajectory. These examine the convergence properties of the incremental 4D-Var method, without updating the reference directory, i.e. when only one outer loop iteration is performed. As part of this experiment they compared the convergence of the incremental method with the convergence of a full low-resolution method (which acted as the control run), over a number of iterations of the inner loop at the beginning of the assimilation process.

In order to observe the convergence of these methods, Laroche and Gauthier studied the energy spectrum of an error field defined as the difference between the vorticity obtained after a given number of iterations and the

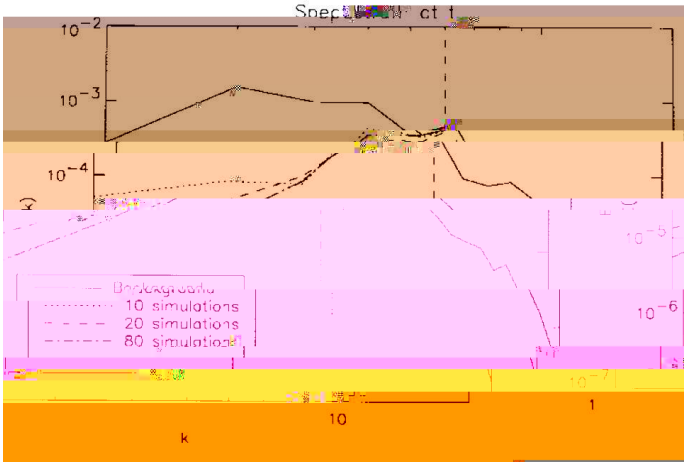


Figure 4.2: Error spectrum at time, t_0 for the incremental approach, [9], the vertical line indicates the mode number k_L

control run created from an integration of the nonlinear model.

In these experiments the models were not used to assimilate data with mode numbers greater than $k_L = 5$. This is because they believed this to be the best choice to enhance the interactions between the resolved and unresolved scales, [9]. Figure 4.2 shows how the error spectrum evolves and converges throughout the first eighty iterations of the inner loop for the incremental approach to the 4D-Var method. As with the work presented in [15], the error spectrum in the larger scales reduces right from the start, although not to the same extent as shown in figure 4.1(a). Then the smaller scales (ignoring mode numbers larger than k_L) begin to reduce in the first 20 iterations, but then after 80 iterations of the inner loop, the power in the smaller scales begins to increase as the solution diverges.

This same divergence phenomenon appears in [15] where the assimilation window is larger than the nonlinear validity timescale of the model.

Figure 4.3 shows the error spectrum for the first eighty iterations of the

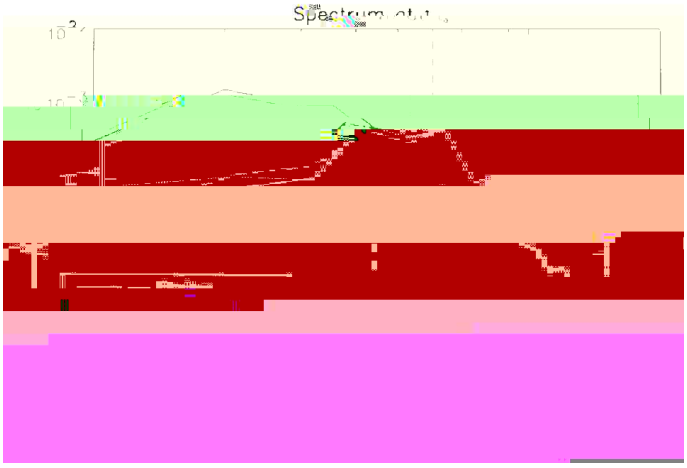


Figure 4.3: Error spectrum at time, t_0, for the full low-resolution approach, [9]

inner loop, using the full low-resolution approach to 4D-Var assimilation.

Again, the larger scales reduce in power quickly, after only a few iterations and the smaller scales take more time to reduce in power. As with Figure 4.2 we see the solution diverging in the small scales, with it being even more pronounced when using the full low-resolution approach. After eighty iterations of the inner loop, the divergence is such that there is more power located in the smaller scales than there was in the initial background field.

In this project we intend to move forward from the pieces of work discussed in this section and investigate the evolution of the forecast error using the incremental 4D-Var method applied to the nonlinear numerical weather prediction model used at the UK Met O□ce. This project has never been done before and we therefore intend to compare the results of this investigation with those obtained by the pieces of work in this chapter.

The next chapter describes the numerical experiments we undertook to

obtain our results.

Chapter 5

Numerical Experiments

In this chapter we run the incremental 4D-Var process on the UK Met O□ce numerical weather prediction model outputting the forecast error every few iterations of the inner loop, in order to investigate the evolution of the scales present in the incremental data.

We describe, as we come to them, the steps taken and the thought processes involved in finding the best combination of accuracy and comparability between results to display the forecast error at each iteration.

5.1 2D Periodogram

We begin by running the incremental 4D-Var method, over one outer loop, (as described in section 3.2) on the Met O□ce numerical weather prediction model. The full operational observation data was used and we examined the incremental data on the full global scale (432x325 numerical grid points), at the surface level.

In order to investigate the best way to display the results, we began by

Figure 5.1: 2D periodogram of the incremental east-west component of the wind field, after 40 inner loop iterations.

analysing the incremental field data of the east-west component of the wind field at the surface, after the full 40 iterations of the inner loop.

This field data was then transformed into Fourier space using the 2D fast Fourier transform, (2.4) and using the 2D periodogram, (2.7), we were able to produce a measure of the power based on the DFT.

Figure 5.1 shows the 2D periodogram of the incremental east-west component of the wind field. The x and y axis display the logarithm of base 10 of the mode numbers in those directions. We display the power in terms of a contour plot, with the brighter colours representing higher powers.

It can be seen that in the bottom left corner we have an area of white (therefore high power), corresponding to a power of $10^{2.19}$

scales are inversely proportional to the mode number and this area has the smallest mode numbers, it corresponds to the largest physical scales. As we

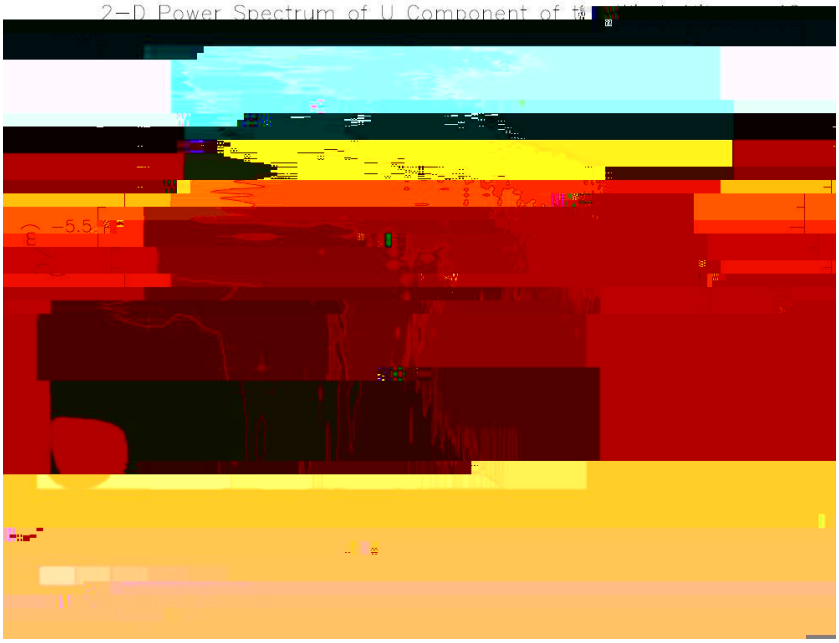


Figure 5.2: 2D periodogram of the incremental u-component of the wind field, after 40 inner loop iterations, using the detrending technique.

to look for a best fit plane to the original data set in physical space. This was then subtracted from the data and the FFT was taken again to transform the results into Fourier space and the 2D periodogram was formed. The detrended 2D periodogram is shown in figure 5.2.

The aliasing problem seems to have been dramatically reduced by the introduction of the detrending method. As can be seen, the highest power is located in the largest scales and as the mode number increases, corresponding to a decrease in the physical scales, the power also reduces. This was expected and is demonstrated in results by Tanguay et al. (Figure 4.1), [15], and by Laroche et al. (Figures 4.2 and 4.3), [9].

The same power scales were used in the production of both 2D peri-

odograms in order to improve comparability of results. It can be seen that after the detrending technique was used on the incremental data, the power in the smallest scales was increased from approximately $10^{-12.23}$ before the detrending to approximately $10^{-9.34}$ after detrending.

To investigate this alteration in the power at small scales we decided to plot the absolute value squared of the Fourier coe□cients of the incremental data, against its corresponding mode number, on a full 1D power series scatter plot, described in the next section.

5.2 The full 1-D power series scatter plot

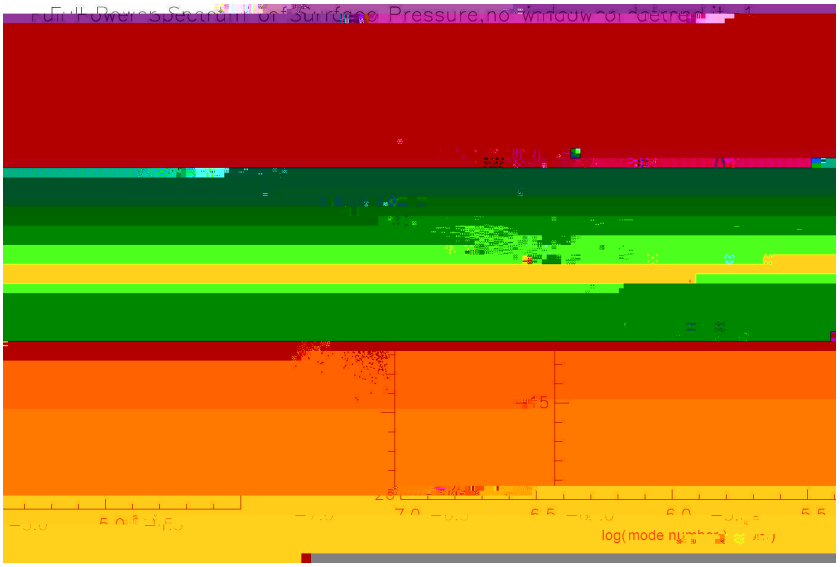


Figure 5.3: Full 1D power series scatter plot, after one inner loop iteration.

Figure 5.4: Full 1D power series scatter plot, after one inner loop iteration and using detrending.

each mode number for the surface pressure increment after one inner loop iteration, without the use of the detrending technique.

The idea of this plot was to help us to understand the increase in power in the small scale data when the detrending technique was applied to the 2D periodogram.

Before the detrending technique was applied the distribution of power appears to be as expected, with more points in the small scale region indicating a lot of noise in the original data. There also appears to be some sort of

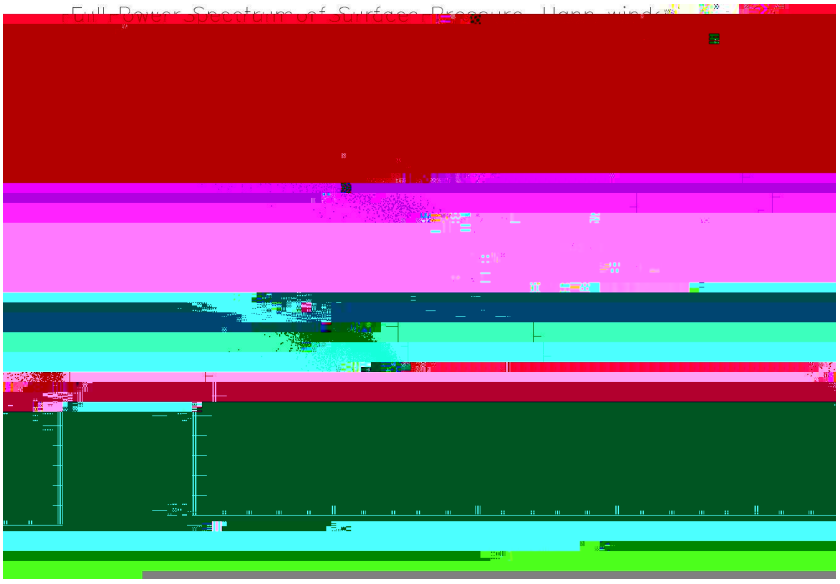
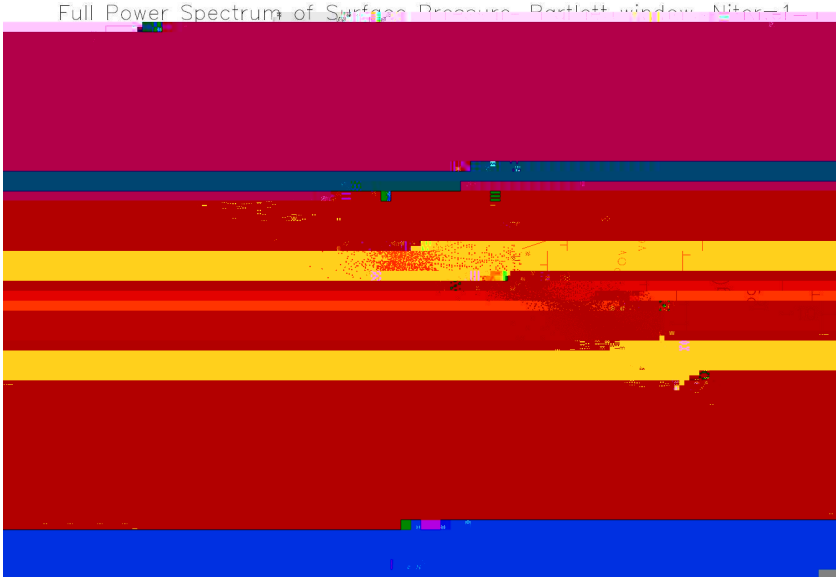


Figure 5.5: Full 1D power series scatter plot, after 1 inner loop iteration, using detrending and the Hann window function.

(figures 2.10a) and b) respectively). We now describe the application of the window functions and the e□ect they have on our results.

The incremental data was detrended as previously, but now, before the transformation into Fourier space, we multiply the data by a windowing function. From section 2.4 we know that the multiplication of a data set by a rectangular window causes discontinuities to form at the edges of the domain. The hope is that, by applying the smoother window functions, we will remove the discontinuity created by the rectangular window and therefore remove the streaks in the scatter plot.

Figure 5.5 shows the same results as Figure 5.4 but with the multiplication of the detrended incremental data by the 2D Hann window and Figure 5.6 shows the same result, using the Bartlett window rather than the Hann


Full Power Spectrum of Surface Pressure, Bartlett window, Nitop=1

This appears to correspond with the results by Tanguay et al., [15], and Laroche et al., [9], who also noted that the larger scales converged first. However, they showed that the intermediate scales then converged followed by the smallest scales. Figures 5.7 and 5.8 show that the smallest scales converged before the intermediate scales, indeed, they appear to converge at a

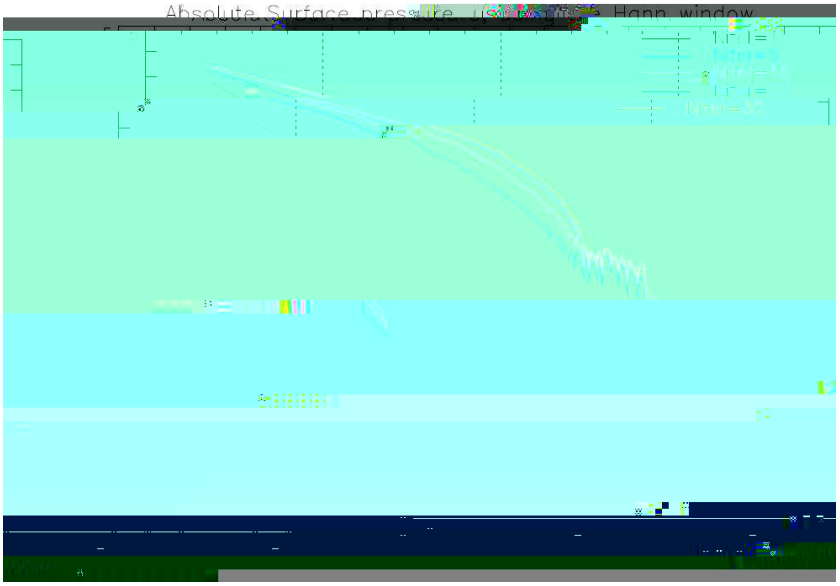


Figure 5.7: Plot showing the binned periodogram data, using the Hann window, of the absolute surface pressure.

mations and is rather more complex than the barotropic vorticity equation. We therefore believe the high frequency region to be an artifact of using the nonlinear model and the 'real-life' data it produces.

To further the comparison between the forecast error at the different inner loop iterations, we now looked at producing a relative plot, which involved dividing the absolute plots (Figures 5.5 and 5.6) by the variance of the incremental data at each corresponding iteration. The resultant plot is shown in figure 5.9.

The result displayed is formed by multiplying the incremental data by the Hann window. Since the results from using the Bartlett and Hann window are so similar we present only the Hann window results.

Figure 5.9 allows us to compare the evolution even more accurately be-

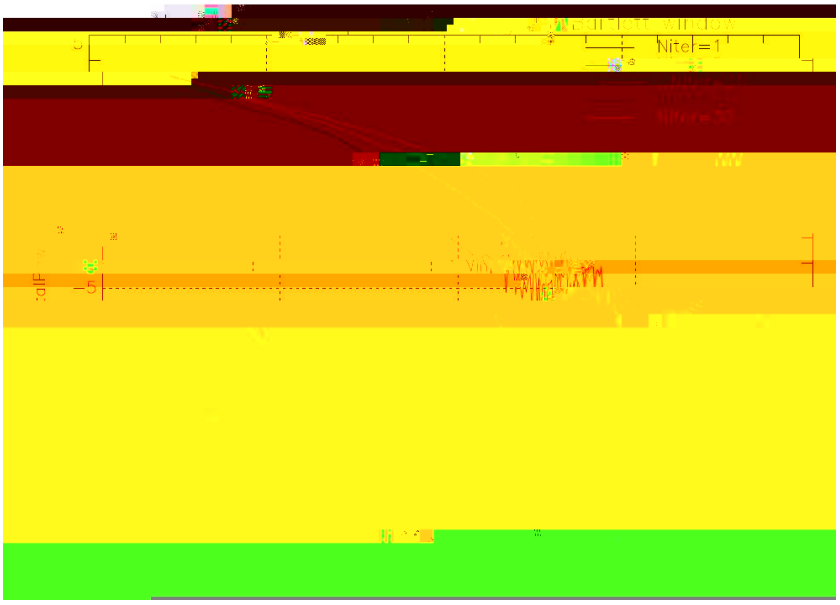


Figure 5.8: Plot showing the binned periodogram data, using the Bartlett window, of the absolute surface pressure.

tween iterations, since the results are now displayed relative to one another. Clearly the relative plot backs up the conclusions drawn from the absolute plots that the larger scales converge before than the intermediate scales, with the smallest scales converging at a similar rate to the largest scales. As we can see, the largest and smallest scales have converged by the 10th iteration with the intermediate scales still to converge at the 30th iteration.

In chapter 6 we conclude by summing up the results obtained in this section and comparing them to previous work. In the next chapter we shall also discuss limitations of out method and what our results mean for future work.

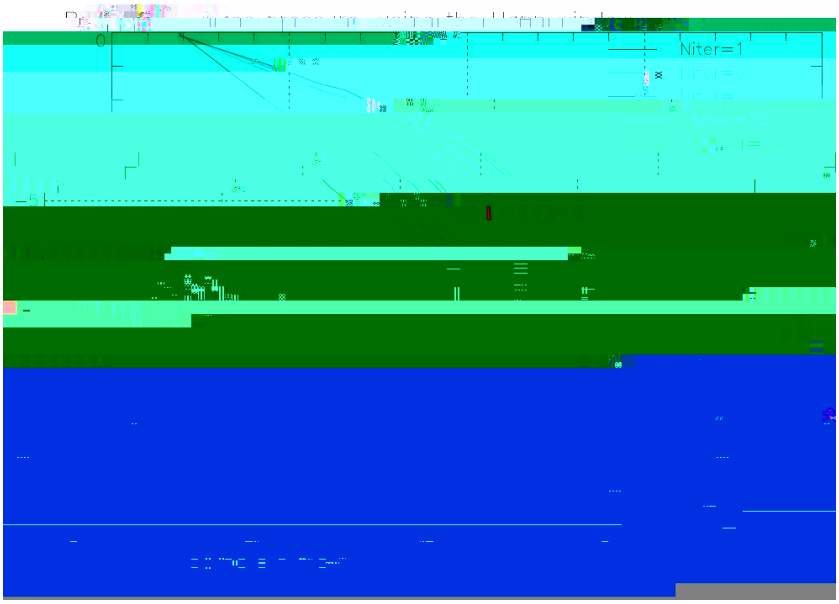


Figure 5.9: Plot showing the binned periodogram data, using the Hann window, of the relative surface pressure.

Chapter 6

Conclusion

In this chapter we draw conclusions based on our results on the evolution of the scales present in the forecast errors of the incremental 4D-Var method, applied to the UK Met O□ce numerical weather prediction model. We then compare these results to the results of previous pieces of work which used the simple barotropic vorticity equation model.

The work previously conducted on the evolution of scales present in the forecast error was only done using the simple barotropic vorticity equation model. In this project we used the nonlinear numerical weather prediction model in order to produce results never obtained before.

In order to view the evolution of the scales present in the forecast error of the incremental 4D-Var method we transformed the data into Fourier space and subsequently formed a 2D periodogram. We discovered that by manipulating and massaging the data we were able to produce an estimate of the power spectrum of the incremental data.

The manipulated 2D periodogram data was then binned and reformed to produce a 1D binned periodogram, which showed the average power in each

mode number bin plotted against the mid-point mode number value of the corresponding bin.

By looking at the incremental results after 1, 5, 10, 20 and 30 iterations of the inner loop, we were able to study the evolution of the scales present in the forecast error. From this we were able to deduce that the error in the largest scales converged quicker than in the intermediate scales, with the smallest scales converging at a similar rate to the largest scales. The results showed that by the 10th iteration the largest and smallest scale errors had converged but the intermediate scale errors still had not converge the-2309-291(t)1(he)]TJ0-23

du09-cin

possible with regards to the physical states, where as the barotropic vorticity

Bibliography

[1] George Arfken (1985) *Mathematical Methods for Physicists*, Miami University, Academic Press, Third Edition, 760-790

[2] J.S. Bendat. and A.G. Piersol (2000) *Random Data Analysis and Measurement Procedures*, Wiley Interscience, Third Edition, 396-397

[3] F. Bouttier and P. Courtier (1999) *Data assimilation concepts and methods*. ECMWF Meteorological Training Course Lecture Series, Web resource:- http://www.ecmwf.int/newsevents/training/rcourse_notes/DATA_ASSIMILATION/ASSIM_CONCEPTS/Assim_concepts11.html, last accessed 17/08/06

[4] K.G. Beauchamp and C.K. Yuen (1979) *Digital Methods for Signal Analysis*, George Allen and Unwin Ltd, 93-95

[5] R.B. Blackman and J.W. Tukey (1959) *The Measurement of Power Spectra*, Dover Publications, 69-85

[6] E. Oran Brigham (1988) *The Fast Fourier Transform and its Applications*, Prentice-Hall International, 83-85 and 240-241

[7] C. Chatfiel (2004) *The Analysis of Time Series - An Introduction*, Chapman and Hall/CRC, Sixth Edition, 126-141

[8] G. Inverarity, Private Communication, 2006

[9] S. Laroche and P. Gauthier (1998) *A validation of the incremental formulation of 4D variational data assimilation in a nonlinear barotropic flow*, Tellus (1998), 50A, 557-572

[10] A. S. Lawless, S. Gratton and N. K. Nichols (2005) *An Investigation of Incremental 4D-Var Using Non-Tangent Linear Models*, Quarterly Journal of the Royal Meteorological Society, Vol. 131, 459-476

[11] N. K. Nichols (2003) *Data Assimilation: Aims and Basic Concepts*, Published in *Data Assimilation for the Earth System*, by R. Swinbank, V. Shutyaev and W.A. Lahoz, Kluwer Academic, 9-20

[12] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery (1992) *Numerical Recipes in Fortran - The Art of Scientific Computing*, Cambridge University Press, Second Edition, 543-545.

[13] U. Ruede updated by C.C. Douglass (1993) *What are Multigrid Methods?*, 1995, Web resource:-

http://www.mgnet.org/mgnet/tutorials
/xwb/mg.html, last accessed 17/06/08

[14] R.B. Stull: *An Introduction to Boundary Layer Meterology*, Kluwer Academic Publications, 295-313

[15] M. Tanguay, P. Bartello, and P. Gauthier (1995) *Four-dimensional data assimilation with a wide range of scales*, Tellus(1995), 47A, 974-997