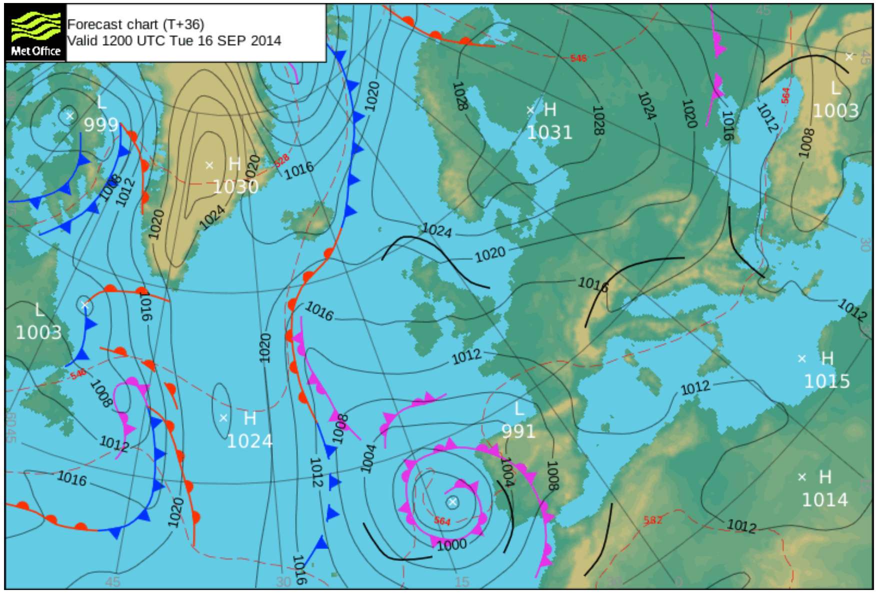
# Variational Data Assimilation




Met Office Website

Nancy Nichols

*Data Assimilation Research Centre*

*University of Reading*

University of Reading

National Centre for Earth Observation

NATURAL ENVIRONMENT RESEARCH COUNCIL

# Outline

- Introduction to data assimilation
- Variational data assimilation
- Incremental 4D variational assimilation
- Treatment of model error
- Conclusions

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# 1.
# The Data Assimilation Problem

# Data Assimilation

## Aim:

Find the best estimate (analysis) of the expected states/parameters of a system, consistent with both observations and the system dynamics given:

- Numerical prediction model
- Observations of the system (over time)
- Background state (prior estimate)
- Estimates of error statistics

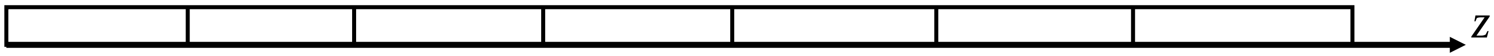
University of Reading
National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example - State Estimation

Diffusion of temperature in a bar




$T^k$ = temperature at grid point $z_k$

States of the system:

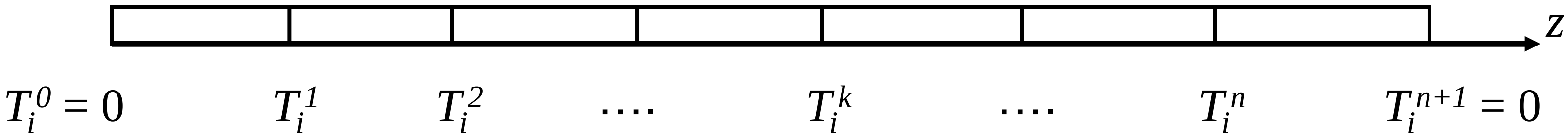
$$\mathbf{T} = \begin{pmatrix} T^1 \\ T^2 \\ \vdots \\ T^n \end{pmatrix}$$

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example - State Estimation

Diffusion of temperature in a bar




$T_i^k$ = temperature at grid point $z_k$ and time $t_i$

States of the system at time $t_i$ :

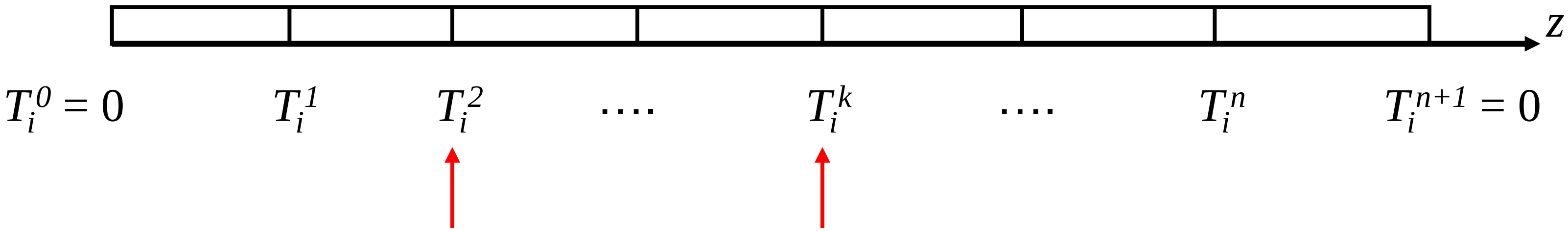
$$\mathbf{T}_i = \begin{pmatrix} T_i^1 \\ T_i^2 \\ \vdots \\ T_i^n \end{pmatrix}$$

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example - Observations

Take observations at grid points at time $t_i$

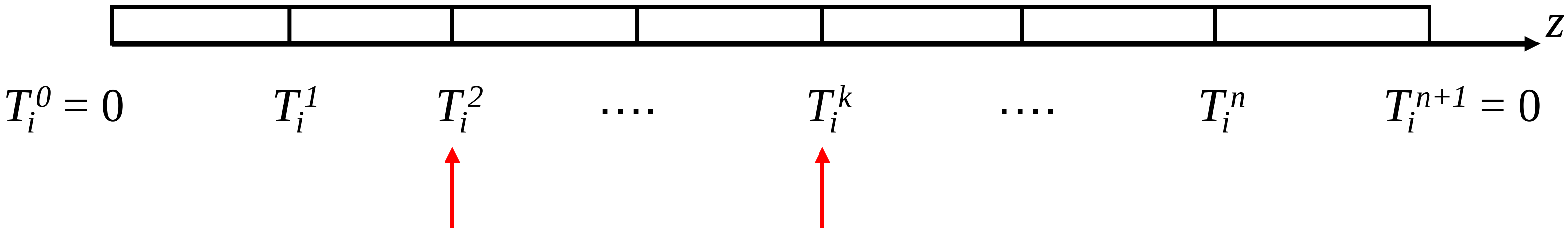



$$\mathbf{y}_i = \begin{pmatrix} T_i^2 + e_i^2 \\ T_i^k + e_i^k \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 & \cdots & 0 \end{pmatrix} \mathbf{T}_i + \mathbf{e}_i$$

where $\mathcal{E}\{\mathbf{e}_i\} = 0\,, \quad \mathcal{E}\{\mathbf{e}_i \mathbf{e}_i^T\} = \mathbf{R}_i$

# Example - Observations

Take observations at grid points at times $t_i$



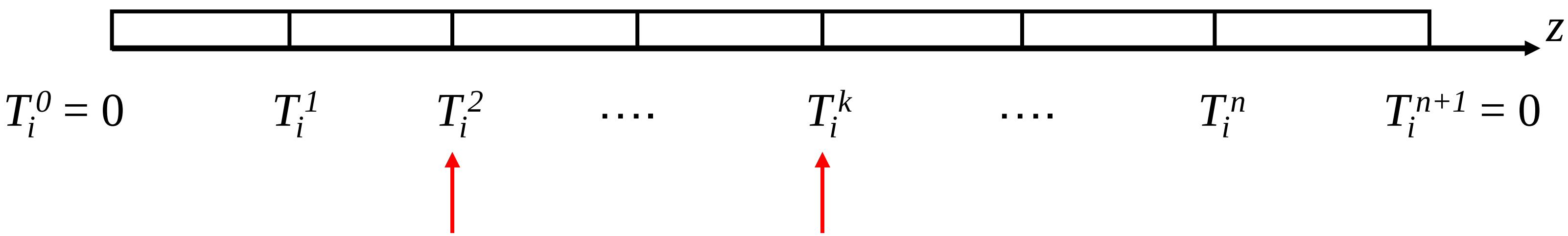

$$\mathbf{y}_i = \begin{pmatrix} T_i^2 + e_i^2 \\ T_i^k + e_i^k \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 & \cdots & 0 \end{pmatrix} \mathbf{T}_i + \mathbf{e}_i$$

implies $$\mathbf{y}_i = \mathbf{H}\mathbf{T}_i + \mathbf{e}_i$$

where $$\mathcal{E}\{\mathbf{e}_i\} = 0\,, \qquad \mathcal{E}\{\mathbf{e}_i\mathbf{e}_i^T\} = \mathbf{R}_i$$

# Example - Observations

Take observations at grid points at times $t_i$




$$\mathbf{y}_i = \begin{pmatrix} T_i^2 + e_i^2 \\ T_i^k + e_i^k \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 & \cdots & 0 \end{pmatrix} \mathbf{T}_i + \mathbf{e}_i$$

implies

$$\boxed{\mathbf{y}_i = \mathbf{H}\mathbf{T}_i + \mathbf{e}_i}$$

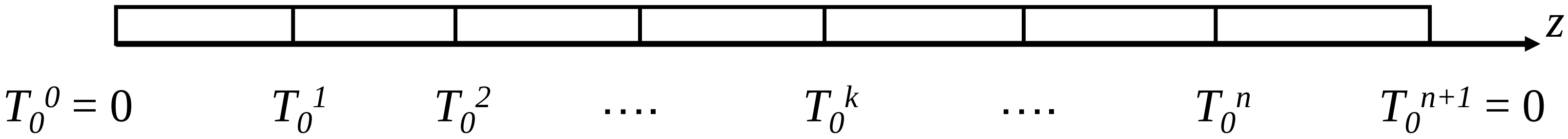
where $\mathcal{E}\{\mathbf{e}_i\} = 0$ , $\mathcal{E}\{\mathbf{e}_i\mathbf{e}_i^T\} = \mathbf{R}_i$

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example - Prior Estimate

Prior estimate at time $t_0$ at all grid points $z_k$




$$\mathbf{T}_b = \begin{pmatrix} T_0^1 + e_0^1 \\ T_0^2 + e_0^2 \\ \vdots \\ T_0^n + e_0^n \end{pmatrix} = \mathbf{T}_0 + \mathbf{e}_b$$

where $\mathcal{E}\{\mathbf{e}_b\} = 0$ , $\mathcal{E}\{\mathbf{e}_b \mathbf{e}_b^T\} = \mathbf{B}$

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example - Data Assimilation Problem

Prior: $$\mathbf{T}_b = \mathbf{T}_0 + \mathbf{e}_b$$

Observations: $$\mathbf{y}_0 = \mathbf{H}\mathbf{T}_0 + \mathbf{e}_0$$

**Question:** can we estimate the state of the system $\mathbf{T}_0$ at $t_o$ from this information? How accurate is the estimate?

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example

Using these equations

implies:

$$\mathbf{y}_0 - \mathbf{H}\mathbf{T}_0 = \mathbf{e}_0$$

= a set of linear equations for $\mathbf{T}_0$ .

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example

Using these equations

implies:

$$\begin{aligned} \mathbf{T}_b - \mathbf{T}_0 &= \mathbf{e}_b \\ \mathbf{y}_0 - \mathbf{H}\mathbf{T}_0 &= \mathbf{e}_0 \end{aligned}$$

= a set of linear equations for $\mathbf{T}_0$ .

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example - Solution

Find the solution that minimizes the error variance and gives the weighted least square error:

$$\min_{\mathbf{T}_0} \quad \mathbf{e}_b^T \mathbf{B}^{-1} \mathbf{e}_b + \mathbf{e}_0^T \mathbf{R}_0^{-1} \mathbf{e}_0$$

# Example - Solution

Find the solution that minimizes the error variance and gives the weighted least square error:

$$\min_{\mathbf{T}_0} \quad \mathbf{e}_b^T \mathbf{B}^{-1} \mathbf{e}_b + \mathbf{e}_0^T \mathbf{R}_0^{-1} \mathbf{e}_0 \quad =$$

$$\min_{\mathbf{T}_0} \; (\mathbf{T}_b - \mathbf{T}_0)^T \mathbf{B}^{-1} (\mathbf{T}_b - \mathbf{T}_0) + \\ + (\mathbf{y}_0 - \mathbf{H}\mathbf{T}_0)^T \mathbf{R}_0^{-1} (\mathbf{y}_0 - \mathbf{H}\mathbf{T}_0)$$

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example - Solution

Find the solution that minimizes the error variance and gives the weighted least square error:

$$\min_{\mathbf{T}_0} \quad \mathbf{e}_b^T \mathbf{B}^{-1} \mathbf{e}_b + \mathbf{e}_0^T \mathbf{R}_0^{-1} \mathbf{e}_0 \quad =$$

$$\min_{\mathbf{T}_0} \; (\mathbf{T}_b - \mathbf{T}_0)^T \mathbf{B}^{-1} (\mathbf{T}_b - \mathbf{T}_0) + \\ + (\mathbf{y}_0 - \mathbf{H}\mathbf{T}_0)^T \mathbf{R}_0^{-1} (\mathbf{y}_0 - \mathbf{H}\mathbf{T}_0)$$

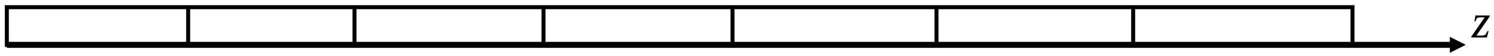
This gives $\mathbf{T}_0$ with minimum variance.

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example - Numerical Model

Difference equation describing diffusion



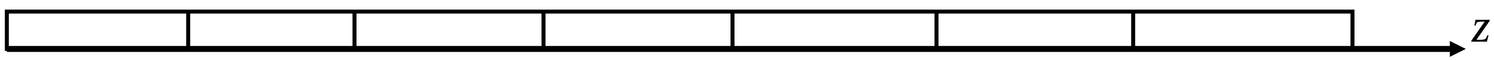

$T_i^0 = 0 \quad T_i^1 \quad T_i^2 \quad \cdots \quad T_i^k \quad \cdots \quad T_i^n \quad T_i^{n+1} = 0$

$$\frac{T_{i+1}^{k} - T_i^{k}}{\delta t} = c\,\frac{T_i^{k+1} - 2T_i^{k} + T_i^{k-1}}{\delta z^2}$$

where $c$ is the diffusion coefficient

# Example - Numerical Model

Difference equation describing diffusion




$T_i^0 = 0$ $\quad T_i^1 \quad T_i^2 \quad \cdots \quad T_i^k \quad \cdots \quad T_i^n \quad T_i^{n+1} = 0$

$$\frac{T_{i+1}^k - T_i^k}{\delta t} = c\,\frac{T_i^{k+1} - 2T_i^k + T_i^{k-1}}{\delta z^2}$$

implies $$T_{i+1}^k = T_i^k + c\,\mu\,(T_i^{k+1} - 2T_i^k + T_i^{k-1})$$

where $c$ is the diffusion coefficient

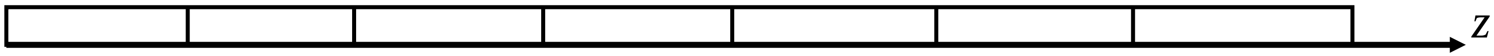
and $\mu = \delta t/\delta z^2$

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example - Numerical Model

Write in matrix-vector form




$T_i^0 = 0$ $\quad T_i^1 \quad T_i^2 \quad \ldots\ldots \quad T_i^k \quad \ldots\ldots \quad T_i^n \quad T_i^{n+1} = 0$

$$T_{i+1}^k = T_i^k + c\,\mu\,(T_i^{k+1} - 2T_i^k + T_i^{k-1})$$

implies

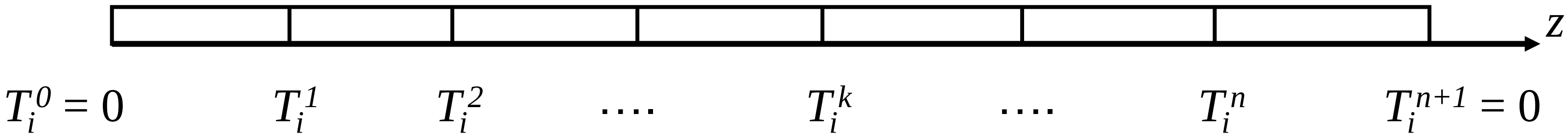
$$\mathbf{T}_{i+1} = \mathbf{T}_i + c\,\mu\,\mathbf{L}\,\mathbf{T}_i \equiv \mathbf{M}\,\mathbf{T}_i$$

where

$$\mathbf{M} = \mathbf{I} + c\,\mu\,\mathbf{L}\ , \quad \mathbf{L} = \begin{pmatrix} -2 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -2 \end{pmatrix}$$

# Example - Numerical Model

Write in matrix-vector form




$$T^k_{i+1} = T^k_i + c\,\mu\,(T^{k+1}_i - 2T^k_i + T^{k-1}_i)$$

implies

$$\mathbf{T}_{i+1} = \mathbf{T}_i + c\,\mu\,\mathbf{L}\,\mathbf{T}_i \equiv \mathbf{M}\,\mathbf{T}_i$$

where

$$\mathbf{M} = \mathbf{I} + c\,\mu\,\mathbf{L}\,, \quad \mathbf{L} = \begin{pmatrix} -2 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -2 \end{pmatrix}$$

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example - System Equations

Prior: $$\mathbf{T}_b = \mathbf{T}_0 + \mathbf{e}_b$$

Model: $$\mathbf{T}_{i+1} = \mathbf{M}\,\mathbf{T}_i$$

Observations: $$\mathbf{y}_i = \mathbf{H}\mathbf{T}_i + \mathbf{e}_i$$

where $$\mathcal{E}\{\mathbf{e}_b\} = 0 \qquad \mathcal{E}\{\mathbf{e}_b\mathbf{e}_b^T\} = \mathbf{B}$$

$$\mathcal{E}\{\mathbf{e}_i\} = 0 \qquad \mathcal{E}\{\mathbf{e}_i\mathbf{e}_i^T\} = \mathbf{R}_i$$

and errors are uncorrelated in time

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example - Data Assimilation Problem

Prior: $$\mathbf{T}_b = \mathbf{T}_0 + \mathbf{e}_b$$

Model: $$\mathbf{T}_{i+1} = \mathbf{M}\,\mathbf{T}_i$$

Observations: $$\mathbf{y}_i = \mathbf{H}\mathbf{T}_i + \mathbf{e}_i$$

**Question:** can we estimate the state of the system $\mathbf{T}_0$ at $t_0$ from this information? How accurate is the estimate?

# Example - YES

Using: $$\mathbf{y}_i = \mathbf{H}\mathbf{T}_i + \mathbf{e}_i = \mathbf{H}\mathbf{M}\mathbf{T}_{i-1} + \mathbf{e}_i$$

implies:

$$\begin{array}{ccccc}
\mathbf{T}_b & - & \mathbf{T}_0 & = & \mathbf{e}_b \\
\mathbf{y}_0 & - & \mathbf{H}\mathbf{T}_0 & = & \mathbf{e}_0 \\
\mathbf{y}_1 & - & \mathbf{H}\mathbf{M}\mathbf{T}_0 & = & \mathbf{e}_1 \\
\mathbf{y}_2 & - & \mathbf{H}\mathbf{M}^2\mathbf{T}_0 & = & \mathbf{e}_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\mathbf{y}_n & - & \mathbf{H}\mathbf{M}^n\mathbf{T}_0 & = & \mathbf{e}_n
\end{array}$$

= a set of linear equations for $\mathbf{T}_0$ .

# Example - Solution

Find the solution that minimizes the error variance and gives the weighted least square error:

$$\min_{\mathbf{T}_0} \; \mathbf{e}_b^T \mathbf{B}^{-1} \mathbf{e}_b + \sum_0^n \mathbf{e}_i^T \mathbf{R}_i^{-1} \mathbf{e}_i \quad =$$

$$\min_{\mathbf{T}_0} \; (\mathbf{T}_b - \mathbf{T}_0)^T \mathbf{B}^{-1} (\mathbf{T}_b - \mathbf{T}_0) + \\ + \sum_{i=0}^{n} (\mathbf{y}_i - \mathbf{H}\mathbf{M}^i \mathbf{T}_0)^T \mathbf{R}_i^{-1} (\mathbf{y}_i - \mathbf{H}\mathbf{M}^i \mathbf{T}_0)$$

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Optimal Estimate

$$\min_{\mathbf{T}_0} \mathcal{J} = \frac{1}{2}(\mathbf{T}_b - \mathbf{T}_0)^T \mathbf{B}^{-1}(\mathbf{T}_b - \mathbf{T}_0) +$$

$$+ \frac{1}{2}\sum_{i=0}^{n}(\mathbf{y}_i - \mathbf{H}\mathbf{T}_i)^T \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathbf{H}\mathbf{T}_i)$$

subject to

$$\mathbf{T}_{i+1} = \mathbf{M}\,\mathbf{T}_i \;, \quad i = 0, 1, \ldots n - 1$$

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Optimal Estimate

$$\min_{\mathbf{T}_0} \mathcal{J} = \frac{1}{2}(\mathbf{T}_b - \mathbf{T}_0)^T \mathbf{B}^{-1}(\mathbf{T}_b - \mathbf{T}_0) +$$

$$+ \frac{1}{2}\sum_{i=0}^{n}(\mathbf{y}_i - \mathbf{H}\mathbf{T}_i)^T \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathbf{H}\mathbf{T}_i)$$

subject to

$$\mathbf{T}_{i+1} = \mathbf{M}\,\mathbf{T}_i \;, \quad i = 0, 1, \ldots n - 1$$

Best Linear Unbiased Estimate

# Optimal Unbiased Estimate

$$\min_{\mathbf{T}_0} \mathcal{J} = \frac{1}{2}(\mathbf{T}_b - \mathbf{T}_0)^T \mathbf{B}^{-1}(\mathbf{T}_b - \mathbf{T}_0) +$$

$$+\frac{1}{2}\sum_{i=0}^{n}(\mathbf{y}_i - \mathbf{H}\mathbf{T}_i)^T \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathbf{H}\mathbf{T}_i)$$

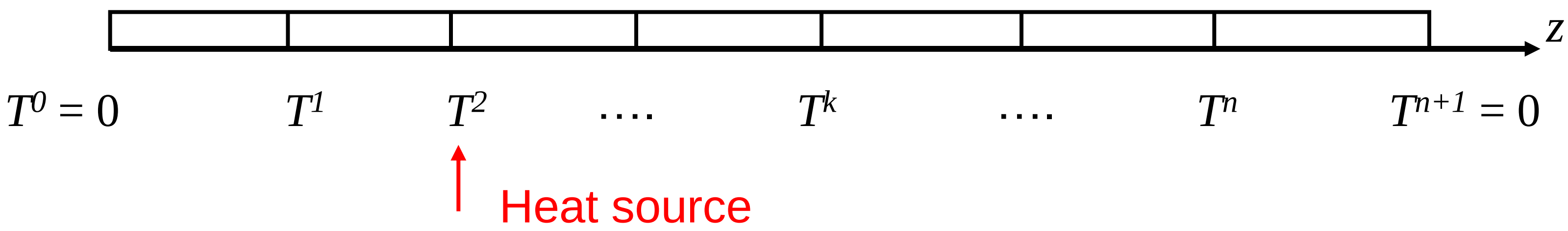
subject to

$$\mathbf{T}_{i+1} = \mathbf{M}\,\mathbf{T}_i \;, \quad i = 0, 1, \ldots n-1$$

## Maximum A Posteriori Likelihood

# Example - Application

Temperature diffusion with source term



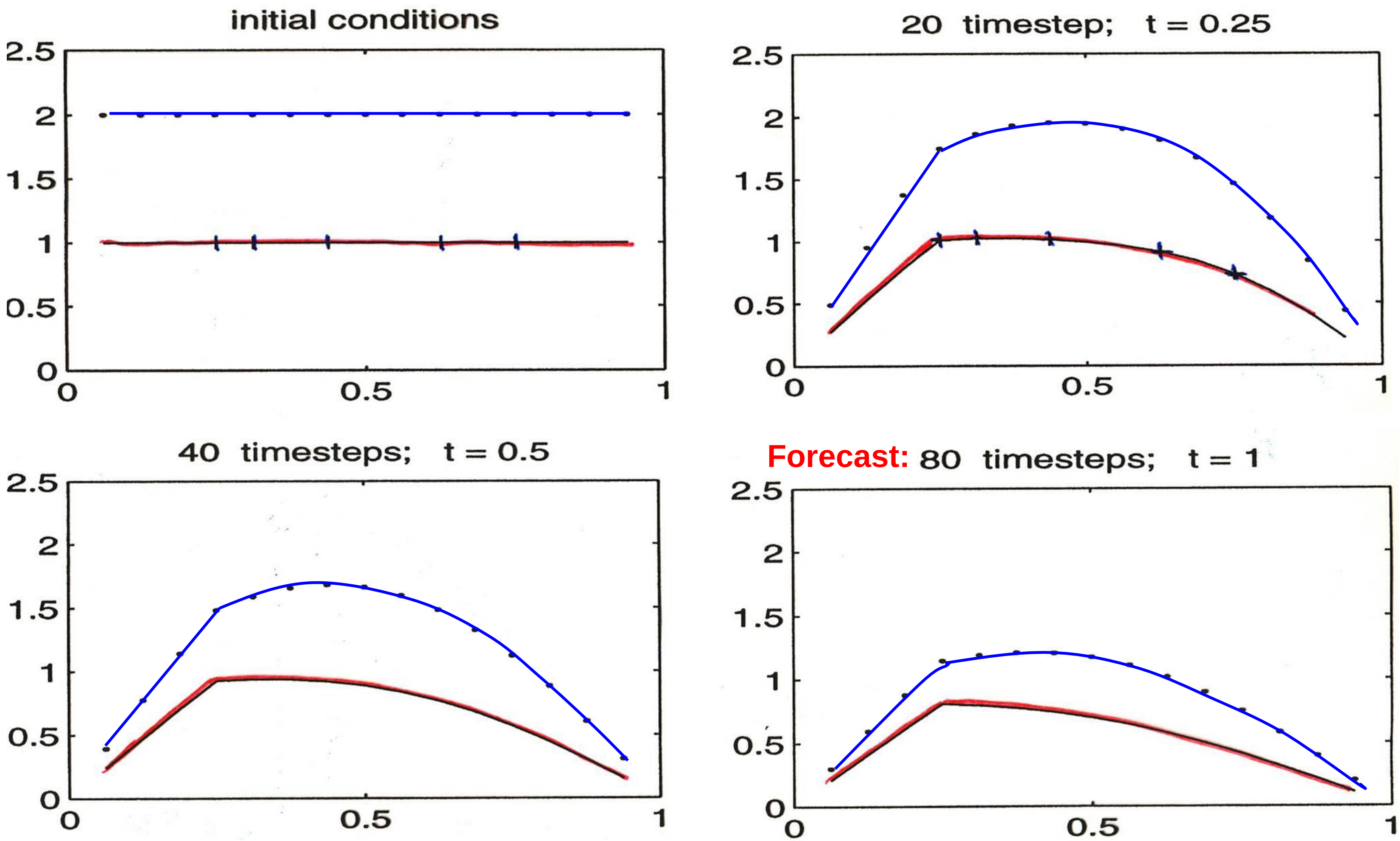

Model:

$$\mathbf{T}_{i+1} = \mathbf{M}\,\mathbf{T}_i + \mathbf{s}_i$$

Twin experiment:

- Truth is solution for $T_0^k = 1$ for all k
- Background is $T_0^k = 2$ for all k
- Observations are from truth with no noise at 5 grid points at every time step for 40 steps

# Heat Equation with Source




Solid = Truth, Dotted = Background, + = Observation, Red = With Assimilation

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# 2.

# Variational Data Assimilation

# 2.

# Variational Data Assimilation

# Optimal Unbiased Estimate

$$\min_{\mathbf{x}_0} \mathcal{J} = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + $$
$$+ \frac{1}{2}\sum_{i=0}^{n}(H[\mathbf{x}_i] - \mathbf{y}_i)^T \mathbf{R}_i^{-1}(H[\mathbf{x}_i] - \mathbf{y}_i)$$

subject to $\mathbf{x}_{i+1} = \mathcal{M}_i(\mathbf{x}_i)\ ,\ i = 0, \ldots, n-1$

$x_b$ - Background state (prior estimate)

$y_i$ - Observations

$H_i$ - Observation operator

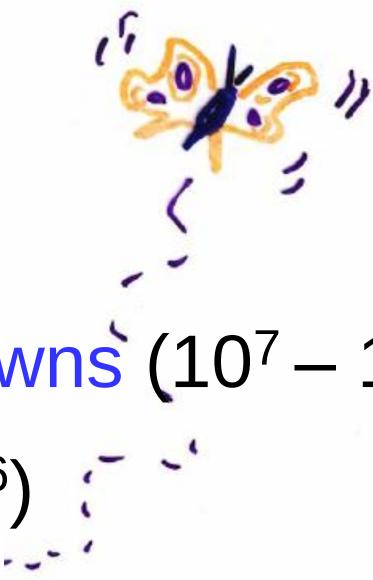
$B$ - Background error covariance matrix

$R_i$ - Observation error covariance matrix

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Significant Properties:



- Very large number of unknowns ($10^7 – 10^8$)
- Few observations ($10^5 – 10^6$)
- System nonlinear unstable/chaotic
- Multi-scale dynamics

# Variational Assimilation

$$\min_{\mathbf{x}_0} \mathcal{J} = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_b) +$$

$$+ \frac{1}{2} \sum_{i=0}^{n} (H[\mathbf{x}_i] - \mathbf{y}_i)^T \mathbf{R}_i^{-1} (H[\mathbf{x}_i] - \mathbf{y}_i)$$

subject to $\mathbf{x}_{i+1} = \mathcal{M}_i(\mathbf{x}_i)$, $i = 0, \ldots, n-1$

Solve iteratively by gradient optimization methods.

Use adjoint methods to find the gradients.

3DVar if n = 0 4DVar if n $\geq$ 1

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Adjoint Model

Define the Lagrangian functional as

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

Then the adjoint equations are

$$\boldsymbol{\lambda}_n = 0$$

$$\boldsymbol{\lambda}_i = \mathbf{M}_i^{T} \boldsymbol{\lambda}_{i+1} - \mathbf{H}_i^T \mathbf{R}_i^{-1} (H_i[\mathbf{x}_i] - \mathbf{y}_i)$$

where $\mathbf{M}_i$ is the linearized dynamical model

and $\mathbf{H}_i$ is the linearized observation operator

# Adjoint Model

**Question** - What are the adjoints?

$\mathbf{M}_i$ is the Jacobian $\frac{\partial \mathcal{M}_i}{\partial \mathbf{x}}$ of the linearized model operator and its adjoint is $\mathbf{M}_i^T$, known as the tangent linear model (TLM)

The adjoint variables $\boldsymbol{\lambda}_k$ measure the sensitivity of the objective function $\mathcal{J}$ to changes in the solutions $\mathbf{x}_k$ of the state equations.

# Adjoint Model

The gradient of $\mathcal{J}$ with respect to the initial condition $\mathbf{x}_0$ is then given by

$$\nabla_0 J = -\lambda_0 + \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b)$$

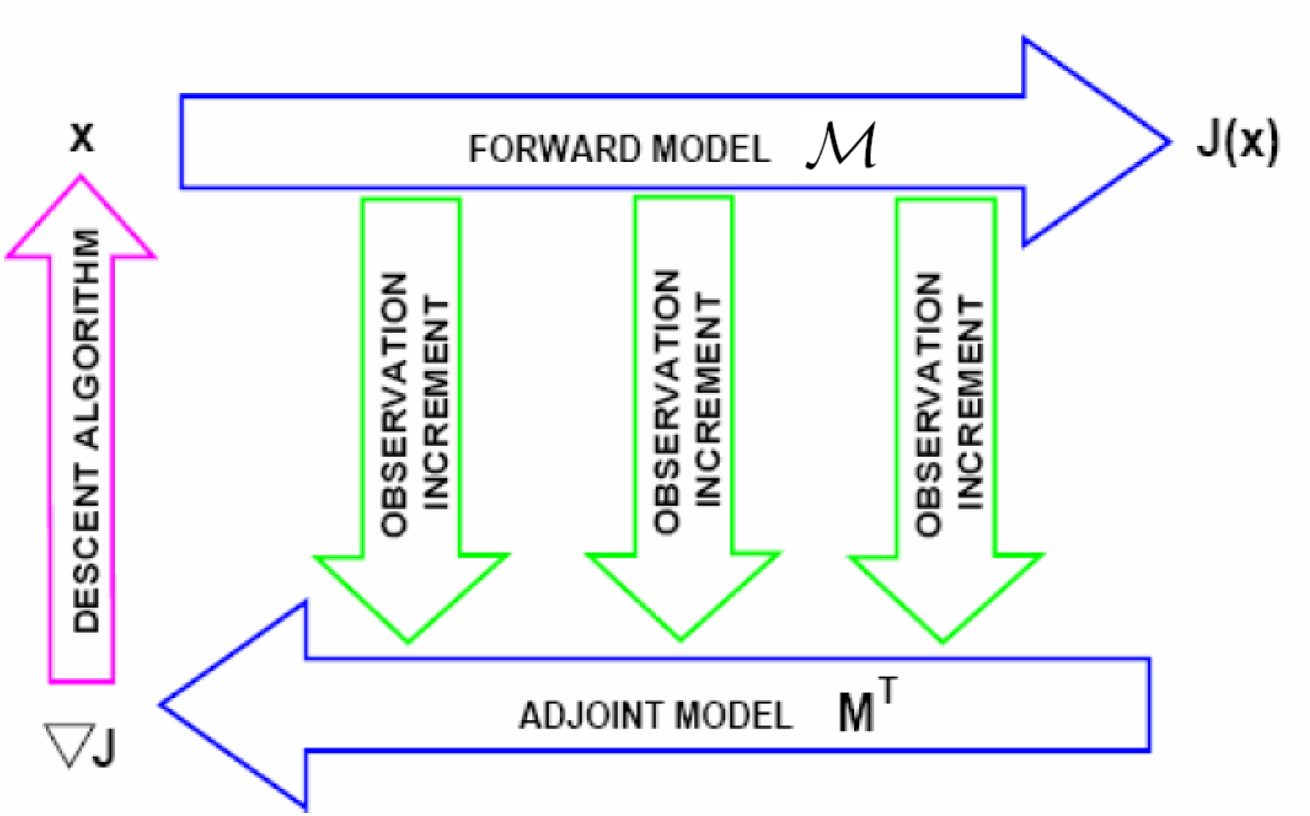
At the optimal the state and adjoint equations must both be satisfied and the gradient must equal to 0 .

# Algorithm

To find the optimal:

- Estimate $\mathbf{x}_0$
- Run the nonlinear model forward; find the 'innovations' $H[\mathbf{x}_i] - \mathbf{y}_i$ and evaluate the objective function $\mathcal{J}$
- Run the adjoint model backward to find $\boldsymbol{\lambda}_0$ and evaluate the gradient $\nabla_0 J$
- Use a gradient nonlinear minimization method to find an improved estimate of $\mathbf{x}_0$
- Repeat until required accuracy is reached.

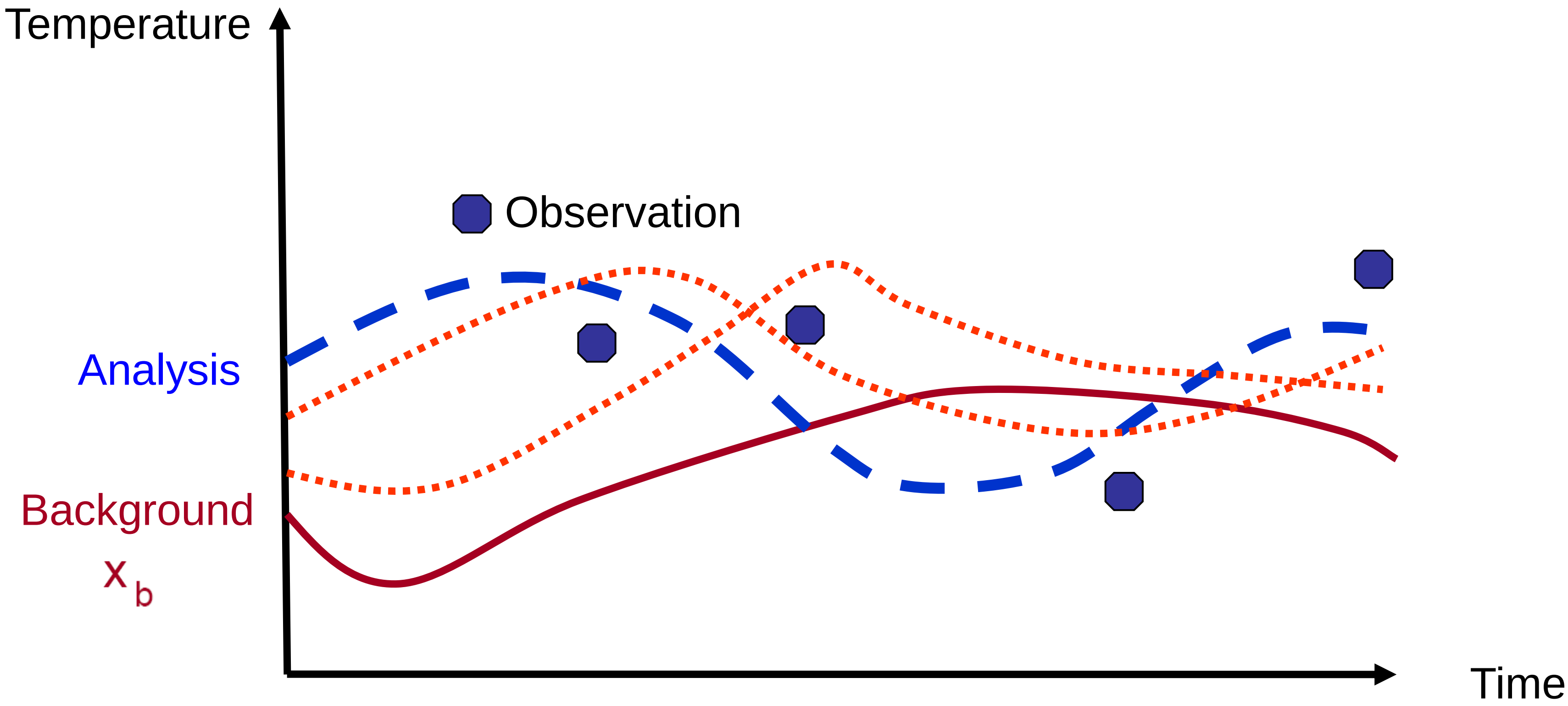
# Algorithm


x
FORWARD MODEL $\mathcal{M}$
J(x)
DESCENT ALGORITHM
OBSERVATION INCREMENT
OBSERVATION INCREMENT
OBSERVATION INCREMENT
ADJOINT MODEL $M^T$
$\nabla J$

# 3.

# Incremental 4D Variational Assimilation

# Incremental 4D-Var




Solve a sequence of linear least squares problems that approximate the nonlinear problem by iteration .

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Incremental 4D-Var

Set $\mathbf{x}_0^{(0)}$ (usually equal to background)

*For k = 0, …, K find:* $\mathbf{x}_{i+1}^{(k)} = \mathcal{M}_i(\mathbf{x}_i^{(k)}), \quad i = 1, \ldots, n$

Solve inner loop linear minimization problem:

$$\tilde{\mathcal{J}}^{(k)}[\delta\mathbf{x}_0^{(k)}] = \frac{1}{2}(\delta\mathbf{x}_0^{(k)} - [\mathbf{x}^b - \mathbf{x}_0{}^{(k)}])^{\mathrm{T}}\mathbf{B}_0^{-1}(\delta\mathbf{x}_0^{(k)} - [\mathbf{x}^b - \mathbf{x}_0{}^{(k)}])$$

$$+ \frac{1}{2}\sum_{i=0}^{n}(\mathbf{H}_i\delta\mathbf{x}_i^{(k)} - \mathbf{d}_i^{(k)})^{\mathrm{T}}\mathbf{R}_i^{-1}(\mathbf{H}_i\delta\mathbf{x}_i^{(k)} - \mathbf{d}_i^{(k)})$$

subject to $\delta\mathbf{x}_{i+1}^{(k)} = \mathbf{M}_i\mathbf{x}_i^{(k)}$ , $\mathbf{d}_i = \mathbf{y}_i - H_i[\mathbf{x}_i^{(k)}]$

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

# Algorithm

To find the optimal:

- Estimate $\mathbf{x}_0$
- Run the nonlinear model forward to find $\mathbf{x}_i$
- Estimate $\delta\,\mathbf{x}_0$ and run the tangent linear model (TLM) forward to find $H_k\delta\mathbf{x}_k - \mathbf{d}_k$ and evaluate the linearized objective function
- Run the adjoint model backward using forcing terms $H_k\delta\mathbf{x}_k - \mathbf{d}_k$ to find $\boldsymbol{\lambda}_0$ and evaluate the gradient of the linearized problem
- Use a gradient minimization method to find an improved estimate of $\delta\,\mathbf{x}_0$
- Update $\mathbf{x}_0$ by adding $\delta\,\mathbf{x}_0$ to old estimate and repeat

# Algorithm

- Incremental 4D-Var without approximations is equivalent to a Gauss-Newton iteration for nonlinear least squares problems.
- In operational implementation we usually approximate the solution procedure:
  - Truncate inner loop iterations
  - Use approximate linear system model
- Theoretical convergence results have been obtained by reference to Gauss-Newton method.

References: Lawless, Gratton and Nichols, *QJ RMetS*, 2005
and Gratton, Lawless and Nichols, *SIAM J on Optimization, 2007*

# Analysis

The analysis $\mathbf{x}_a$ is the optimal solution to the assimilation problem and $\mathbf{x}_a = \mathbf{x}_0 + \mathbf{e}_a$. The uncertainty is given by

$$\mathcal{E}\{\mathbf{e}_a\mathbf{e}_a^T\} \equiv \mathbf{A} = (\mathbf{B}^{-1} + \hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}})^{-1}$$

where

$$\hat{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1\mathbf{M}_0 \\ \mathbf{H}_2\mathbf{M}_1\mathbf{M}_0 \\ \vdots \\ \mathbf{H}_n\mathbf{M}_{n-1}\ldots\mathbf{M}_0 \end{pmatrix} \qquad \hat{\mathbf{R}} = \begin{pmatrix} \mathbf{R}_0 & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{R}_1 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{R}_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \mathbf{R}_n \end{pmatrix}$$

# Conditioning of the Problem

Accuracy/rate of convergence depend on the condition number = $\lambda_{max} / \lambda_{min}$ of the Hessian:

$$\mathbf{A} = \mathbf{B}^{-1} + \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}}$$

where

$$\hat{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1\mathbf{M}_{0,1} \\ \vdots \\ \mathbf{H}_n\mathbf{M}_{0,n} \end{pmatrix} \qquad \hat{\mathbf{R}} = \begin{pmatrix} \mathbf{R}_0 & 0 & \cdots & 0 \\ 0 & \mathbf{R}_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{R}_n \end{pmatrix}$$

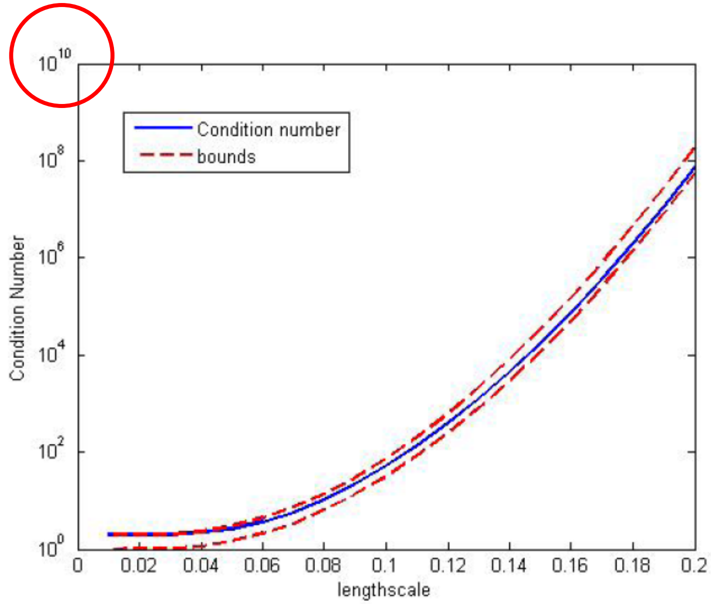
$$\mathbf{M}_{0,k} = \frac{\partial \mathcal{M}_{0,k}}{\partial \mathbf{x}}|_{\mathbf{x}_0} \qquad \mathbf{H}_k = \frac{\partial \mathcal{H}_k}{\partial \mathbf{x}}|_{\mathcal{M}_{0,k}(\mathbf{x}_0)}$$

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Conditioning of Hessian

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




Periodic Gaussian Exponential

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

Blue = condition number    Red = bounds

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Preconditioning the Hessian

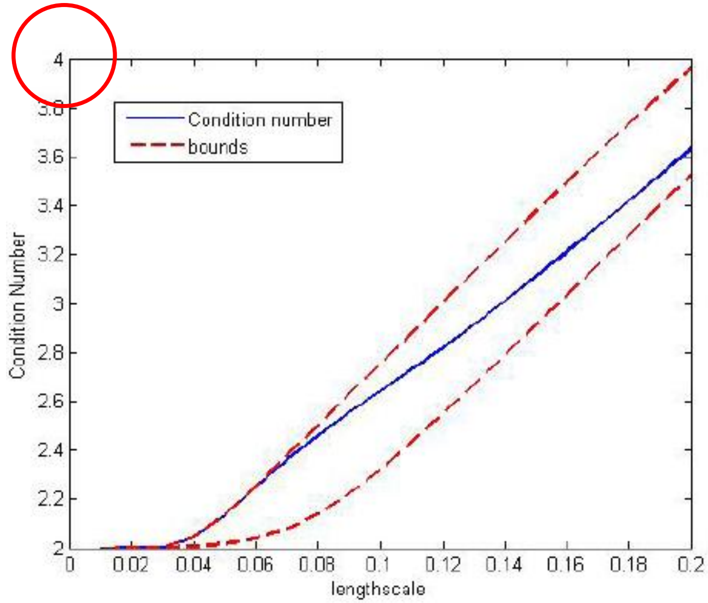
To improve conditioning transform to new variable :

- $\mathbf{z} = \mathrm{B}^{1/2} \; (\mathbf{x}_0 - \mathbf{x}_0^{\mathbf{b}})$
- Uncorrelated variables
- Equivalent to preconditioning by
- Hessian of transformed problem is

$$\mathbf{I} + \mathbf{B}^{1/2}\hat{\mathbf{H}}^{T}\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}}\mathbf{B}^{1/2}$$

# Preconditioned Hessian

Condition Number of Preconditioned Hessian vs Correlation Length Scale




Periodic Gaussian Exponential

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

Blue = condition number    Red = bounds

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Convergence Rates of CG in 4D – using SOAR Correlation Matrix

| | Iterations | |
|---|---|---|
| Lengthscale | Unprecond | Precond |
| 0.01 | 8 | 8 |
| 0.1 | 54 | 11 |
| 0.2 | 187 | 12 |
| 0.3 | 361 | 12 |

Haben et al, 2011

# 4.

# Model Error

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example - Effects of Model Error

Model: Linear Advection 1-D Upwind Scheme

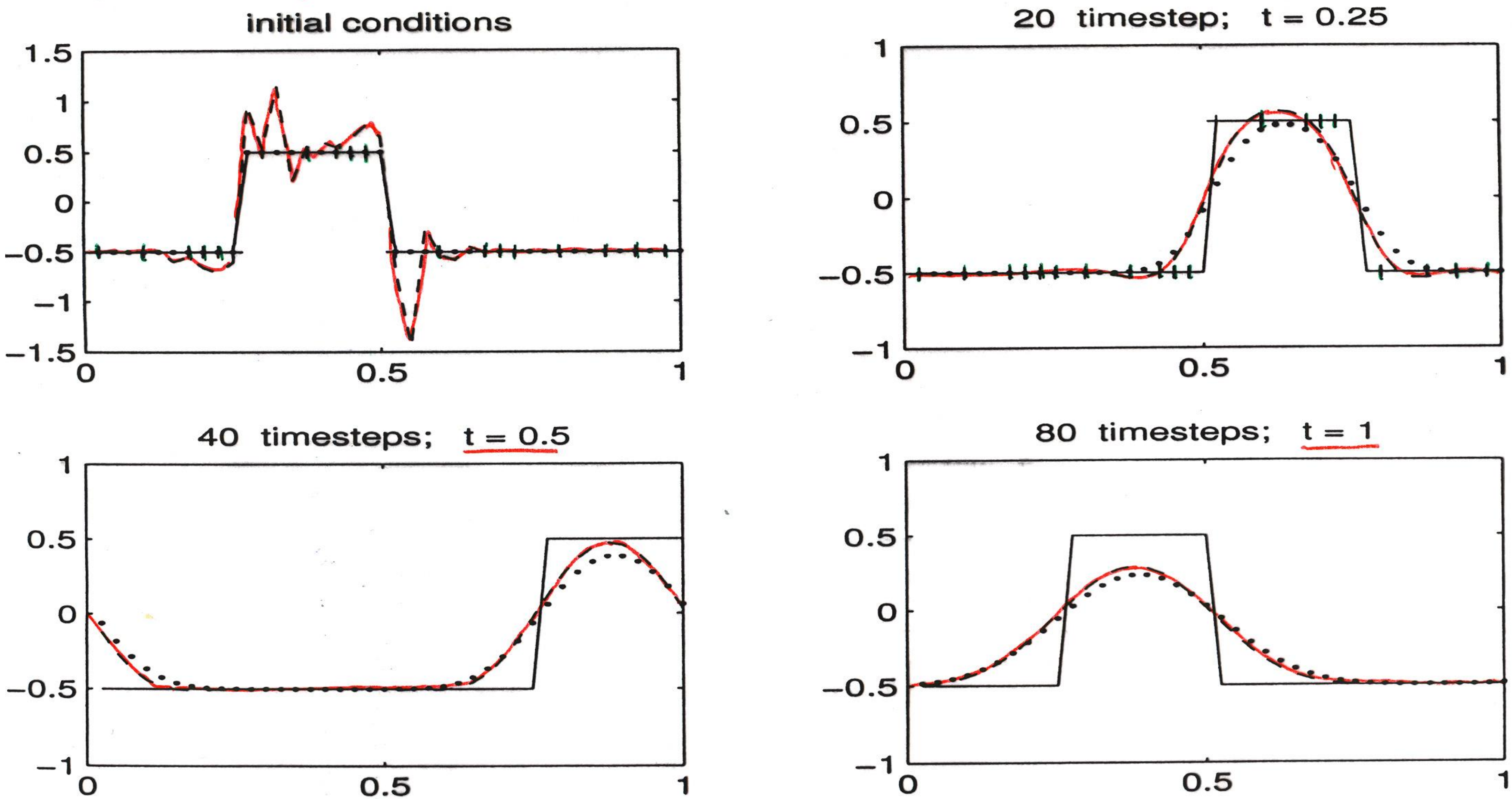
Initial conditions: Square wave

Boundary conditions: Periodic

Stepsize: $t = 1/80$ $x = 1/40$

Observations: Exact solution to $u_t + u_x = 0$ at 20 unevenly spaced points at each time step

# WAVE EQUATION – 4D VARIATIONAL ASSIMILATION




Solid = Truth, Dotted = Background, + = Observation, Red = With Assimilation

# System Equations

Prior: $$\mathbf{x}_b = \mathbf{x}_0 + \mathbf{e}_b$$

Model: $$\mathbf{x}_{i+1} = \mathcal{M}_i(\mathbf{x}_i) \ + \ \boldsymbol{\epsilon}_i \ ,$$

Observations: $$\mathbf{y}_i = H_i[\mathbf{x}_i] + \mathbf{e}_i$$

where
$$\mathcal{E}\{\mathbf{e}_b\} = 0 \qquad \mathcal{E}\{\mathbf{e}_b\mathbf{e}_b^T\} = \mathbf{B}$$
$$\mathcal{E}\{\mathbf{e}_i\} = 0 \qquad \mathcal{E}\{\mathbf{e}_i\mathbf{e}_i^T\} = \mathbf{R}_i$$
$$\mathcal{E}\{\boldsymbol{\epsilon}_i\} = 0 \qquad \mathcal{E}\{\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i^T\} = \mathbf{Q}_i$$

and errors are uncorrelated in time

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Variational Assimilation with Model Error

$$\min_{\mathbf{x}_0, \boldsymbol{\epsilon}_i} \mathcal{J} = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}_0^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) +$$

$$+ \frac{1}{2} \sum_{i=0}^{n} (H_i[\mathbf{x}_i] - \mathbf{y}_i)^T \mathbf{R}_i^{-1} (H_i[\mathbf{x}_i] - \mathbf{y}_i) +$$

$$+ \frac{1}{2} \sum_{k=0}^{N} \boldsymbol{\epsilon}_i^T \mathbf{Q}_i^{-1} \boldsymbol{\epsilon}_i \,,$$

subject to

$$\mathbf{x}_{i+1} = \mathcal{M}_i(\mathbf{x}_i) \; + \; \boldsymbol{\epsilon}_i \,,$$

$$i = 0, \ldots, n-1$$

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Variational Assimilation with Model Error

$$\min_{\mathbf{x}_0, \boldsymbol{\epsilon}_i} \mathcal{J} = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}_0^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) +$$

$$+\frac{1}{2} \sum_{i=0}^{n} (H_i[\mathbf{x}_i] - \mathbf{y}_i)^T \mathbf{R}_i^{-1} (H_i[\mathbf{x}_i] - \mathbf{y}_i) +$$

$$+\frac{1}{2} \sum_{k=0}^{N} \boldsymbol{\epsilon}_i^T \mathbf{Q}_i^{-1} \boldsymbol{\epsilon}_i$$

subject to

$$\mathbf{x}_{i+1} = \mathcal{M}_i(\mathbf{x}_i) + \boldsymbol{\epsilon}_i \,,$$

$$i = 0, \ldots, n-1$$

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Adjoint Method

Can solve using the adjoint technique as before. Now the adjoints are increased by an additional set of adjoint variables giving the sensitivity of the objective function $\mathcal{J}$ with respect to each of the model error variables $\epsilon_i$ .

At present this is too expensive for real time forecasting, but simplifications can be used.

# Augmented Method

One approach is to augment the dynamic equations with a simple model for the dynamics of the errors. Then we only need to estimate the initial error $\epsilon_0$ . The additional adjoints can then be calculated efficiently. If it is assumed that the error is a constant 'bias' error then the gradients can be found directly from the previous adjoint equations.

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Example - Effects of Model Error

Model: Linear Advection 1-D Upwind Scheme

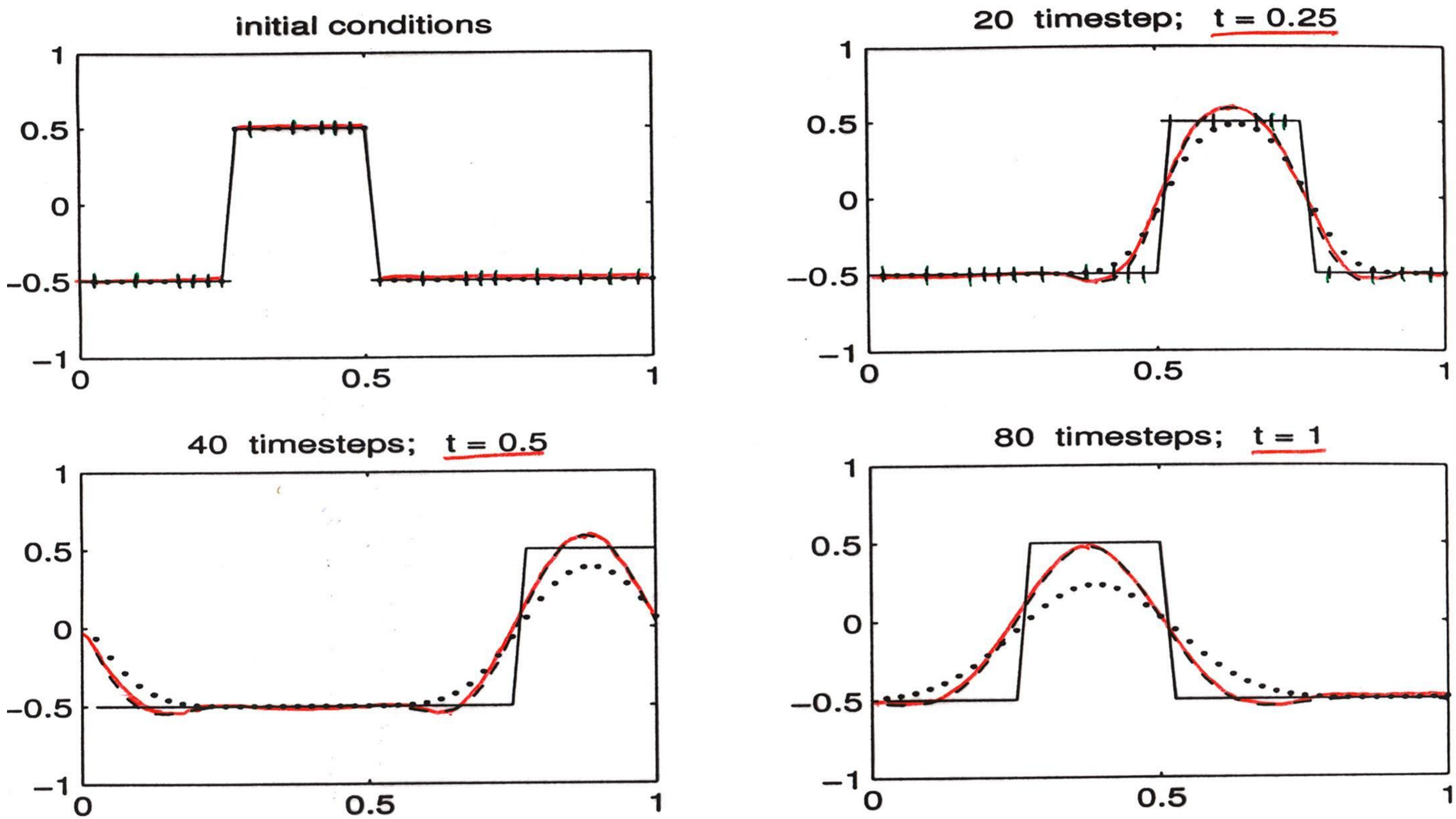
Initial conditions: Square wave

Boundary conditions: Periodic

Stepsize: t = 1/80 x = 1/40

Observations: Exact solution to $u_t + u_x = 0$ at 20 unevenly spaced points at each time step

# WAVE EQUATION - 4D VARIATIONAL ASSIMILATION




Solid = Truth, Dotted = Background, + = Observation, Red = With Assimilation

Evolving Error Model

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

# Application

# Simple assimilation

Model: FOAM global model: 1$^{o}$ horizontal resolution

Data assimilated: thermal profiles (including TAO moorings) and surface temperature (no salinity)

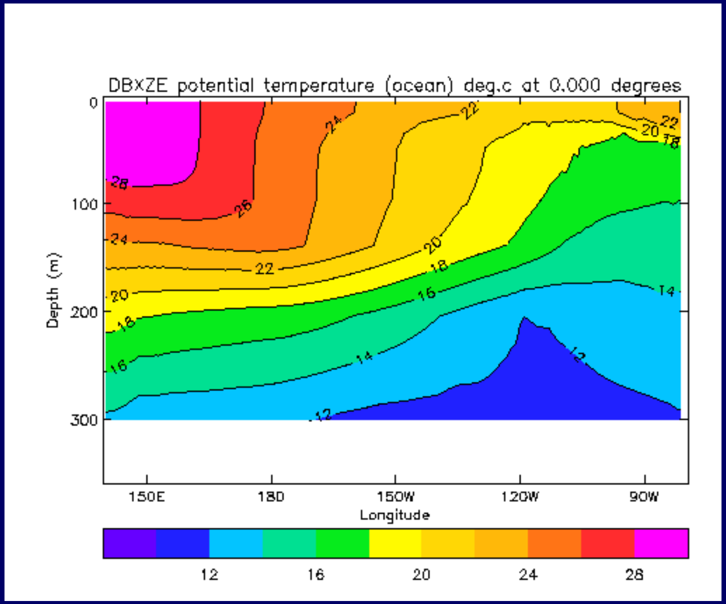
Assimilation method: analysis correction scheme

Surface fluxes: climatological wind stresses (Hellerman-Rosenstein) and heat fluxes

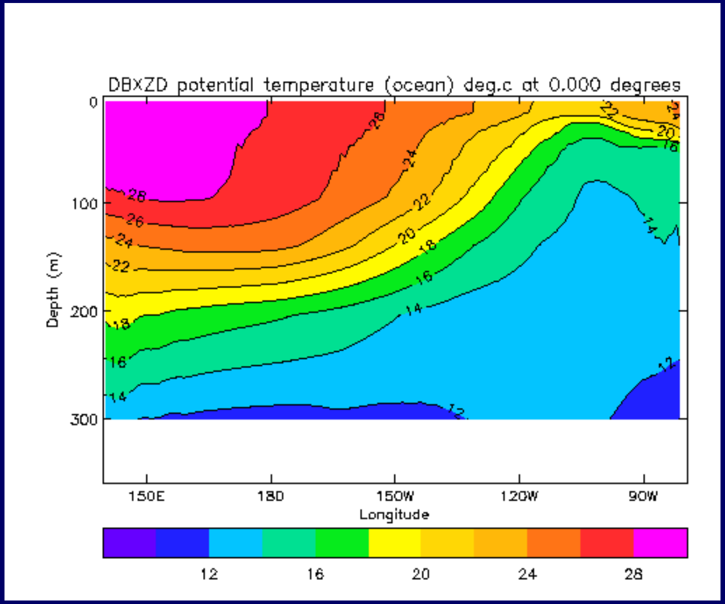
Period: 1995

# Effect of simple data assimilation

No assimilation




With assimilation



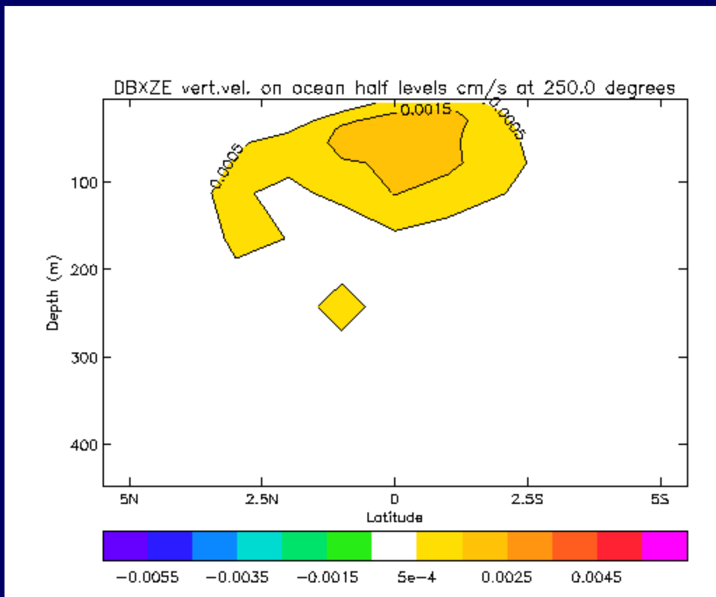

surface

300 m

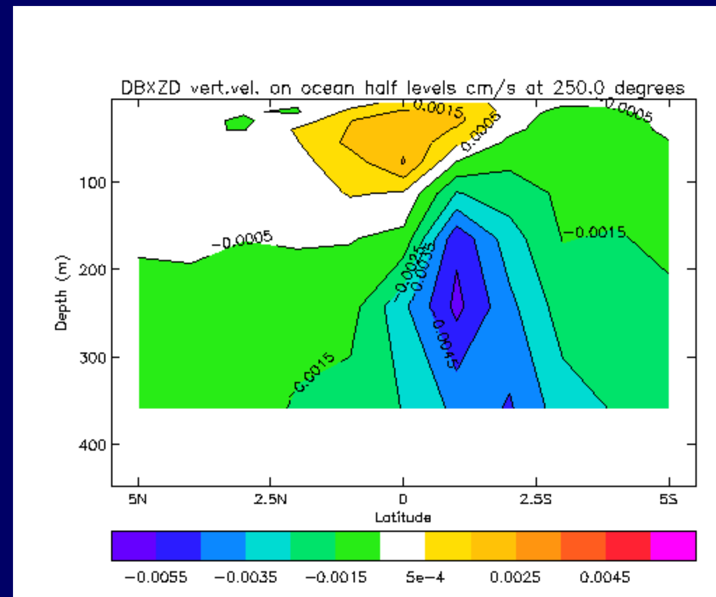
Annual mean potential temperatures ($^{o}$C) along the equatorial Pacific

# Effect of simple data assimilation

No assimilation




With assimilation




surface

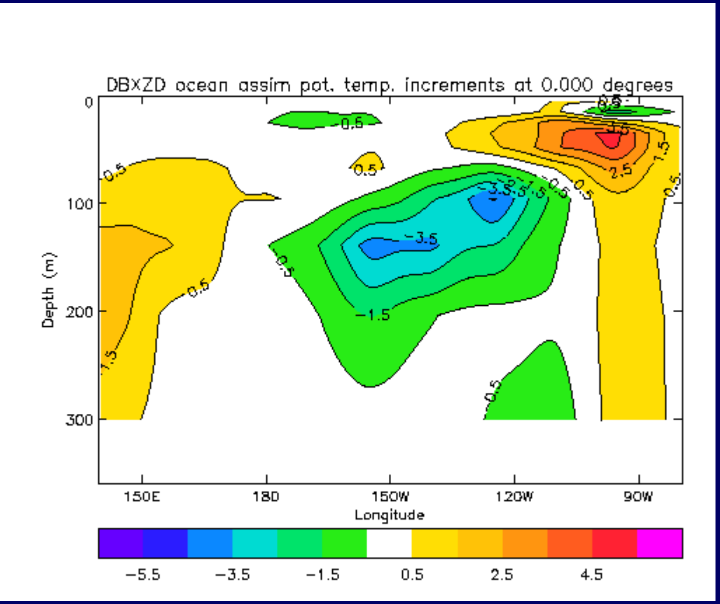
400 m

Annual mean vertical velocities at 110 $^oW$ (5 $^oN$ to 5 $^oS$) contour interval = $10^{-3}$ cm/s = 1 m/day

The University of Reading

Met Office

# Effect of simple data assimilation

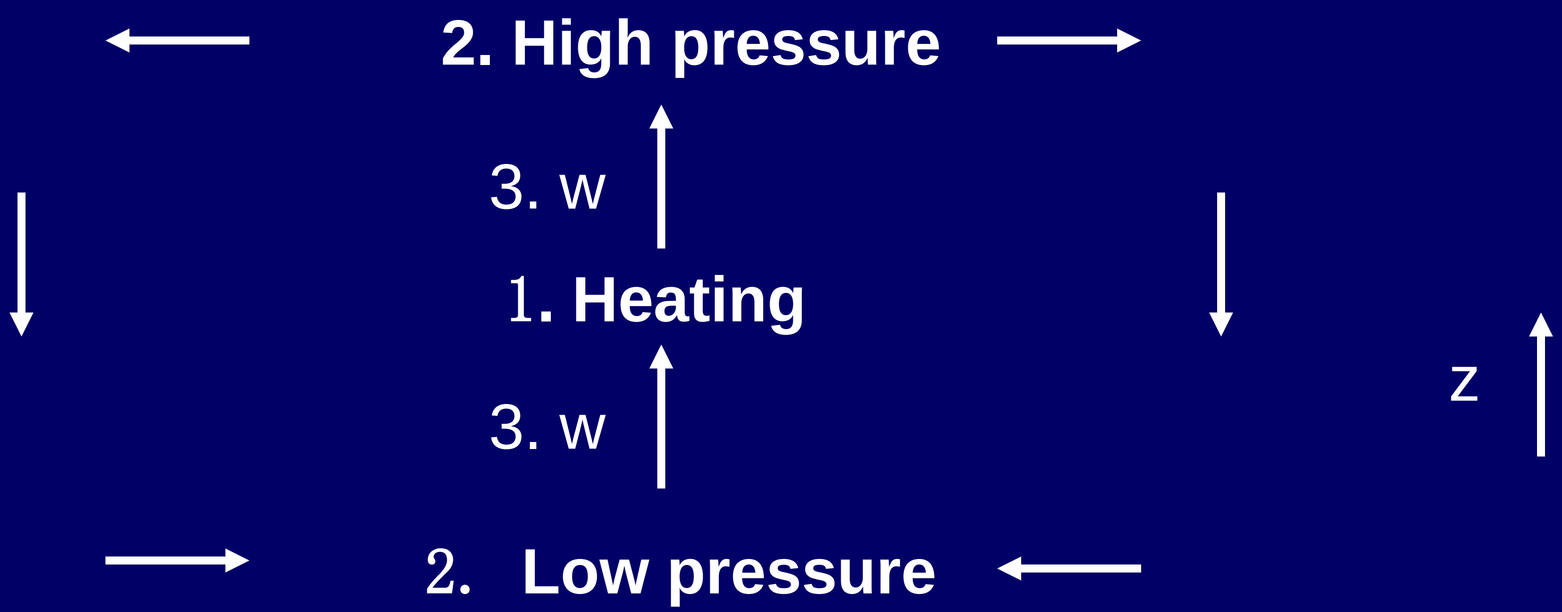



surface

300 m

Annual mean temperature increment from assimilation along the equatorial Pacific (contour interval =°C per month)

# Circulations induced by assimilation at equator where model is cold



# Central ideas

1. Where thermal increments of the same sign are repeatedly being made the balance of forces in the model is incorrect

2. Pressure fields in the opposite sense to those generated by the standard data assimilation increments need to be accumulated and applied

3. These increments are of small amplitude and large spatial scale so should not cause instabilities

# Control theory & augmented state

1. In control theory a state *x(t)* is evolved using a model *f* and observations *y*

$$x_t^f = f(x_{t-1}^a) \quad ; \quad x_t^a - x_t^f = K(y_t - h(x_t^f))$$

2. To control biases the state is extended/augmented by a bias, b(t), which is evolved and updated

$$x_t^f = f^x(x_{t-1}^a, b_{t-1}^a) \quad ; \quad b_t^f = f^b(x_{t-1}^a, b_{t-1}^a)$$

$$x_t^a - x_t^f = K^x(y_t - h(x_t^f)) \quad ; \quad b_t^a - b_t^f = K^b(y_t - h(x_t^f))$$

The University of Reading

Met Office

# Pressure correction method

1. The bias includes only scalar variables which contribute to the pressure field

2. For these variables $K^b = -\lambda\, K^x \quad ; \quad 0 < \lambda \; \square \; 1$

3. The model's pressure field is calculated using the sum of the bias and model scalar fields

4. The model for the evolution of the bias is:

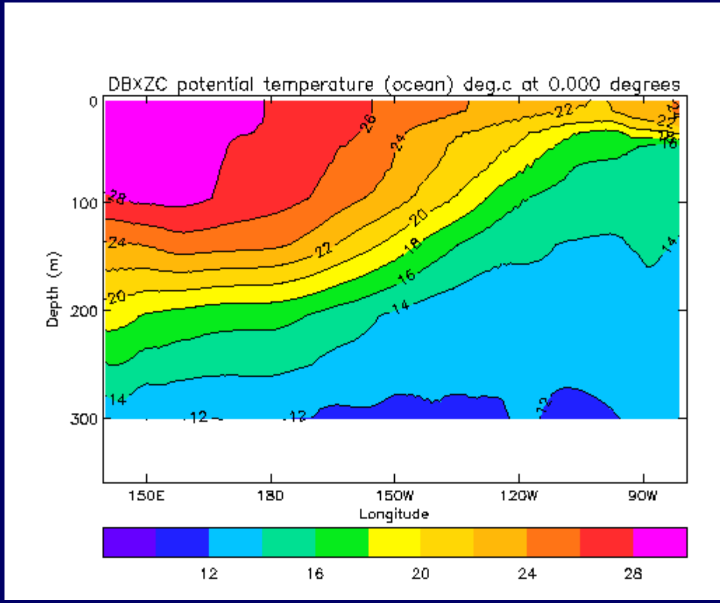
$$b_t^f = b_{t-1}^a$$

The University of Reading

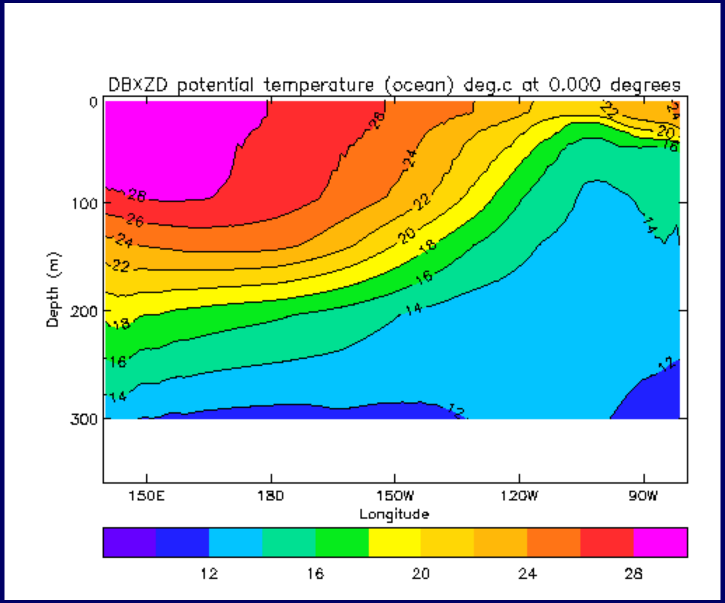
Met Office

# Repeat assimilation using pressure correction method with $\gamma = \varepsilon / 10$

Pressure correction




Original assimilation




surface

300 m

Annual mean potential temperatures (°C) along the equatorial Pacific



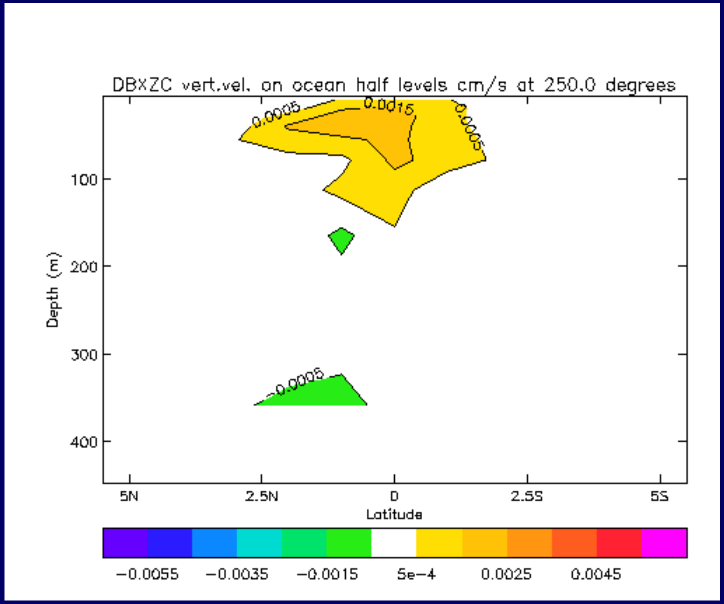
The University of Reading

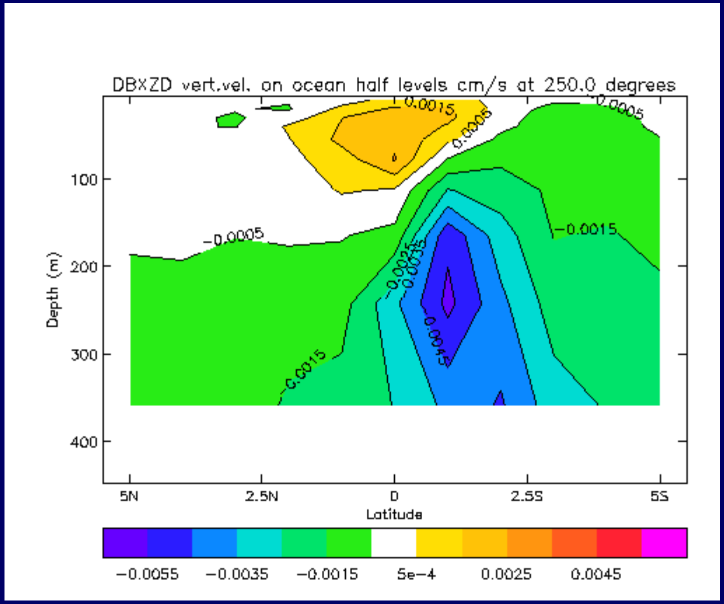
Met Office

# Repeat assimilation using pressure correction method with $\gamma = \varepsilon / 10$

Pressure correction

Original assimilation




surface

400 m

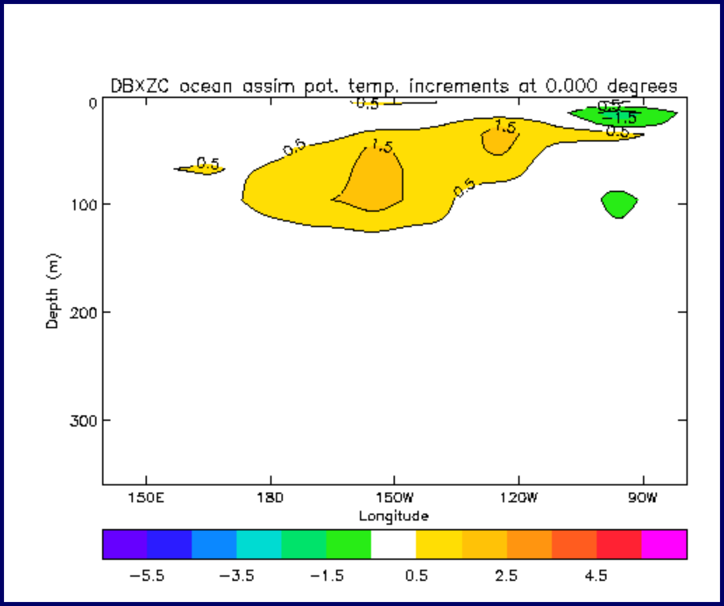
Annual mean vertical velocities
at 110 °W (5 °N to 5 °S) contour interval =
$10^{-3}$ cm/s = 1 m/day

The University of Reading

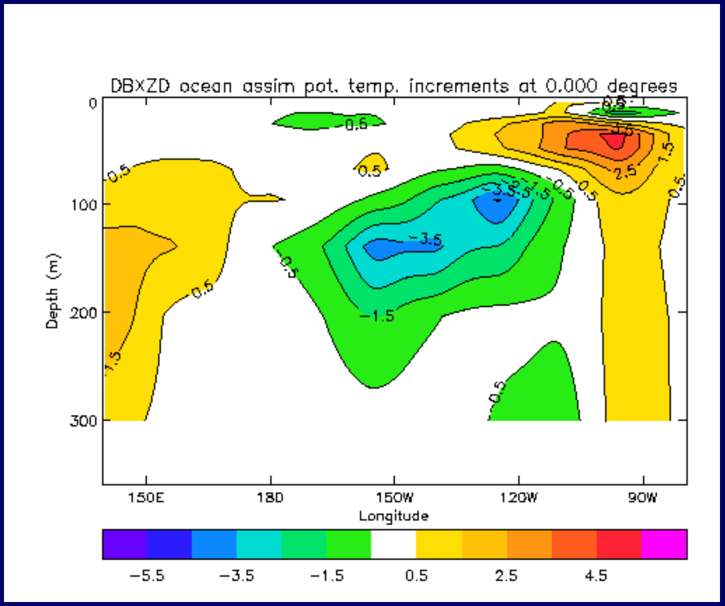
Met Office

# Repeat assimilation using pressure correction method with $\gamma = \varepsilon / 10$

Pressure correction




Original assimilation




surface

300 m

Annual mean temperature increment from assimilation along the equatorial Pacific (contour interval =ºC per month)

The University of Reading

Met Office

# Concluding summary

1. Simple assimilation of thermal data into an OGCM drives unrealistic motions within equatorial belt

2. A “pressure correction” method has been developed to control these motions using control theory ideas

3. It enables a better balanced assimilation of thermal data within the equatorial belt of OGCMs

4. There is a need to trial the method for seasonal forecasts

# 5.

# Conclusions

# Conclusions

4D Variational Data Assimilation is a powerful technique for estimating and predicting the states of very large environmental systems. It is used in major operational forecasting centres. The method can be adapted to a wide variety of problems and can be simplified by using approximations in the procedure.

Many challenges left!

University of Reading

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL