Mathematics of Data Assimilation-
Complex Systems in Numerical Weather Prediction

Melina Freitag

Department of Mathematical Sciences
University of Bath

BICS meeting 'The Maths of Complex Systems'
6th February 2008

Outline

What is Data Assimilation?

Loose definition

Estimation and prediction (analysis) of an unknown, true state by combining observations and system dynamics (model output).

What is Data Assimilation?

Loose definition

Estimation and prediction (analysis) of an unknown, true state by combining observations and system dynamics (model output).

Some examples

- Navigation (collect observations and produce velocity corrections)
- Geophysics (values of some model parameter must be obtained from the observed data)
- Medical imaging

What is Data Assimilation?

Loose definition

Estimation and prediction (analysis) of an unknown, true state by combining observations and system dynamics (model output).

Some examples

- Navigation (collect observations and produce velocity corrections)
- Geophysics (values of some model parameter must be obtained from the observed data)
- Medical imaging
- Numerical weather prediction

The atmosphere

- The atmosphere is a complex system!
- Mathematical modelling, observations and mathematical Data Assimilation help to understand this complex multi-scale system

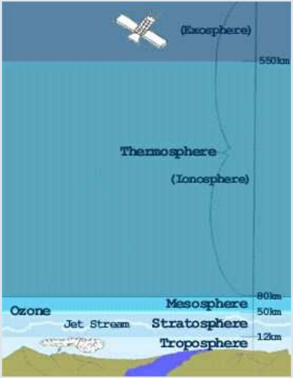


Figure: http://www.usoe.k12.ut.us

The atmosphere

- The atmosphere is a complex system!
- Mathematical modelling, observations and mathematical Data Assimilation help to understand this complex multi-scale system

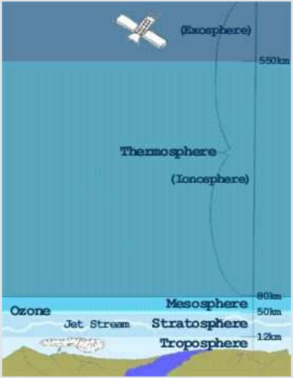


Figure: http://www.usoe.k12.ut.us

⇒"Aspects of Ionosphere modelling" Nathan (see talk later this afternoon)

The weather (NWP)

Data Assimilation in NWP

Estimate the state of the atmosphere $\mathbf{x}_i$ at a certain time/certain times i.

A priori information $\mathbf{x}^B$

- background state (usual previous forecast)

Data Assimilation in NWP

Estimate the state of the atmosphere $\mathbf{x}_i$ at a certain time/certain times i.

A priori information $\mathbf{x}^B$

- background state (usual previous forecast)

Observations $\mathbf{y}$

- Satellites
- Ships and buoys
- Surface stations
- Airplanes

Data Assimilation in NWP

Estimate the state of the atmosphere $\mathbf{x}_i$ at a certain time/certain times i.

A priori information $\mathbf{x}^B$

- background state (usual previous forecast)

Models

- a model how the atmosphere evolves in time (imperfect)

$$\mathbf{x}_{i+1} = M(\mathbf{x}_i)$$

Observations $\mathbf{y}$

- Satellites
- Ships and buoys
- Surface stations
- Airplanes

Data Assimilation in NWP

Estimate the state of the atmosphere $\mathbf{x}_i$ at a certain time/certain times i.

A priori information $\mathbf{x}^B$

- background state (usual previous forecast)

Models

- a model how the atmosphere evolves in time (imperfect)

$$\mathbf{x}_{i+1} = M(\mathbf{x}_i)$$

- a function linking model space and observation space (imperfect)

$$\mathbf{y}_i = H(\mathbf{x}_i)$$

Observations $\mathbf{y}$

- Satellites
- Ships and buoys
- Surface stations
- Airplanes

Data Assimilation in NWP

Estimate the state of the atmosphere $\mathbf{x}_i$ at a certain time/certain times i.

A priori information $\mathbf{x}^B$

- background state (usual previous forecast)

Models

- a model how the atmosphere evolves in time (imperfect)

$$\mathbf{x}_{i+1} = M(\mathbf{x}_i)$$

- a function linking model space and observation space (imperfect)

$$\mathbf{y}_i = H(\mathbf{x}_i)$$

Observations $\mathbf{y}$

- Satellites
- Ships and buoys
- Surface stations
- Airplanes

Assimilation algorithms

- used to find an (approximate) state of the atmosphere $\mathbf{x}_i$ at times i (usually $i = 0$)
- using this/these states a forecast for future states of the atmosphere can be obtained

Outline

Schematics of DA

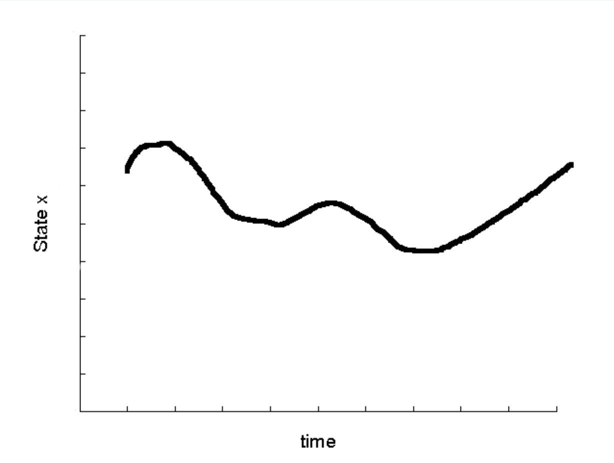


Figure: Background state $\mathbf{x}^B$

Schematics of DA

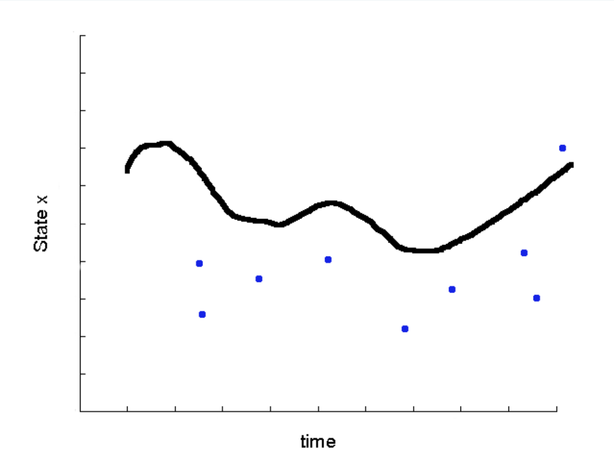


Figure: Observations $\mathbf{y}$

Schematics of DA

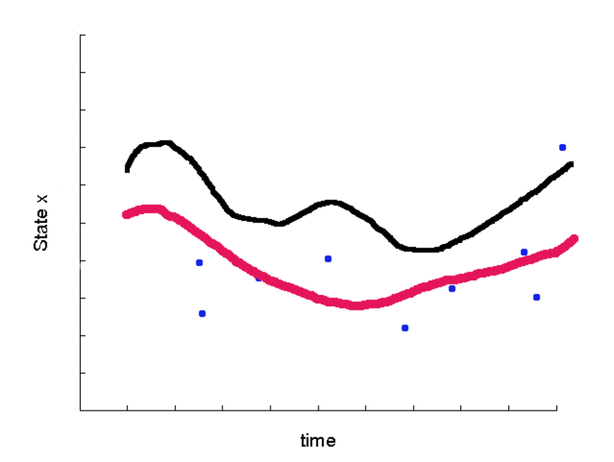


Figure: Analysis $\mathbf{x}^A$ (consistent with observations and model dynamics)

Data Assimilation in NWP

Underdeterminacy

- Size of the state vector $\mathbf{x}$: $432 \times 320 \times 50 \times 7 = \mathcal{O}(10^7)$
- Number of observations (size of $\mathbf{y}$): $\mathcal{O}(10^5 - 10^6)$
- Operator H (nonlinear!) maps from state space into observations space: $\mathbf{y} = H(\mathbf{x})$

Notation

- $\mathbf{x}^{\text{Truth}}$: True state
- $\mathbf{x}^{B}$: Background state (taken from previous forecast)
- $\mathbf{x}^{A}$: Analysis (estimation of the true state after the DA)

Outline

Data Assimilation in NWP

We are looking for the state of the atmosphere $\mathbf{x}_i$ at a certain time/certain times i.

Apriori information $\mathbf{x}^B$

- background state (usual previous forecast) has errors!

Models

- a model how the atmosphere evolves in time (imperfect)

$$\mathbf{x}_{i+1} = M(\mathbf{x}_i) + \text{error}$$

- a function linking model space and observation space (imperfect)

$$\mathbf{y}_i = H(\mathbf{x}_i) + \text{error}$$

Observations $\mathbf{y}$ has errors!

- Satellites
- Ships and buoys
- Surface stations
- Airplanes

Assimilation algorithms

- used to find an (approximate) state of the atmosphere $\mathbf{x}_i$ at times i (usually $i = 0$)
- using this/these states a forecast for future states of the atmosphere can be obtained

Error variables

Modelling the errors

- background error $\varepsilon^B = \mathbf{x}^B - \mathbf{x}^{\text{Truth}}$ of average $\overline{\varepsilon}^B$ and covariance

$$\mathbf{B} = \overline{(\varepsilon^B - \overline{\varepsilon}^B)(\varepsilon^B - \overline{\varepsilon}^B)^T}$$

- observation error $\varepsilon^O = \mathbf{y} - H(\mathbf{x}^{\text{Truth}})$ of average $\overline{\varepsilon}^O$ and covariance

$$\mathbf{R} = \overline{(\varepsilon^O - \overline{\varepsilon}^O)(\varepsilon^O - \overline{\varepsilon}^O)^T}$$

Error variables

Modelling the errors

- background error $\varepsilon^B = \mathbf{x}^B - \mathbf{x}^{\text{Truth}}$ of average $\overline{\varepsilon}^B$ and covariance

$$\mathbf{B} = \overline{(\varepsilon^B - \overline{\varepsilon}^B)(\varepsilon^B - \overline{\varepsilon}^B)^T}$$

- observation error $\varepsilon^O = \mathbf{y} - H(\mathbf{x}^{\text{Truth}})$ of average $\overline{\varepsilon}^O$ and covariance

$$\mathbf{R} = \overline{(\varepsilon^O - \overline{\varepsilon}^O)(\varepsilon^O - \overline{\varepsilon}^O)^T}$$

Assumptions

- Linearised observation operator: $H(\mathbf{x}) - H(\mathbf{x}^B) = \mathbf{H}(\mathbf{x} - \mathbf{x}^B)$
- Nontrivial errors: $\mathbf{B}$, $\mathbf{R}$ are positive definite
- Unbiased errors: $\overline{\mathbf{x}^B - \mathbf{x}^{\text{Truth}}} = \overline{\mathbf{y} - H(\mathbf{x}^{\text{Truth}})} = 0$
- Uncorrelated errors: $\overline{(\mathbf{x}^B - \mathbf{x}^{\text{Truth}})(\mathbf{y} - H(\mathbf{x}^{\text{Truth}}))^T} = 0$

Optimal least-squares estimater

Cost function minimisation (3D-Var)

Solution of the variational optimisation problem $\mathbf{x}^A = \arg\min J(\mathbf{x})$ where

$$
\begin{aligned}
J(\mathbf{x}) &= (\mathbf{x}-\mathbf{x}^B)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^B) + (\mathbf{y}-H(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}-H(\mathbf{x})) \\
&= J_B(\mathbf{x}) + J_O(\mathbf{x})
\end{aligned}
$$

- $\mathbf{B}^{-1}$ expensive!

Optimal least-squares estimater

Cost function minimisation (3D-Var)

Solution of the variational optimisation problem $\mathbf{x}^A = \arg\min J(\mathbf{x})$ where

$$\begin{aligned} J(\mathbf{x}) &= (\mathbf{x}-\mathbf{x}^B)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^B) + (\mathbf{y}-H(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}-H(\mathbf{x})) \\ &= J_B(\mathbf{x}) + J_O(\mathbf{x}) \end{aligned}$$

- $\mathbf{B}^{-1}$ expensive!

Interpolation equations

$$\mathbf{x}^A = \mathbf{x}^B + \mathbf{K}(\mathbf{y} - H(\mathbf{x}^B)), \quad \text{where}$$
$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \quad \mathbf{K}\ldots\text{gain matrix}$$

- expensive!

Four-dimensional variational assimilation (4D-Var)

Minimise the cost function

$$J(\mathbf{x}_0) = (\mathbf{x}_0 - \mathbf{x}^B)^T \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}^B) + \sum_{i=0}^{n} (\mathbf{y}_i - H_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1} (\mathbf{y}_i - H_i(\mathbf{x}_i))$$

subject to model dynamics $\mathbf{x}_i = M_{0\rightarrow i}\mathbf{x}_0$

Four-dimensional variational assimilation (4D-Var)

Minimise the cost function

$$J(\mathbf{x}_0) = (\mathbf{x}_0 - \mathbf{x}^B)^T \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}^B) + \sum_{i=0}^{n} (\mathbf{y}_i - H_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1} (\mathbf{y}_i - H_i(\mathbf{x}_i))$$

subject to model dynamics $\mathbf{x}_i = M_{0\to i}\mathbf{x}_0$

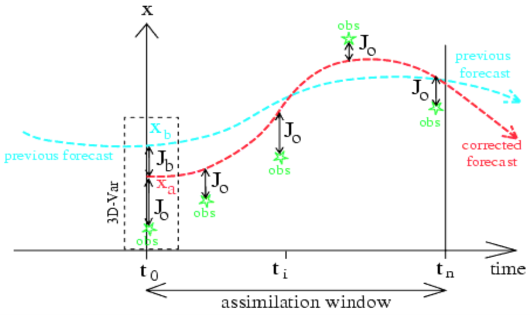
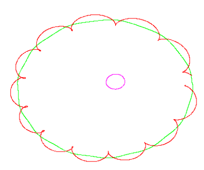


Figure: Copyright:ECMWF

Example - Three-Body Problem

Motion of three bodies in a plane, two position ($\mathbf{q}$) and two momentum ($\mathbf{p}$) coordinates for each body $\alpha = 1, 2, 3$



Equations of motion

$$
\begin{aligned}
H(\mathbf{q}, \mathbf{p}) &= \frac{1}{2} \sum_{\alpha} \frac{|\mathbf{p}_\alpha|^2}{m_\alpha} - \sum \sum_{\alpha < \beta} \frac{m_\alpha m_\beta}{|\mathbf{q}_\alpha - \mathbf{q}_\beta|} \\
\frac{d\mathbf{q}_\alpha}{dt} &= \frac{\partial H}{\partial \mathbf{p}_\alpha} \\
\frac{d\mathbf{p}_\alpha}{dt} &= -\frac{\partial H}{\partial \mathbf{q}_\alpha}
\end{aligned}
$$

Example - Three-Body problem

- solver: partitioned Runge-Kutta scheme with time step $h = 0.001$
- observations are taken as noise from the truth trajectory
- background is given from a perturbed initial condition
- assimilation window is taken 300 time steps
- minimisation of cost function J using a Gauss-Newton method

$$\nabla J(\mathbf{x}_0) = 0$$

$$\nabla\nabla J(\mathbf{x}_0^j)\Delta\mathbf{x}_0^j = -\nabla J(\mathbf{x}_0^j), \quad \mathbf{x}_0^{j+1} = \mathbf{x}_0^j + \Delta\mathbf{x}_0^j$$

- subsequent forecast is take 5000 time steps

Example- Three-Body problem

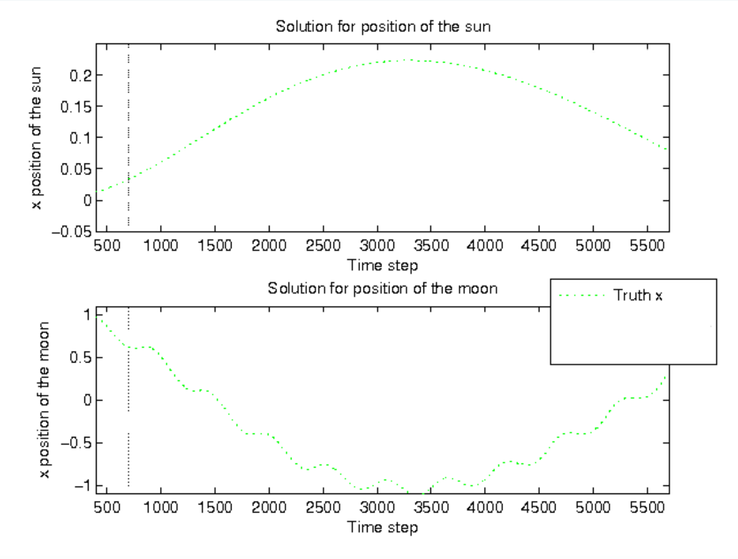


Figure: Truth trajectory

Example- Three-Body problem

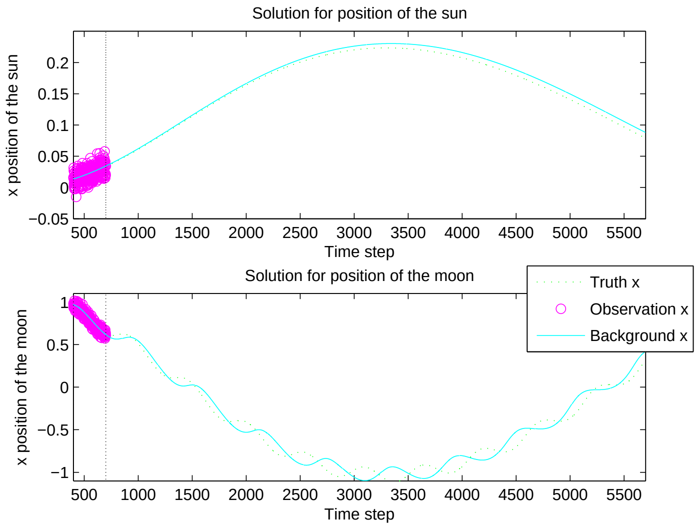


Figure: Truth trajectory with observations and background

Example- Three-Body problem

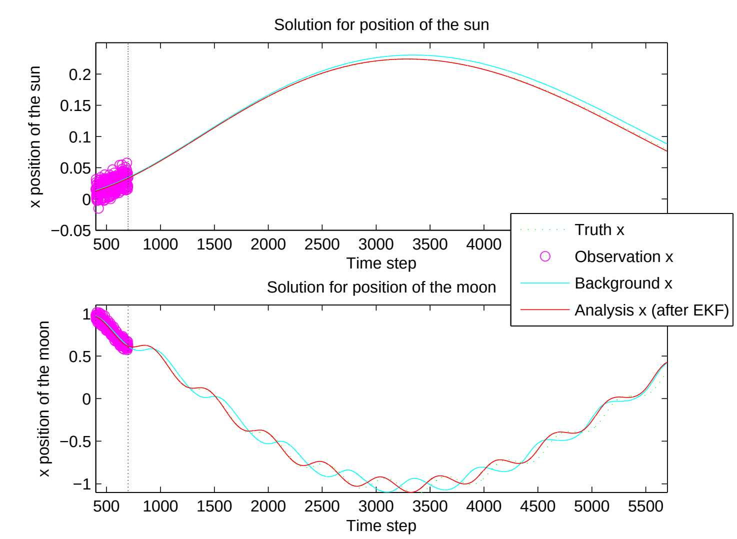


Figure: Analysis

Example- Three-Body problem

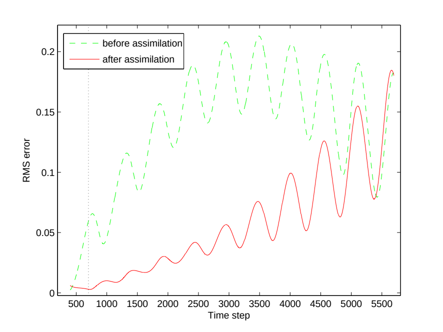


Figure: RMS error

Example- Three-Body problem

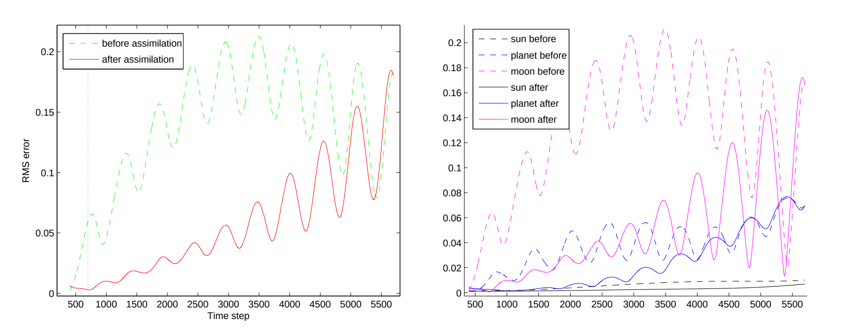


Figure: RMS error

The Kalman Filter Algorithm

- Sequential data assimilation
- covariance matrices are updated at each step $\mathbf{P}^F$, $\mathbf{P}^A$

State and error covariance forecast

$$
\begin{aligned}
\text{State forecast} \quad \mathbf{x}_{i+1}^F &= \mathbf{M}_{i+1,i}\mathbf{x}_i^A \\
\text{Error covariance forecast} \quad \mathbf{P}_{i+1}^F &= \mathbf{M}_{i+1,i}\mathbf{P}_i^A\mathbf{M}_{i+1,i}^T + \mathbf{Q}_i
\end{aligned}
$$

The Kalman Filter Algorithm

- Sequential data assimilation
- covariance matrices are updated at each step $\mathbf{P}^F$, $\mathbf{P}^A$

State and error covariance forecast

$$\begin{aligned} \text{State forecast} \quad \mathbf{x}_{i+1}^F &= \mathbf{M}_{i+1,i}\mathbf{x}_i^A \\ \text{Error covariance forecast} \quad \mathbf{P}_{i+1}^F &= \mathbf{M}_{i+1,i}\mathbf{P}_i^A\mathbf{M}_{i+1,i}^T + \mathbf{Q}_i \end{aligned}$$

State and error covariance analysis

$$\begin{aligned} \text{Kalman gain} \quad \mathbf{K}_i &= \mathbf{P}_i^F\mathbf{H}_i^T(\mathbf{H}_i\mathbf{P}_i^F\mathbf{H}_i^T + \mathbf{R}_i)^{-1} \\ \text{State analysis} \quad \mathbf{x}_i^A &= \mathbf{x}_i^F + \mathbf{K}_i(\mathbf{y}_i - \mathbf{H}_i\mathbf{x}_i^F) \\ \text{Error covariance of analysis} \quad \mathbf{P}_i^A &= (\mathbf{I} - \mathbf{K}_i\mathbf{H}_i)\mathbf{P}_i^F \end{aligned}$$

Example - Three-Body Problem

- same setup as before
- Compare using $\mathbf{B} = \mathbf{I}$ with using a flow-dependent matrix $\mathbf{B}$ which was generated by a Kalman Filter before the assimilation starts (see G. Inverarity (2007))

Example - Three-Body Problem

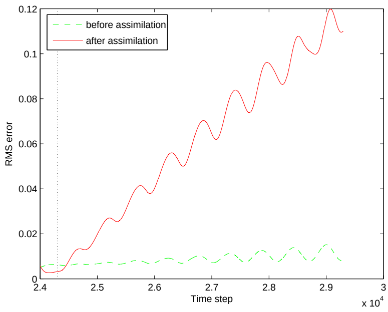


Figure: 4D-Var with $\mathbf{B} = \mathbf{I}$

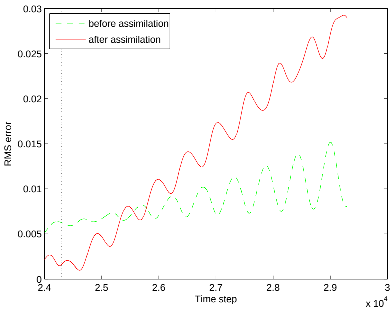


Figure: 4D-Var with $\mathbf{B} = \mathbf{P}^A$

Outline

Problems with Data Assimilation

- DA is computational very expensive, one cycle is much more expensive than the actual forecast (many applications of the tangent linear model and adjoint)

Problems with Data Assimilation

- DA is computational very expensive, one cycle is much more expensive than the actual forecast (many applications of the tangent linear model and adjoint)
- estimation of the $\mathbf{B}$-matrix is hard
  - in operational DA $\mathbf{B}$ is about $10^7 \times 10^7$
  - $\mathbf{B}$ should be flow-dependent but in practice often static
  - $\mathbf{B}$ needs to be modeled and diagonalised since $\mathbf{B}^{-1}$ too expensive to compute (”control variable transform”)

Problems with Data Assimilation

- DA is computational very expensive, one cycle is much more expensive than the actual forecast (many applications of the tangent linear model and adjoint)
- estimation of the $\mathbf{B}$-matrix is hard
  - in operational DA $\mathbf{B}$ is about $10^7 \times 10^7$
  - $\mathbf{B}$ should be flow-dependent but in practice often static
  - $\mathbf{B}$ needs to be modeled and diagonalised since $\mathbf{B}^{-1}$ too expensive to compute (”control variable transform”)
- many assumptions are not valid
  - errors non-Gaussian, data have biases
  - forward model operator M is not exact and also non-linear and system dynamics are chaotic
  - minimisation of the cost function needs close initial guess, small assimilation window

Problems with Data Assimilation

- DA is computational very expensive, one cycle is much more expensive than the actual forecast (many applications of the tangent linear model and adjoint)
- estimation of the **B**-matrix is hard
  - in operational DA **B** is about $10^7 \times 10^7$
  - **B** should be flow-dependent but in practice often static
  - **B** needs to be modeled and diagonalised since $\mathbf{B}^{-1}$ too expensive to compute ("control variable transform")
- many assumptions are not valid
  - errors non-Gaussian, data have biases
  - forward model operator M is not exact and also non-linear and system dynamics are chaotic
  - minimisation of the cost function needs close initial guess, small assimilation window
- model error not included

Outline

Model error and perturbation error

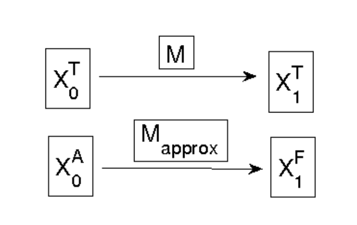


Figure: One assimilation window (6 hours)

$$
\begin{aligned}
\mathbf{x}_1^F - \mathbf{x}_1^{\text{Truth}} &= M_{\text{appr}}(\mathbf{x}_0^A) - M(\mathbf{x}_0^{\text{Truth}}) \\
&= \underbrace{M_{\text{appr}}(\mathbf{x}_0^A) - M_{\text{appr}}(\mathbf{x}_0^{\text{Truth}})}_{\text{Perturbation error}} + \underbrace{M_{\text{appr}}(\mathbf{x}_0^{\text{Truth}}) - M(\mathbf{x}_0^{\text{Truth}})}_{\text{Model error}}
\end{aligned}
$$

Model error and perturbation error

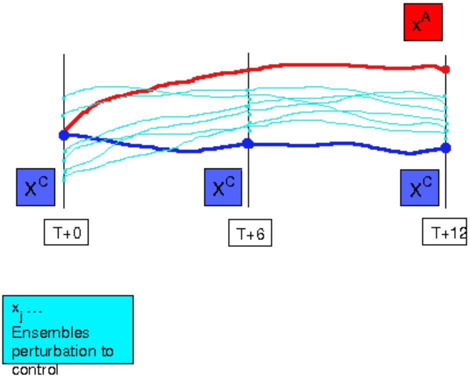


Figure: Model error and perturbation error

Model error

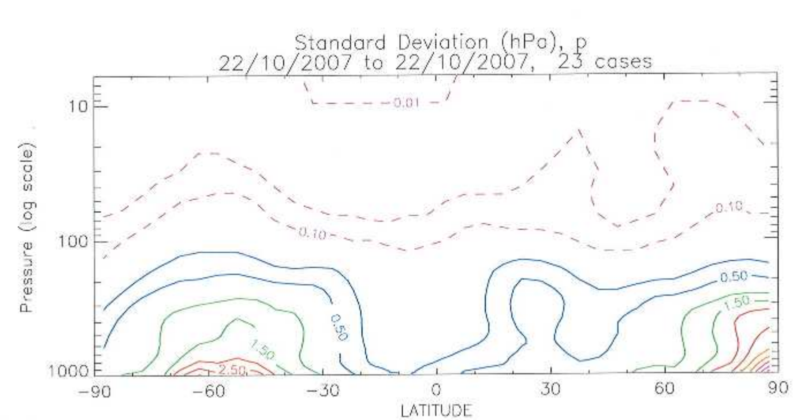


Figure: Perturbation error after 7 hours (Copyright: MetOffice)

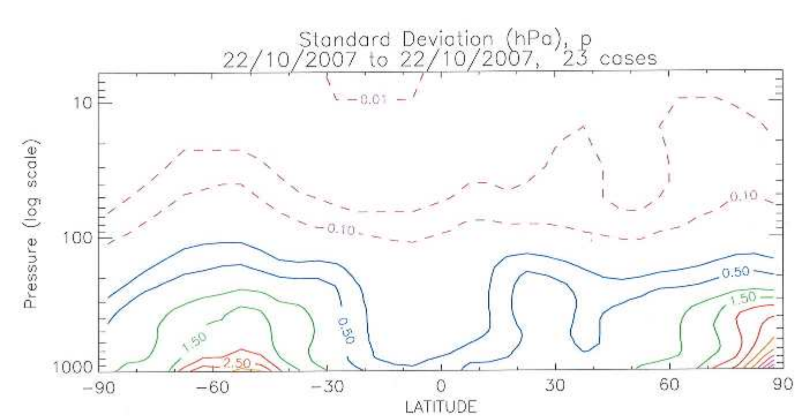


Figure: Perturbation error after 12 hours (Copyright: MetOffice)

Model error

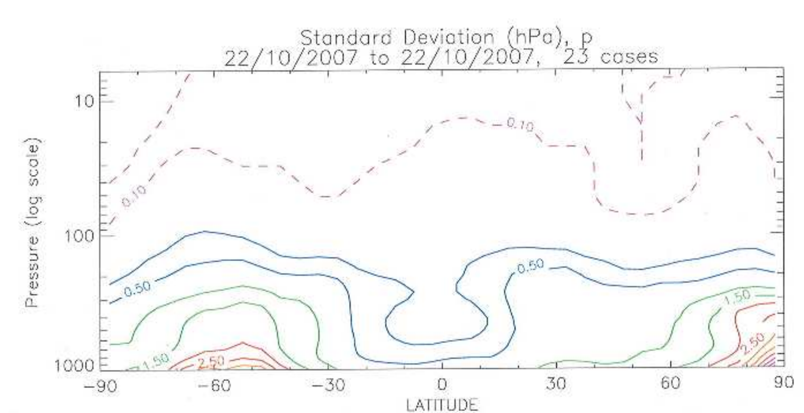


Figure: Model error after 12 hours (Copyright: MetOffice)

Plan

- design a simple chaotic model of reduced order (Lorenz model)
- include several time scales (to model the atmosphere)

Plan

- design a simple chaotic model of reduced order (Lorenz model)
- include several time scales (to model the atmosphere)
- include model error and analyse influence of this model error onto the DA scheme
- analyse the influence of the error made by the numerical approximation (part of the model error) on the error in the DA scheme
- investigate several assimilation algorithms and optimisation strategies to reduce existing errors

Plan

- design a simple chaotic model of reduced order (Lorenz model)
- include several time scales (to model the atmosphere)
- include model error and analyse influence of this model error onto the DA scheme
- analyse the influence of the error made by the numerical approximation (part of the model error) on the error in the DA scheme
- investigate several assimilation algorithms and optimisation strategies to reduce existing errors
- improve the representation of multi-scale behaviour in the atmosphere in existing DA methods
- improve the forecast of small scale features (like convective storms)

Plan

- design a simple chaotic model of reduced order (Lorenz model)
- include several time scales (to model the atmosphere)
- include model error and analyse influence of this model error onto the DA scheme
- analyse the influence of the error made by the numerical approximation (part of the model error) on the error in the DA scheme
- investigate several assimilation algorithms and optimisation strategies to reduce existing errors
- improve the representation of multi-scale behaviour in the atmosphere in existing DA methods
- improve the forecast of small scale features (like convective storms)
- Theme D: Numerical methods for multi-scale modelling