Variational Data Assimilation

Nancy Nichols
Data Assimilation Research Centre
University of Reading
Outline

• Introduction to data assimilation
• Variational data assimilation
• Incremental 4D variational assimilation
• Treatment of model error
• Conclusions
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
Example - State Estimation

Diffusion of temperature in a bar

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

$T^k = \text{temperature at grid point } z_k$

States of the system:

$T = \begin{pmatrix} T^1 \\ T^2 \\ \vdots \\ T^n \end{pmatrix}$
Example - State Estimation

Diffusion of temperature in a bar

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

\[ T_i^k = \text{temperature at grid point } z_k \text{ and time } t_i \]

States of the system at time \( t_i \):

\[ T_i = \begin{pmatrix} T_i^1 \\ T_i^2 \\ \vdots \\ T_i^n \end{pmatrix} \]
Example - Observations

Take observations at grid points at time $t_i$

$T_{i0}^0 = 0 \quad T_i^1 \quad T_i^2 \quad \ldots \quad T_i^k \quad \ldots \quad T_i^n \quad T_{i+1}^n = 0$

$y_i = \begin{pmatrix} T_{i2}^2 + e_{i2}^2 \\ T_{ik}^k + e_{ik}^k \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 & \cdots & 0 \end{pmatrix} T_i + e_i$

where $\mathcal{E}\{e_i\} = 0$, $\mathcal{E}\{e_i e_i^T\} = R_i$
Example - Observations

Take observations at grid points at times $t_i$

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

$$y_i = \begin{pmatrix} T_i^2 & e_i^2 \\ T_i^k & e_i^k \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 & \cdots & 0 \end{pmatrix} T_i + e_i$$

implies

$$y_i = HT_i + e_i$$

where

$$\mathbb{E}\{e_i\} = 0, \quad \mathbb{E}\{e_i e_i^T\} = R_i$$
Example - Observations

Take observations at grid points at times \( t_i \)

\[
T_{i0} = 0 \quad T_{i1} \quad T_{i2} \quad \ldots \quad T_{ik} \quad \ldots \quad T_{in} \quad T_{i,n+1} = 0
\]

\[
y_i = \begin{pmatrix}
T_{i2}^2 + e_i^2 \\
T_{ik}^2 + e_i^2
\end{pmatrix} = \begin{pmatrix}
0 & 1 & 0 & \cdots & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 1 & \cdots & 0
\end{pmatrix} \begin{pmatrix}
T_i \\
e_i
\end{pmatrix}
\]

implies

\[
y_i = HT_i + e_i
\]

where

\[
\mathcal{E}\{e_i\} = 0, \quad \mathcal{E}\{e_i e_i^T\} = R_i
\]
Example - Prior Estimate

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

\[ T_0^0 = 0 \quad T_0^1 \quad T_0^2 \quad \ldots \quad T_0^k \quad \ldots \quad T_0^n \quad T_0^{n+1} = 0 \]

\[
T_b = \begin{pmatrix}
T_0^1 + e_0^1 \\
T_0^2 + e_0^2 \\
\vdots \\
T_0^n + e_0^n
\end{pmatrix} = T_0 + e_b
\]

where

\[ \mathcal{E}\{e_b\} = 0 , \quad \mathcal{E}\{e_be_b^T\} = B \]
Example - Data Assimilation Problem

Prior: \[ T_b = T_0 + e_b \]

Observations: \[ y_0 = HT_0 + e_0 \]

**Question:** can we estimate the state of the system \( T_0 \) at \( t_0 \) from this information? How accurate is the estimate?
Example

Using these equations implies:

\[ y_0 - HT_0 = e_0 \]

= a set of linear equations for \( T_0 \).
Example

Using these equations

implies: \begin{align*}
T_b - T_0 &= e_b \\
\gamma_0 - HT_0 &= e_0
\end{align*}

= a set of linear equations for \( T_0 \).
Example - Solution

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

\[
\min_{T_0} \quad e_b^T B^{-1} e_b + e_0^T R_0^{-1} e_0
\]
Example - Solution

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

\[
\min_{T_0} \ e_b^T B^{-1} e_b + e_0^T R_0^{-1} e_0 = \\
\min_{T_0} (T_b - T_0)^T B^{-1} (T_b - T_0) + \\
+ (y_0 - HT_0)^T R_0^{-1} (y_0 - HT_0)
\]
Example - Solution

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

\[
\min_{T_0} \ e_b^T B^{-1} e_b + e_0^T R_0^{-1} e_0 =
\]

\[
\min_{T_0} \ (T_b - T_0)^T B^{-1} (T_b - T_0) + \]

\[
+ (y_0 - HT_0)^T R_0^{-1} (y_0 - HT_0)
\]

This gives \( T_0 \) with \text{minimum variance.}
Difference equation describing diffusion

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

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

where \( c \) is the diffusion coefficient
Example - Numerical Model

Difference equation describing diffusion

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

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
\[ \mu = \delta t / \delta z^2 \]
Example - Numerical Model

Write in matrix-vector form

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

\[ T_i^{k+1} = T_i^k + c \mu (T_i^{k+1} - 2T_i^k + T_i^{k-1}) \]

implies

\[ T_{i+1} = T_i + c \mu \mathbf{L} T_i \equiv \mathbf{M} T_i \]

where

\[ \mathbf{M} = \mathbf{I} + c \mu \mathbf{L}, \quad \mathbf{L} = \begin{pmatrix} -2 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -2 \end{pmatrix} \]
Example - Numerical Model

Write in matrix-vector form

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

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

implies

\[ T_{i+1} = T_i + c \mu L T_i \equiv M T_i \]

where

\[ M = I + c \mu L \]

\[ L = \begin{pmatrix} -2 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 1 & -2 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -2 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 1 & -2 \end{pmatrix} \]
Example - System Equations

Prior:
\[ T_b = T_0 + e_b \]

Model:
\[ T_{i+1} = M T_i \]

Observations:
\[ y_i = H T_i + e_i \]

where
\[ \mathcal{E}\{e_b\} = 0 \quad \mathcal{E}\{e_b e_b^T\} = B \]
\[ \mathcal{E}\{e_i\} = 0 \quad \mathcal{E}\{e_i e_i^T\} = R_i \]

and errors are uncorrelated in time
Example - Data Assimilation Problem

Prior: \[ T_b = T_0 + e_b \]

Model: \[ T_{i+1} = M T_i \]

Observations: \[ y_i = HT_i + e_i \]

**Question:** can we estimate the state of the system \( T_0 \) at \( t_0 \) from this information? How accurate is the estimate?
Example - YES

Using: \[ y_i = HT_i + e_i = HMT_{i-1} + e_i \]

implies:

\[ T_b - T_0 = e_b \]
\[ y_0 - HT_0 = e_0 \]
\[ y_1 - HMT_0 = e_1 \]
\[ y_2 - HM^2T_0 = e_2 \]
\[ \vdots \]
\[ y_n - HM^nT_0 = e_n \]

= a set of linear equations for \( T_0 \).
Example - Solution

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

$$\min_{T_0} e_b^T B^{-1} e_b + \sum_{i=0}^{n} e_i^T R_i^{-1} e_i =$$

$$\min_{T_0} (T_b - T_0)^T B^{-1} (T_b - T_0) +$$

$$+ \sum_{i=0}^{n} (y_i - H M^i T_0)^T R_i^{-1} (y_i - H M^i T_0)$$
Optimal Estimate

\[
\min_{T_0} J = \frac{1}{2} (T_b - T_0)^T B^{-1} (T_b - T_0) + \\
+ \frac{1}{2} \sum_{i=0}^{n} (y_i - HT_i)^T R_i^{-1} (y_i - HT_i)
\]

subject to

\[
T_{i+1} = MT_i, \quad i = 0, 1, \ldots n - 1
\]
Optimal Estimate

\[
\min_{T_0} J = \frac{1}{2} (T_b - T_0)^T B^{-1} (T_b - T_0) + \\
\frac{1}{2} \sum_{i=0}^{n} (y_i - HT_i)^T R_i^{-1} (y_i - HT_i)
\]

subject to

\[
T_{i+1} = M T_i , \quad i = 0, 1, \ldots n - 1
\]

Best Linear Unbiased Estimate
Optimal Unbiased Estimate

\[
\min_{T_0} J = \frac{1}{2} (T_b - T_0)^T B^{-1} (T_b - T_0) + \\
+ \frac{1}{2} \sum_{i=0}^{n} (y_i - HT_i)^T R_i^{-1} (y_i - HT_i)
\]

subject to

\[
T_{i+1} = M T_i , \quad i = 0, 1, ... n - 1
\]

Maximum A Posteriori Likelihood
Example - Application

Temperature diffusion with source term

\[ T^0 = 0 \quad T^1 \quad T^2 \quad \ldots \quad T^k \quad \ldots \quad T^n \quad T^{n+1} = 0 \]

Heat source

Model:

\[ T_{i+1} = M T_i + s_i \]

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

- \( \mathbf{x}_b \) - Background state (prior estimate)
- \( \mathbf{y}_i \) - Observations
- \( H_i \) - Observation operator
- \( \mathbf{B} \) - Background error covariance matrix
- \( \mathbf{R}_i \) - Observation error covariance matrix
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

\[
\min_{x_0} J = \frac{1}{2}(x_0 - x_b)^T B^{-1}(x_0 - x_b) + \\
+ \frac{1}{2} \sum_{i=0}^{n} (H[x_i] - y_i)^T R_i^{-1}(H[x_i] - y_i)
\]

subject to \[ x_{i+1} = M_i(x_i), \quad i = 0, \ldots, n - 1 \]

Solve iteratively by gradient optimization methods.
Use adjoint methods to find the gradients.

3DVar if \( n = 0 \) \quad 4DVar if \( n \geq 1 \)
Adjoint Model

Define the Lagrangian functional as

$$L = \mathcal{J} + \sum_{t=1}^{n-1} \lambda_{i+1}^T (x_{i+1} - M_i(x_i)).$$

Then the adjoint equations are

$$\lambda_n = 0$$

$$\lambda_i = M_i^T \lambda_{i+1} - H_i^T R_i^{-1}(H_i[x_i] - y_i)$$

where $M_i$ is the linearized dynamical model and $H_i$ is the linearized observation operator.
Adjoint Model

**Question** - What are the adjoints?

\( M_i \) is the Jacobian \( \frac{\partial M_i}{\partial x} \) of the linearized model operator and its **adjoint** is \( M_i^T \), known as the tangent linear model (TLM).

The **adjoint variables** \( \lambda_k \) measure the **sensitivity** of the objective function \( J \) to changes in the solutions \( x_k \) of the state equations.
Adjoint Model

The gradient of $J$ with respect to the initial condition $x_0$ is then given by

$$\nabla_0 J = -\lambda_0 + B^{-1}(x_0 - x_b)$$

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

- Estimate $x_0$
- Run the nonlinear model forward; find the ‘innovations’ $H[x_i] - y_i$ and evaluate the objective function $\mathcal{J}$
- Run the adjoint model backward to find $\lambda_0$ and evaluate the gradient $\nabla_0 J$
- Use a gradient nonlinear minimization method to find an improved estimate of $x_0$
- Repeat until required accuracy is reached.
Algorithm
3. Incremental 4D Variational Assimilation
Solve a sequence of linear least squares problems that approximate the nonlinear problem by iteration.
Incremental 4D-Var

Set $x_0^{(0)}$ (usually equal to background)

For $k = 0, \ldots, K$ find:

$$x_{i+1}^{(k)} = M_i(x_i^{(k)}), \quad i = 1, \ldots, n$$

Solve inner loop linear minimization problem:

$$\tilde{J}(k)[\delta x_0^{(k)}] = \frac{1}{2} (\delta x_0^{(k)} - [x^b - x_0^{(k)}])^T B_0^{-1} (\delta x_0^{(k)} - [x^b - x_0^{(k)}])$$

$$+ \frac{1}{2} \sum_{i=0}^n (H_i \delta x_i^{(k)} - d_i^{(k)})^T R_i^{-1} (H_i \delta x_i^{(k)} - d_i^{(k)})$$

subject to

$$\delta x_{i+1}^{(k)} = M_i x_i^{(k)} , \quad d_i = y_i - H_i [x_i^{(k)}]$$

Update:

$$x_0^{(k+1)} = x_0^{(k)} + \delta x_0^{(k)}$$
Algorithm

To find the optimal:

- Estimate $x_0$
- Run the nonlinear model forward to find $x_i$
- Estimate $\delta x_0$ and run the tangent linear model (TLM) forward to find $H_k \delta x_k - d_k$ and evaluate the linearized objective function
- Run the adjoint model backward using forcing terms $H_k \delta x_k - d_k$ to find $\lambda_0$ and evaluate the gradient of the linearized problem
- Use a gradient minimization method to find an improved estimate of $\delta x_0$
- Update $x_0$ by adding $\delta x_0$ 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.

Analysis

The analysis $x_a$ is the optimal solution to the assimilation problem and $x_a = x_0 + e_a$. The uncertainty is given by

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

where

$$\hat{H} = \begin{pmatrix} H_0 \\ H_1 M_0 \\ H_2 M_1 M_0 \\ \vdots \\ H_n M_{n-1} \ldots M_0 \end{pmatrix}$$

$$\hat{R} = \begin{pmatrix} R_0 & 0 & 0 & \ldots & 0 \\ 0 & R_1 & 0 & \ldots & 0 \\ 0 & 0 & R_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & R_n \end{pmatrix}$$
Conditioning of the Problem

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

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

where

$$\hat{H} = \begin{pmatrix} H_0 \\ H_1 M_{0,1} \\ \vdots \\ H_n M_{0,n} \end{pmatrix}$$

$$\hat{R} = \begin{pmatrix} R_0 & 0 & \cdots & 0 \\ 0 & R_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R_n \end{pmatrix}$$

$$M_{0,k} = \frac{\partial M_{0,k}}{\partial x} \bigg|_{x_0}$$

$$H_k = \frac{\partial H_k}{\partial x} \bigg|_{M_{0,k}(x_0)}$$
Conditioning of Hessian

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

Periodic Gaussian Exponential

\[ B_{ij} = \sigma_b^2 \exp \left( \frac{-r_{i,j}^2}{2L^2} \right) \]

Blue = condition number    Red = bounds
Preconditioning the Hessian

To improve conditioning transform to new variable:

- $z = B^{1/2} (x_0 - x_0^b)$
- Uncorrelated variables
- Equivalent to preconditioning by
- Hessian of transformed problem is

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

Condition Number of Preconditioned Hessian vs Correlation Length Scale

Periodic Gaussian Exponential

\[ B_{ij} = \sigma_b^2 \exp \left( -\frac{r_{i,j}^2}{2L^2} \right) \]

Blue = condition number  Red = bounds
Convergence Rates of CG in 4D – using SOAR Correlation Matrix

<table>
<thead>
<tr>
<th>Lengthscale</th>
<th>Unprecond</th>
<th>Precond</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>0.1</td>
<td>54</td>
<td>11</td>
</tr>
<tr>
<td>0.2</td>
<td>187</td>
<td>12</td>
</tr>
<tr>
<td>0.3</td>
<td>361</td>
<td>12</td>
</tr>
</tbody>
</table>

Haben et al, 2011
4. Model Error
Example - Effects of Model Error

Model: Linear Advection 1-D Upwind Scheme

Initial conditions: Square wave

Boundary conditions: Periodic

Steplsize: $t = \frac{1}{80} \quad x = \frac{1}{40}$

Observations: Exact solution to $u_t + u_x = 0$ at 20 unevenly spaced points at each time step
Solid = Truth,  Dotted = Background,  + = Observation,  Red = With Assimilation
System Equations

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

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

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

where \( \mathbb{E}\{\mathbf{e}_b\} = 0 \quad \mathbb{E}\{\mathbf{e}_b\mathbf{e}_b^T\} = \mathbf{B} \)

\( \mathbb{E}\{\mathbf{e}_i\} = 0 \quad \mathbb{E}\{\mathbf{e}_i\mathbf{e}_i^T\} = \mathbf{R}_i \)

\( \mathbb{E}\{\mathbf{\epsilon}_i\} = 0 \quad \mathbb{E}\{\mathbf{\epsilon}_i\mathbf{\epsilon}_i^T\} = \mathbf{Q}_i \)

and errors are uncorrelated in time
Variational Assimilation with Model Error

\[
\min_{x_0, \epsilon_i} J = \frac{1}{2} (x_0 - x_0^b)^T B_0^{-1} (x_0 - x_0^b) + \\
\frac{1}{2} \sum_{i=0}^{n} (H_i [x_i] - y_i)^T R_i^{-1} (H_i [x_i] - y_i) + \\
\frac{1}{2} \sum_{k=0}^{N} \epsilon_i^T Q_i^{-1} \epsilon_i ,
\]

subject to

\[
x_{i+1} = M_i (x_i) + \epsilon_i ,
\]

\[i = 0, \ldots, n - 1\]
Variational Assimilation with Model Error

\[
\min_{x_0, \epsilon_i} \mathcal{J} = \frac{1}{2} (x_0 - x^b_0)^T B_0^{-1} (x_0 - x^b_0) + \\
\frac{1}{2} \sum_{i=0}^{n} (H_i [x_i] - y_i)^T R_i^{-1} (H_i [x_i] - y_i) + \\
\frac{1}{2} \sum_{k=0}^{N} \epsilon_i^T Q_i^{-1} \epsilon_i
\]

subject to

\[
x_{i+1} = M_i(x_i) + \epsilon_i, \quad i = 0, \ldots, n - 1
\]
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.
Example - Effects of Model Error

Model: Linear Advection 1-D Upwind Scheme

Initial conditions: Square wave

Boundary conditions: Periodic

Stepsize: \( t = \frac{1}{80} \quad x = \frac{1}{40} \)

Observations: Exact solution to \( u_t + u_x = 0 \) at 20 unevenly spaced points at each time step
Solid = Truth, Dotted = Background, + = Observation, Red = With Assimilation

Evolving Error Model
Application
Simple assimilation

Model: FOAM global model: $1^\circ$ 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

Annual mean potential temperatures (°C) along the equatorial Pacific
Effect of simple data assimilation

No assimilation

With assimilation

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

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

1. Heating
2. Low pressure
3. w
4. z

2. High pressure
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 $y$

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

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

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

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

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

2. For these variables \[ K^b = -\lambda K^x \quad ; \quad 0 < \lambda \leq 1 \]

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:

\[ b^f_t = b^a_{t-1} \]
Repeat assimilation using pressure correction method with $\gamma = \varepsilon / 10$

Pressure correction

Original assimilation

Annual mean potential temperatures (°C) along the equatorial Pacific
Repeat assimilation using pressure correction method with $\gamma = \varepsilon / 10$

Pressure correction

Original assimilation

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

400 m
Repeat assimilation using pressure correction method with $\gamma = \epsilon / 10$

Pressure correction

Original assimilation

---

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