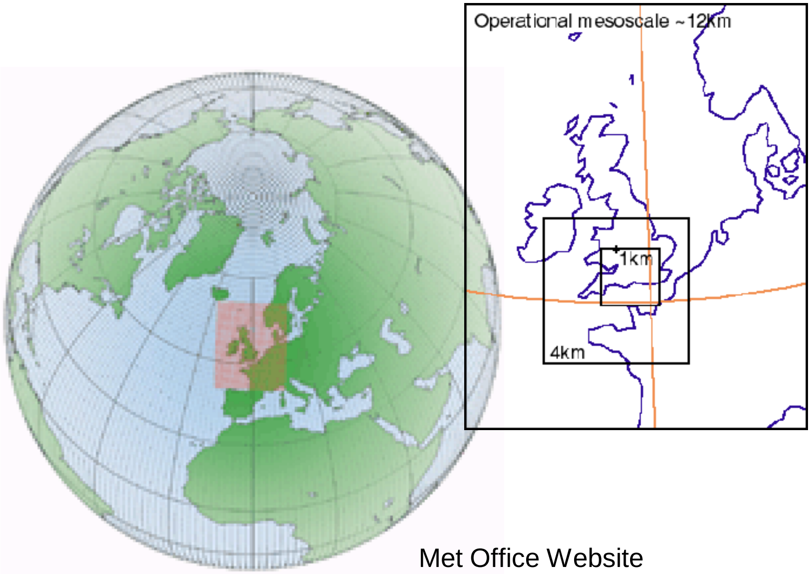
Issues in High Resolution Data Assimilation



Met Office Website

Nancy Nichols

Gillian Baxter, Sarah Dance, Amos Lawless, Sue Ballard*

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

* Met Office

University of Reading

NATURAL
ENVIRONMENT
RESEARCH COUNCIL

National Centre for Earth Observation

NATURAL ENVIRONMENT RESEARCH COUNCIL

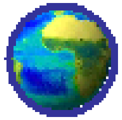
NCEO: **Delivering world-class science by unlocking the full potential of Earth Observation to monitor, diagnose and predict environmental and climate change, and ensuring that scientific advances translate into public good.**

T1: EO for climate diagnosis and prediction (Integrating Theme)

- Exploiting EO to improve national capability for climate prediction over timescales from months to decades
- Reducing and quantifying uncertainty

Leaders: Prof K Haines (Reading)
Prof A Slingo (Reading)
Dr S Laxon (UCL)
Prof P Cox (Exeter)

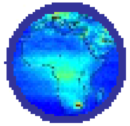
T2: Monitoring, diagnosis, re-analysis and prediction of the global carbon cycle



- Understanding the feedbacks between physical and biological processes involving the carbon cycle, in order to predict changes in carbon fluxes at the Earth's surface

Leaders: Prof S Quegan (Sheffield)
Prof J Aiken (PML)

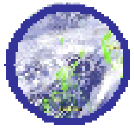
T3: Atmospheric composition: air quality and climate



- Developing an integrated approach to the analysis of satellite measurements, to provide new information on atmospheric composition and aerosols for air-pollution forecasting and testing climate models

Leaders: Dr B Kerridge (RAL)
Prof M Chipperfield (Leeds)

T4: High resolution predictions of hazardous weather, floods and water resources



- Developing capability to forecast hazardous weather and hydrological consequences
- Understanding multiscale dynamics
- Developing novel assimilation techniques for highly non-linear processes

Leaders: Prof R Gurney (U of Reading)
Dr S Dance (U of Reading)

T5: Cryosphere and polar oceans

- Using new EO data to quantify changes in the mass balance of the cryosphere and to develop new models to represent the relevant processes in coupled climate prediction models
- Determining the impact of polar melt on the circulation of the ocean

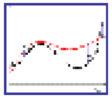
Leader: Dr S Laxon (UCL)

T6: Dynamic Earth and geo-hazards

- Using global satellite measurements of the Earth's surface and volcanic gas emissions to advance knowledge of processes responsible for earthquakes, tsunamis and volcanoes, and hence developing better warning systems

Leader: Prof B Parsons (Oxford)

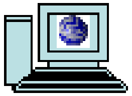
T7: Data assimilation and treatment of uncertainty (Cross Cutting Theme)



- Developing the theory of data assimilation, including methods to treat data and model uncertainty, to underpin applications in NCEO and partner agencies
- Promoting collaboration with groups funded by the EPSRC on research into the underpinning theory

Leader: Prof I Roulstone (Surrey)

EO informatics (Underpinning Theme)



- Exploiting and developing e-science and new data informatics technologies in order to make EO data and derived products from multiple data centres more easily accessible, and to ensure proper archiving of data in line with NERC policy

Leaders: Dr B Lawrence (RAL)
Dr G Robinson (Reading)

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

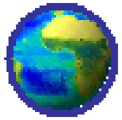
University of Reading

T1: EO for climate diagnosis and prediction (Integrating Theme)

- Exploiting EO to improve national capability for climate prediction over timescales from months to decades
- Reducing and quantifying uncertainty

Leaders: Prof K Haines (Reading)
Prof A Slingo (Reading)
Dr S Laxon (UCL)
Prof P Cox (Exeter)

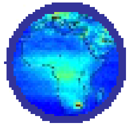
T2: Monitoring, diagnosis, re-analysis and prediction of the global carbon cycle



- Understanding the feedbacks between physical and biological processes involving the carbon cycle, in order to predict changes in carbon fluxes at the Earth's surface

Leaders: Prof S Quegan (Sheffield)
Prof J Aiken (PML)

T3: Atmospheric composition: air quality and climate

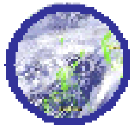


- Developing an integrated approach to the analysis of satellite measurements, to provide new information on atmospheric composition and aerosols for air-pollution forecasting and testing climate models

Leaders: Dr B Kerridge (RAL)
Prof M Chipperfield (Leeds)

T4: High resolution predictions of hazardous weather, floods and water resources

*



- Developing capability to forecast hazardous weather and hydrological consequences
- Understanding multiscale dynamics
- Developing novel assimilation techniques for highly non-linear processes

Leaders: Prof R Gurney (U of Reading)
Dr S Dance (U of Reading)

T5: Cryosphere and polar oceans

- Using new EO data to quantify changes in the mass balance of the cryosphere and to develop new models to represent the relevant processes in coupled climate prediction models
- Determining the impact of polar melt on the circulation of the ocean

Leader: Dr S Laxon (UCL)

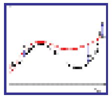
T6: Dynamic Earth and geo-hazards

- Using global satellite measurements of the Earth's surface and volcanic gas emissions to advance knowledge of processes responsible for earthquakes, tsunamis and volcanoes, and hence developing better warning systems

Leader: Prof B Parsons (Oxford)

T7: Data assimilation and treatment of uncertainty (Cross Cutting Theme)

*



- Developing the theory of data assimilation, including methods to treat data and model uncertainty, to underpin applications in NCEO and partner agencies
- Promoting collaboration with groups funded by the EPSRC on research into the underpinning theory

Leader: Prof I Roulstone (Surrey)

EO informatics (Underpinning Theme)

- Exploiting and developing e-science and new data informatics technologies in order to make EO data and derived products from multiple data centres more easily accessible, and to ensure proper archiving of data in line with NERC policy

Leaders: Dr B Lawrence (RAL)
Dr G Robinson (Reading)

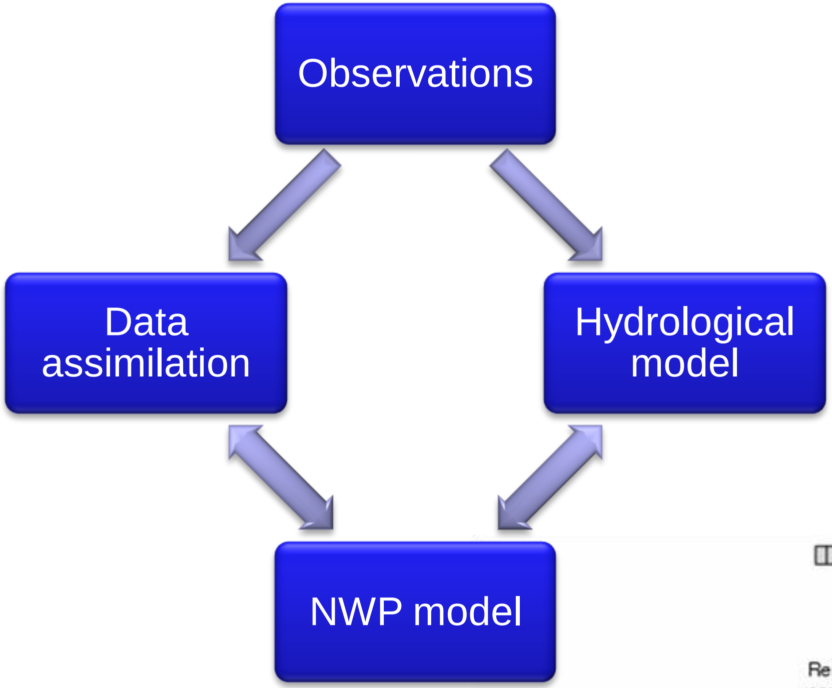
National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

Outline

- Motivation
- Challenges
- Multi-scale Modelling
- Summary and Outlook

Hazardous weather and flooding


Observations

Data assimilation

Hydrological model

NWP model

bbc.co.uk


Birmingham tornado 2005

bbc.co.uk

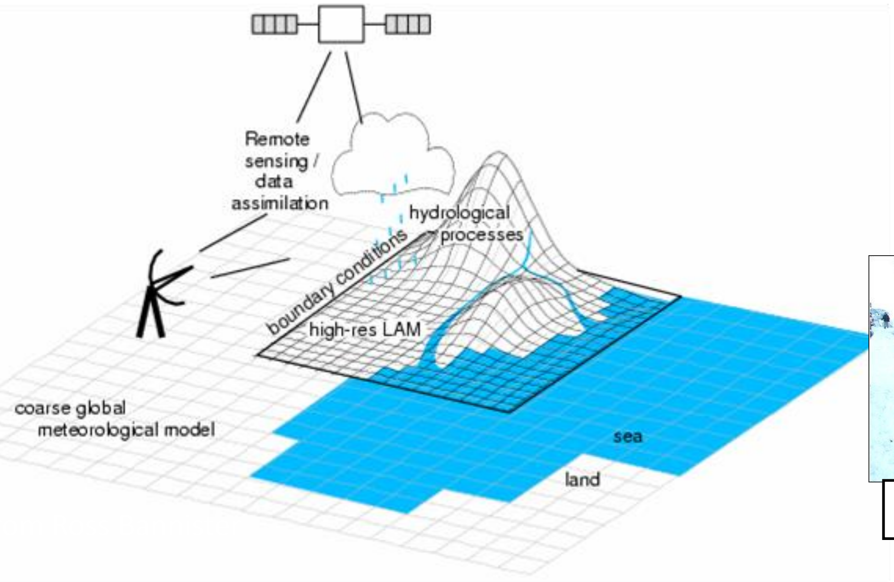

Heathrow fog, Christmas 2006

Leeds University


Boscastle storm 2004

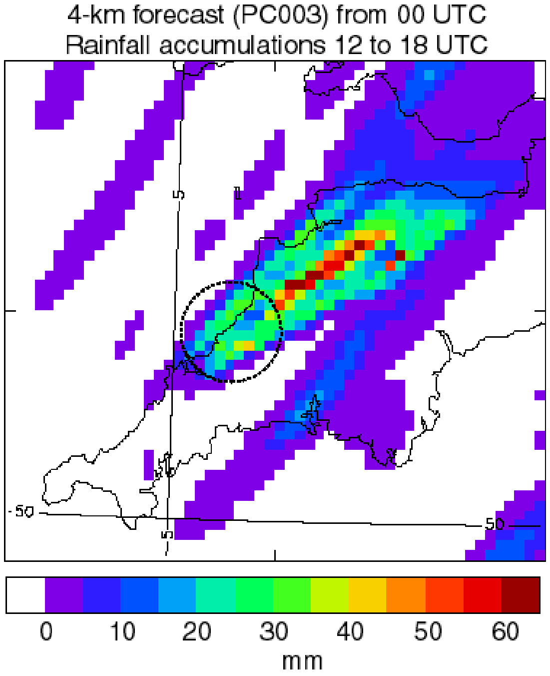
bbc.co.uk


Snow, 2009



- New observation types providing detail on required scales

- Operational storm-scale (1.5km) limited area models now expected – possibly higher resolution in future.


4-km forecast (PC003) from 00 UTC
Rainfall accumulations 12 to 18 UTC

- Improvements in hydrological models, including increased interest in the use of more sophisticated data assimilation techniques.

Data assimilation on convective scales is a NEW problem – very different in character from assimilation on synoptic scales.

What are the challenges?

Challenges

1. Observations
2. Background Covariances
3. Multi-scale Dynamics / Coupled Systems
4. Nonlinearity and Uncertainty
5. Model Reduction

Challenges

1. Observations
2. Background Covariances
3. Multi-scale Dynamics / Coupled Systems
4. Nonlinearity and Uncertainty
5. Model Reduction

Multi-Scale Dynamics

Strong dynamical forcings and feedback exist between synoptic and storm-scale systems. In high resolution convective models:

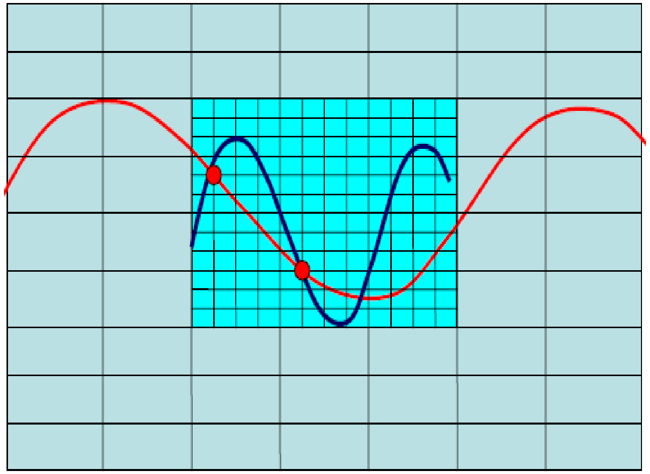
- need to update fine-scale information while preserving large scale information
- need lateral boundary conditions for nested limited area models from synoptic-scale data
- need to retain rapid convergence of all important scales in the optimization algorithm

Question:

How are different scales treated in a LAM?

- Study aliasing problems in limited area models: examine how different wave lengths are projected onto the limited area analysis, using a simple nested advection-diffusion model.

- Examine methods for combining longer wave-lengths from the global model with shorter wave-lengths from the LAM.



National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

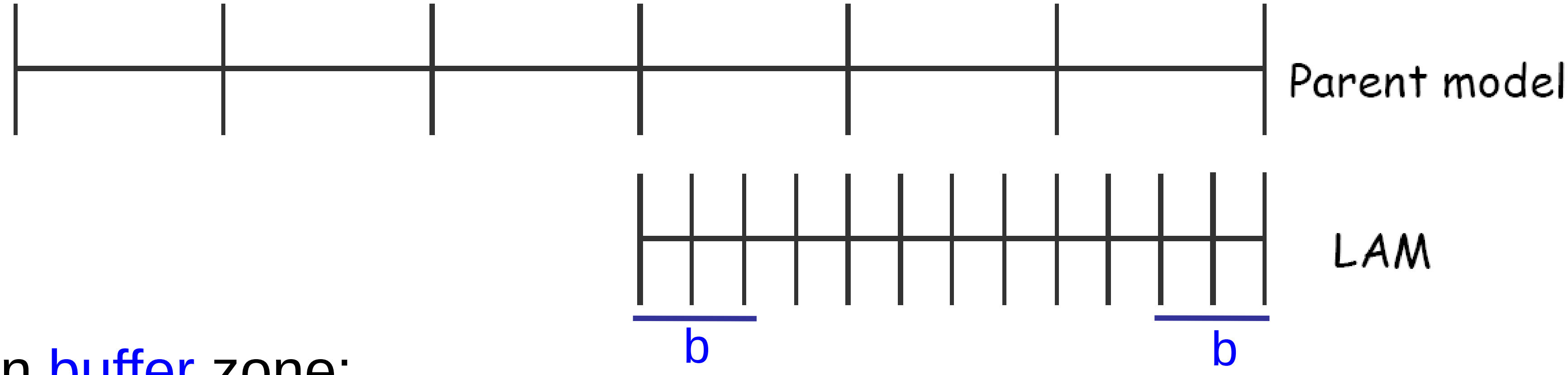
University of Reading

Model

The 1D linear advection-diffusion equation

$$u_t + c u_x = \sigma u_{xx}$$

with periodic boundary conditions for the parent model and the parent analysis for the LAM boundaries. Discretization is explicit time, up-wind advection and centred diffusion.



In buffer zone:

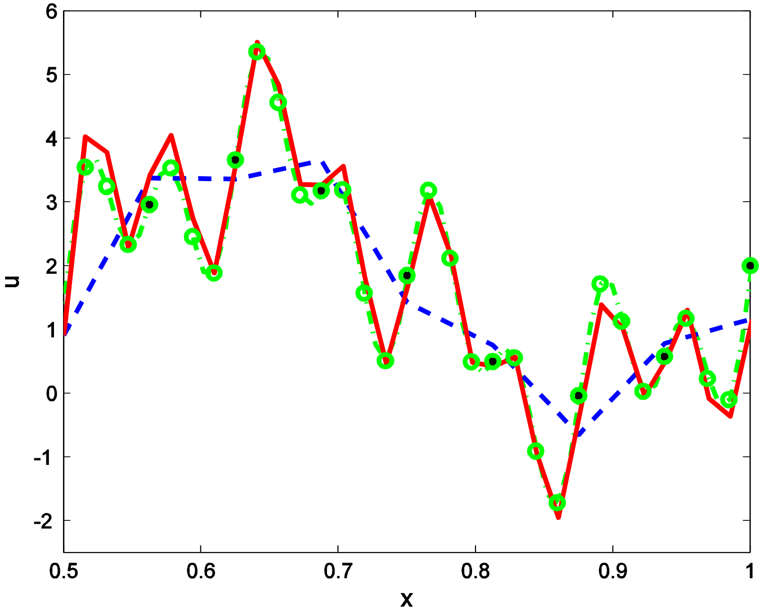
$$u_i^{new} = (1-\omega)u_i^L + \omega u_i^G \quad , \qquad \omega = 1 - [(i-1)/b]$$

Assimilation System

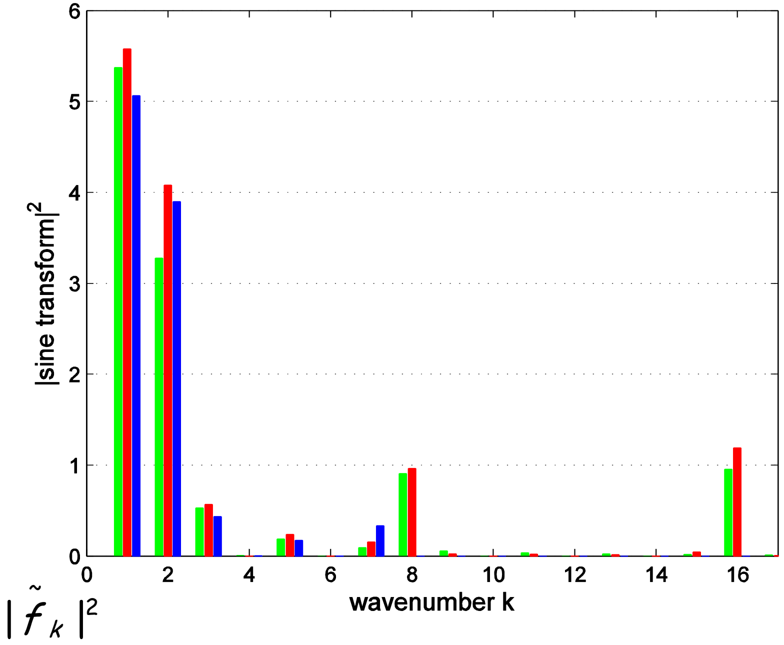
- Uses 4DVar
- Transforms to spectral space using double-sine control variable transform
- Perfect observations at all points
- LAM boundary conditions from parent analysis
- Davies Relaxation at LAM boundaries
- High Resolution LAM = 4 x parent
- High Resolution truth = 2 x LAM

Experiment 1 : Long and short waves

$$\text{truth} = 2\sin(x/4) + 2\sin(2x) + \sin(8x) + \sin(16x)$$



LAM domain



Power spectrum

o observations   --- truth   --- parent analysis   --- LAM analysis

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

Summary:

- Higher resolution allows higher wave-numbers to be captured by the LAM
- A large proportion of the “long wave” information is aliased onto wave-number k=1
- Some “long wave” information is aliased onto higher wave-numbers
- These conclusions can be shown to hold mathematically for a general case using discrete Fourier transforms

Assimilation in Spectral Space

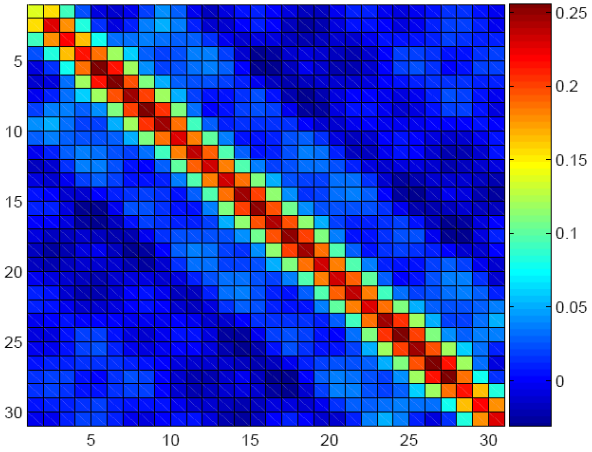
$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + \frac{1}{2}\sum_{k=0}^{t}(\mathbf{y}_k - h_k(\mathbf{x}_k))^T \mathbf{R}_k^{-1}(\mathbf{y}_k - h_k(\mathbf{x}_k))$$

sine transform: $x = Uz$, $\mathbf{\Sigma}^{-1} = \mathbf{U}^T \mathbf{B}^{-1} \mathbf{U}$

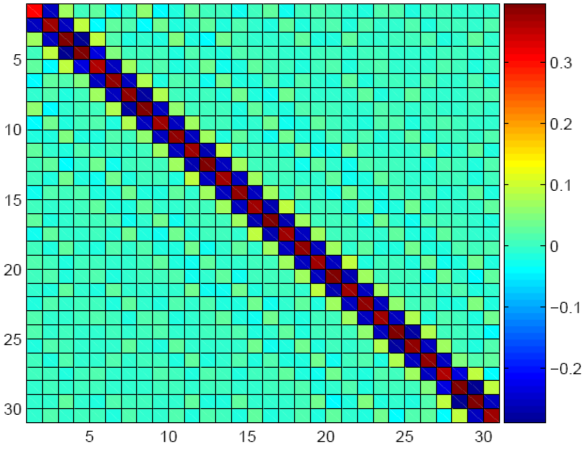
$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{z}_0 - \mathbf{z}_b)^T \boxed{\mathbf{\Sigma}^{-1}}(\mathbf{z}_0 - \mathbf{z}_b) + \frac{1}{2}\sum_{k=0}^{t}(\mathbf{y}_k - h_k(\mathbf{U}\mathbf{z}_k))^T \mathbf{R}_k^{-1}(\mathbf{y}_k - h_k(\mathbf{U}\mathbf{z}_k))$$

Background Matrix B

B_1



B_2



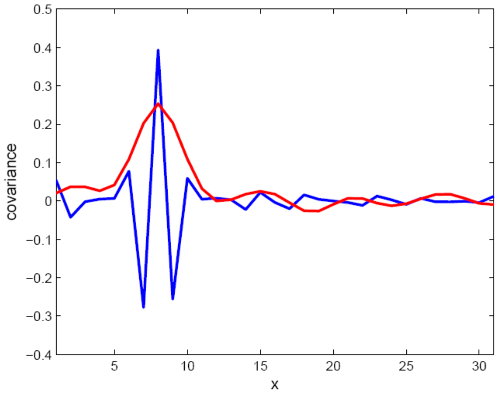
$\Sigma_1 = \mathrm{diag}\{1.0, 0.5, 0.1, 0.01, 0.005\}$

$\Sigma_2 = \mathrm{diag}\{0.005, 0.01, 0.1, 0.5, 1.0\}$

Correlation structure

Red = B_1

Blue = B_2



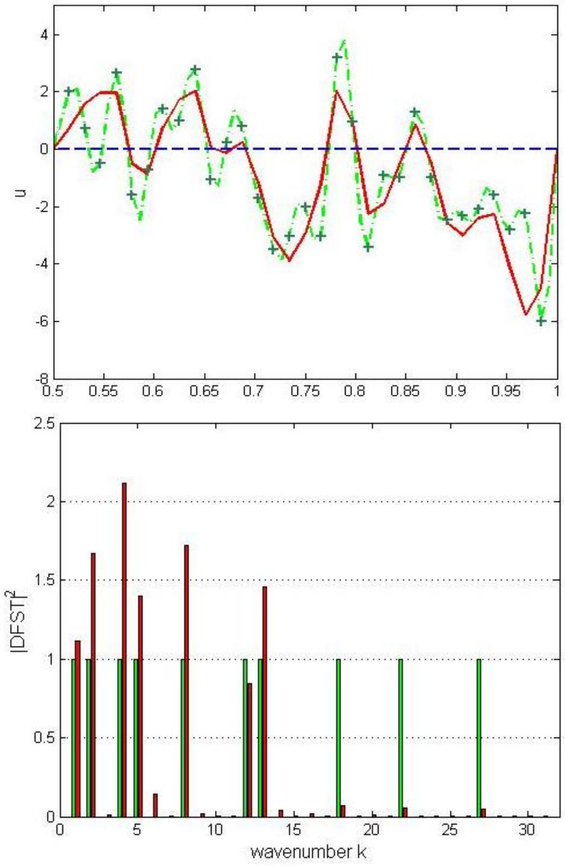
National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

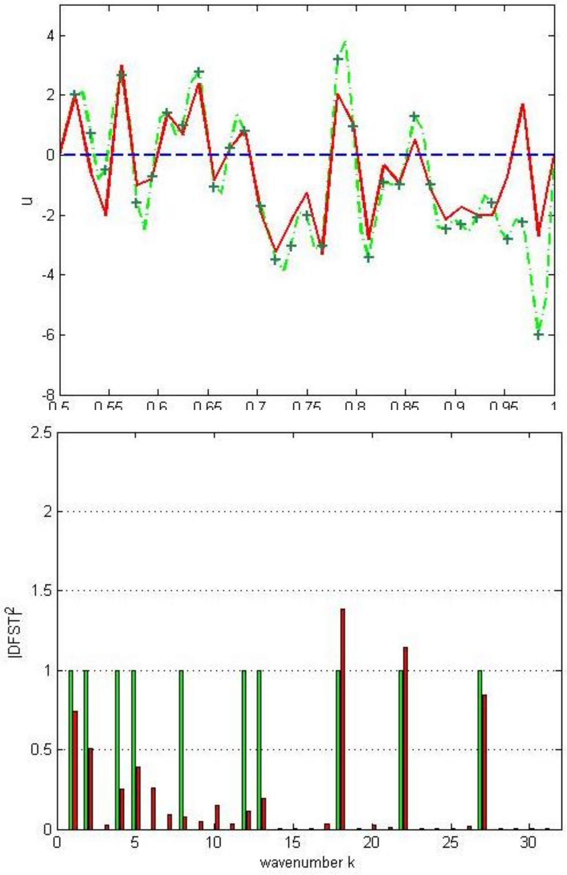
University of Reading

Different weightings in spectral space on background

$$\Sigma^{-1} = \left(\begin{array}{c:c} 0 & \\ \hline & I \end{array}\right)$$

$$\Sigma^{-1} = \left(\begin{array}{c:c} I & \\ \hline & 0 \end{array}\right)$$





LAM domain

Power spectrum

o observations --- truth --- parent analysis --- LAM analysis

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

Conclusions

- Wave-lengths shorter than the resolution of the global models can be analysed in the LAM, but longer wave-lengths may be incorrectly represented due to aliasing.

- Weighting global background differentially in spectral space can affect scales analysed in LAM model.

G.M. Baxter, *PhD Thesis*, 2009

Further Work - ???

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

Further Work - ???

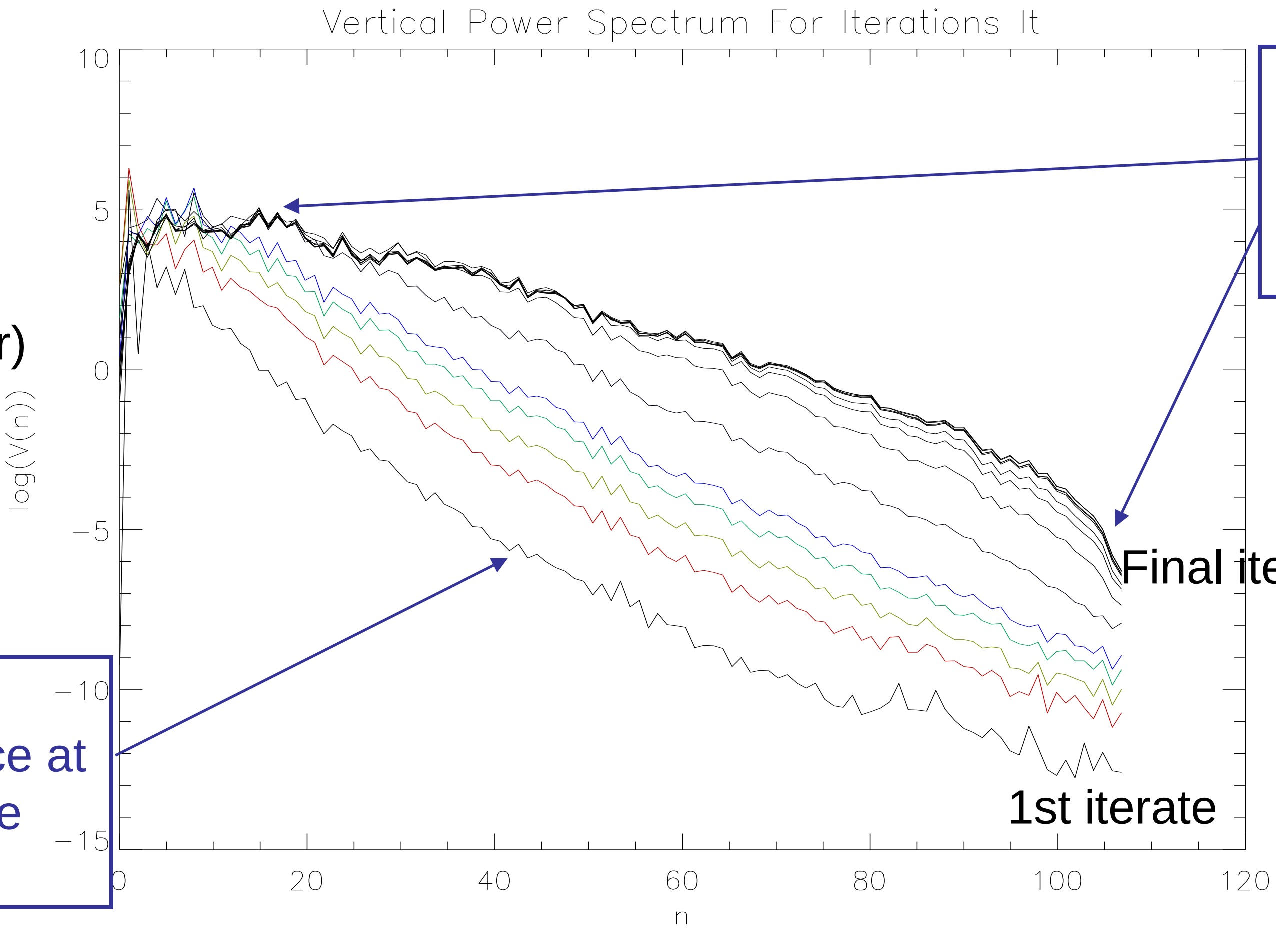
- Test methods for combining long wave information from Global models with high frequency information from the Lam via control variable transforms more generally
- Improved treatment of boundary conditions
- Scale-dependence of 4DVar convergence

Multi-scale systems (2)

- Convergence of the inner loop of the Met Office incremental 4DVar data assimilation system at different Fourier scales has been analysed. Multi-level optimization methods are planned for development.
- Conditioning of the linearized minimization problem as a function of the length-scales in the background covariances and as a function of the observation variances.

Fourier spectrum of pressure increment at lowest model level as inner loop converges.

Log (Power)



Faster convergence at large and small scales

Final iterate

Slower convergence at intermediate scales

1st iterate

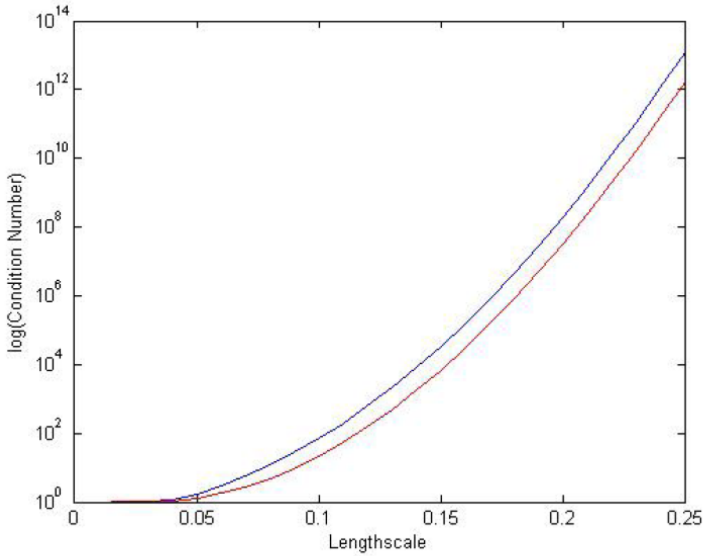
Wave number

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

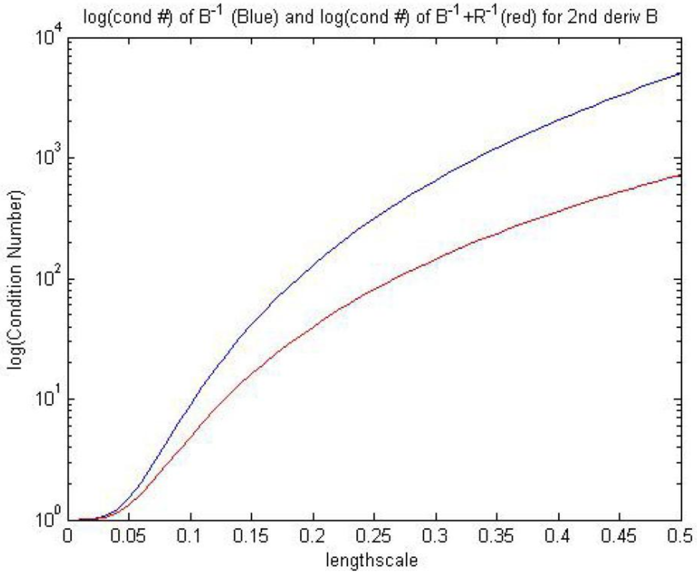
Conditioning of 3DVar

Condition Number of $(B^{-1} + HR^{-1}H^{T})$ vs Length Scale



Periodic Gaussian Exponential

$$\mathbf{B}_{ij} = \sigma_b^2 \exp\left(\frac{-r_{i,j}^2}{2L^2}\right)$$



Laplacian 2nd Derivative

$$\mathbf{B}^{-1} = \gamma^{-1}\left(\mathbf{I} + \frac{l^4}{2\Delta x^4}(\mathbf{L})^2\right)$$

Blue = no obs Red = with obs variances 0.1 / 0.2

Results:

- The Met Office inner loop converges more slowly at mid-wave-lengths. Multigrid approach might improve rates.
- Conditioning of inner linear system decreases with the length scales in the background error covariance matrix.
- Conditioning is improved by the addition of the observations

Haben et al., *Internal Reports*, 2009

The Discrete Fourier sine Transform of a function f_j is

$$\hat{f}_k = \sum_{j=1}^{N-1} f_j \sin(\pi jk/N)$$

where k is the wavenumber and N is the number of gridpoints