Weak Constraint 4D-Var

Yannick Trémolet

ECMWF Training Course - Data Assimilation

March 14, 2014
Outline

1 Introduction

2 Theoretical Maximum Likelihood Formulation

3 Practical 4D Variational Data Assimilation
   - Model Error Forcing Control Variable
   - 4D State Control Variable

4 Covariance Matrix

5 Results
   - Constant Model Error Forcing
   - Systematic Model Error
   - Is it model error?

6 Towards a long assimilation window

7 Summary and Discussion
Outline

1 Introduction

2 Theoretical Maximum Likelihood Formulation

3 Practical 4D Variational Data Assimilation
   • Model Error Forcing Control Variable
   • 4D State Control Variable

4 Covariance Matrix

5 Results
   • Constant Model Error Forcing
   • Systematic Model Error
   • Is it model error?

6 Towards a long assimilation window

7 Summary and Discussion
4D Variational Data Assimilation

4D-Var comprises the minimisation of:

\[ J(x) = \frac{1}{2} [\mathcal{H}(x) - y]^T R^{-1} [\mathcal{H}(x) - y] + \frac{1}{2} (x_0 - x_b)^T B^{-1} (x_0 - x_b) + \frac{1}{2} \mathcal{F}(x)^T C^{-1} \mathcal{F}(x) \]

- \(x\) is the 4D state of the atmosphere over the assimilation window.
- \(\mathcal{H}\) is a 4D observation operator, accounting for the time dimension.
- \(\mathcal{F}\) represents the remaining theoretical knowledge after background information has been accounted for (balance, DFI...).
- Control variable reduces to \(x_0\) using the hypothesis: \(x_i = M_i(x_{i-1})\).
- The solution is a trajectory of the model \(M\) even though it is not perfect...
A typical assumption in data assimilation is to ignore model error (bias and random).

The perfect model assumption limits the length of the analysis window that can be used to roughly 12 hours.

Model bias can affect assimilation of some observations (radiance data in the stratosphere).

In weak constraint 4D-Var, we define the model error as

$$\eta_i = x^t_i - M_i(x^t_{i-1}) \quad \text{for } i = 1, \ldots, n$$

and we allow it to be non-zero.

Note: $x^t$ represents the true atmospheric state which is of course not known.
Outline

1 Introduction

2 Theoretical Maximum Likelihood Formulation

3 Practical 4D Variational Data Assimilation
   - Model Error Forcing Control Variable
   - 4D State Control Variable

4 Covariance Matrix

5 Results
   - Constant Model Error Forcing
   - Systematic Model Error
   - Is it model error?

6 Towards a long assimilation window

7 Summary and Discussion
We can derive the weak constraint cost function using Bayes’ rule:

\[ p(x_0 \cdots x_n | x_b; y_0 \cdots y_n) = \frac{p(x_b; y_0 \cdots y_n | x_0 \cdots x_n) p(x_0 \cdots x_n)}{p(x_b; y_0 \cdots y_n)} \]

The denominator is independent of \( x_0 \cdots x_n \).

The term \( p(x_b; y_0 \cdots y_n | x_0 \cdots x_n) \) simplifies to:

\[ p(x_b | x_0) \prod_{i=0}^{n} p(y_i | x_i) \]

Hence

\[ p(x_0 \cdots x_n | x_b; y_0 \cdots y_n) \propto p(x_b | x_0) \left[ \prod_{i=0}^{n} p(y_i | x_i) \right] p(x_0 \cdots x_n) \]
\[ p(x_0 \cdots x_n | x_b; y_0 \cdots y_n) \propto p(x_b | x_0) \left[ \prod_{i=0}^{n} p(y_i | x_i) \right] p(x_0 \cdots x_n) \]

- Taking minus the logarithm gives the cost function:

\[ J(x_0 \cdots x_n) = -\log p(x_b | x_0) - \sum_{i=0}^{n} \log p(y_i | x_i) - \log p(x_0 \cdots x_n) \]

- The terms involving \( x_b \) and \( y_i \) are the background and observation terms of the strong constraint cost function.
- The final term is new. It represents the \textit{a priori} probability of the sequence of states \( x_0 \cdots x_n \).
Given the sequence of states $x_0 \cdots x_n$, we can calculate the corresponding model errors:

$$
\eta_i = x_i - M_i(x_{i-1}) \quad \text{for } i = 1, \ldots, n
$$

We can use our knowledge of the statistics of model error to define

$$
p(x_0 \cdots x_n) \equiv p(x_0; \eta_1 \cdots \eta_n)
$$

One possibility is to assume that model error is uncorrelated in time. In this case:

$$
p(x_0 \cdots x_n) \equiv p(x_0)p(\eta_1) \cdots p(\eta_n)
$$

If we take $p(x_0) = \text{const.}$ (all states equally likely), and $p(\eta_i)$ as Gaussian with covariance matrix $Q_i$, weak constraint 4D-Var adds the following term to the cost function:

$$
\frac{1}{2} \sum_{i=1}^{n} \eta_i^T Q_i^{-1} \eta_i
$$
For Gaussian, temporally-uncorrelated model error, the weak constraint 4D-Var cost function is:

\[
J(x) = \frac{1}{2} (x_0 - x_b)^T B^{-1} (x_0 - x_b) \\
+ \frac{1}{2} \sum_{i=0}^{n} [H_i(x_i) - y_i]^T R_i^{-1} [H_i(x_i) - y_i] \\
+ \frac{1}{2} \sum_{i=1}^{n} [x_i - M_i(x_{i-1})]^T Q_i^{-1} [x_i - M_i(x_{i-1})]
\]

- Do not reduce the control variable using the model and retain the 4D nature of the control variable.
- Account for the fact that the model contains some information but is not exact by adding a model error term to the cost function.
- The model \( M \) is not verified exactly: it is a weak constraint.
- If model error is correlated in time, the model error term contains additional cross-correlation blocks.
Outline

1 Introduction

2 Theoretical Maximum Likelihood Formulation

3 Practical 4D Variational Data Assimilation
   - Model Error Forcing Control Variable
   - 4D State Control Variable

4 Covariance Matrix

5 Results
   - Constant Model Error Forcing
   - Systematic Model Error
   - Is it model error?

6 Towards a long assimilation window

7 Summary and Discussion
Outline

1 Introduction

2 Theoretical Maximum Likelihood Formulation

3 Practical 4D Variational Data Assimilation
   - Model Error Forcing Control Variable
   - 4D State Control Variable

4 Covariance Matrix

5 Results
   - Constant Model Error Forcing
   - Systematic Model Error
   - Is it model error?

6 Towards a long assimilation window

7 Summary and Discussion
\[ J(x_0, \eta) = \frac{1}{2} \sum_{i=0}^{n} [\mathcal{H}(x_i) - y_i]^T R_i^{-1} [\mathcal{H}(x_i) - y_i] \]
\[ + \frac{1}{2} (x_0 - x_b)^T B^{-1} (x_0 - x_b) + \frac{1}{2} \eta^T Q^{-1} \eta \]

with \( x_i = M_i(x_{i-1}) + \eta_i \).

- \( \eta_i \) has the dimension of a 3D state,
- \( \eta_i \) represents the instantaneous model error,
- \( \eta_i \) is propagated by the model.
- All results shown later are for constant forcing over the length of one assimilation window, i.e. for correlated model error.
TL and AD models can be used with little modification,

- Information is propagated between observations and initial condition control variable by TL and AD models.
Outline

1 Introduction

2 Theoretical Maximum Likelihood Formulation

3 Practical 4D Variational Data Assimilation
   - Model Error Forcing Control Variable
   - 4D State Control Variable

4 Covariance Matrix

5 Results
   - Constant Model Error Forcing
   - Systematic Model Error
   - Is it model error?

6 Towards a long assimilation window

7 Summary and Discussion
Use $x = \{x_i\}_{i=0,...,n}$ as the control variable.

Nonlinear cost function:

$$J(x) = \frac{1}{2} (x_0 - x_b)^T B^{-1} (x_0 - x_b)$$

$$+ \frac{1}{2} \sum_{i=0}^{n} [H(x_i) - y_i]^T R_i^{-1} [H(x_i) - y_i]$$

$$+ \frac{1}{2} \sum_{i=1}^{n} [M(x_{i-1}) - x_i]^T Q_i^{-1} [M(x_{i-1}) - x_i]$$

In principle, the model is not needed to compute the $J_0$ term.

In practice, the control variable will be defined at regular intervals in the assimilation window and the model used to fill the gaps.
Model integrations within each time-step (or sub-window) are independent:
- Information is not propagated across sub-windows by TL/AD models,
- Natural parallel implementation.

Tangent linear and adjoint models:
- Can be used without modification,
- Propagate information between observations and control variable within each sub-window.

Several 4D-Var cycles are coupled and optimised together.
Outline

1 Introduction

2 Theoretical Maximum Likelihood Formulation

3 Practical 4D Variational Data Assimilation
   - Model Error Forcing Control Variable
   - 4D State Control Variable

4 Covariance Matrix

5 Results
   - Constant Model Error Forcing
   - Systematic Model Error
   - Is it model error?

6 Towards a long assimilation window

7 Summary and Discussion
Model Error Covariance Matrix

- An easy choice is $Q = \alpha B$.
- If $Q$ and $B$ are proportional, $\delta x_0$ and $\eta$ are constrained in the same directions, may be with different relative amplitudes.
- They both predominantly retrieve the same information.

- $B$ can be estimated from an ensemble of 4D-Var assimilations.
- Considering the forecasts run from the 4D-Var members:
  - At a given step, each model state is supposed to represent the same *true* atmospheric state,
  - The tendencies from each of these model states should represent possible evolutions of the atmosphere from that same *true* atmospheric state,
  - The differences between these tendencies can be interpreted as possible uncertainties in the model or realisations of *model error*.

- $Q$ can be estimated by applying the statistical model used for $B$ to tendencies instead of analysis increments.
- $Q$ has narrower correlations and smaller amplitudes than $B$. 
Most of the techniques developed to model $B$ can be re-used to model $Q$ (spectral technique, wavelets, filtering...).

Obtaining samples of model error is much more difficult:

- Currently, tendency differences between integrations of the members of an ensemble are used as a proxy for samples of model error.
- Use results from stochastic representation of uncertainties in EPS.
- Compare the covariances of $\eta$ produced by the current system with the matrix $Q$ being used.

It is possible to derive an estimate of $HQHT$ from cross-covariances between observation departures produced from pairs of analyses with different length windows (R. Todling).

- This produces a projection of $Q$ on a fixed observing network, not in the full matrix.

Characterising the statistical properties of model error is one of the main current problems in data assimilation and ensemble forecasting.
Outline

1 Introduction

2 Theoretical Maximum Likelihood Formulation

3 Practical 4D Variational Data Assimilation
   - Model Error Forcing Control Variable
   - 4D State Control Variable

4 Covariance Matrix

5 Results
   - Constant Model Error Forcing
   - Systematic Model Error
   - Is it model error?

6 Towards a long assimilation window

7 Summary and Discussion
Outline

1. Introduction
2. Theoretical Maximum Likelihood Formulation
3. Practical 4D Variational Data Assimilation
   - Model Error Forcing Control Variable
   - 4D State Control Variable
4. Covariance Matrix
5. Results
   - Constant Model Error Forcing
   - Systematic Model Error
   - Is it model error?
6. Towards a long assimilation window
7. Summary and Discussion
Results: Fit to observations

AMprofiler-windspeed Std Dev N.Amer

Background Departure

Analysis Departure

- Fit to observations is more uniform over the assimilation window.
- Background fit improved only at the start: error varies in time?
Mean Model Error Forcing

Temperature
Model level 11 (≈5hPa)
July 2004

Mean M.E. Forcing
M.E. Mean Increment
Control Mean Increment

Min = -0.05, Max = 0.10, RMS Global=0.02, N.hem=0.01, S.hem=0.03, Tropics=0.01
Outline

1 Introduction

2 Theoretical Maximum Likelihood Formulation

3 Practical 4D Variational Data Assimilation
   - Model Error Forcing Control Variable
   - 4D State Control Variable

4 Covariance Matrix

5 Results
   - Constant Model Error Forcing
   - Systematic Model Error
   - Is it model error?

6 Towards a long assimilation window

7 Summary and Discussion
Weak Constraint 4D-Var with Cycling Term

- Model error is not only random: there are biases.
- For random model error, the 4D-Var cost function is:

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

- For systematic model error, we might consider:

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

- Test case: can we address the model bias in the stratosphere?
Monthly Mean Model Error (Temperature (K/12h), July 2008)
Weak Constraint 4D-Var with Cycling Term

No Cycling Term

With Cycling Term

Monthly Mean Analysis Increment (Temperature (K), July 2008)
Weak Constraint 4D-Var with Cycling Term

AMSU-A Background departures, Channels 13 and 14

RADIANCES FROM METOP / AMSU-A CHANNEL 13
MEAN FIRST GUESS DEPARTURE (OBS-FG) (USED)
DATA PERIOD = 2008070100 - 2008073112
EXP = f57z
Min: -0.883688    Max: 0.90642    Mean: -0.084109

RADIANCES FROM METOP / AMSU-A CHANNEL 13
MEAN FIRST GUESS DEPARTURE (OBS-FG) (USED)
DATA PERIOD = 2008070100 - 2008073112
EXP = f8j2
Min: -0.592767    Max: 0.48862    Mean: -0.026685

RADIANCES FROM METOP / AMSU-A CHANNEL 14
MEAN FIRST GUESS DEPARTURE (OBS-FG) (USED)
DATA PERIOD = 2008070100 - 2008073112
EXP = f57z
Min: -1.6020      Max: 1.7330     Mean: 0.016017

RADIANCES FROM METOP / AMSU-A CHANNEL 14
MEAN FIRST GUESS DEPARTURE (OBS-FG) (USED)
DATA PERIOD = 2008070100 - 2008073112
EXP = f8j2
Min: -1.0986      Max: 0.973196   Mean: 0.099372

Control

Model Error
The short term forecast is improved with the model error cycling. Weak constraint 4D-Var can correct for seasonal bias (partially).
Outline

1 Introduction

2 Theoretical Maximum Likelihood Formulation

3 Practical 4D Variational Data Assimilation
   - Model Error Forcing Control Variable
   - 4D State Control Variable

4 Covariance Matrix

5 Results
   - Constant Model Error Forcing
   - Systematic Model Error
     - Is it model error?

6 Towards a long assimilation window

7 Summary and Discussion
The only significant source of observations in the box is aircraft data (Denver airport).

Removing aircraft data in the box eliminates the spurious forcing.
Aircraft Temperature Bias
USA ACARS and TEMP 00z temperature biases
Monthly averages for asc, desc and cruise level

Observations are biased.

Figure from Lars Isaksen
The mean temperature increment is smaller with weak constraint 4D-Var (Stratosphere only, June 1993).
The work on model error has helped identify other sources of error in the system (balance term).
Observation error or Model Error?

Observation error bias correction can compensate for model error.
Outline

1 Introduction

2 Theoretical Maximum Likelihood Formulation

3 Practical 4D Variational Data Assimilation
   - Model Error Forcing Control Variable
   - 4D State Control Variable

4 Covariance Matrix

5 Results
   - Constant Model Error Forcing
   - Systematic Model Error
   - Is it model error?

6 Towards a long assimilation window

7 Summary and Discussion
6-hour sub-windows:

- Better than 6-hour 4D-Var: two cycles are coupled through $J_q$,
- Better than 12-hour 4D-Var: more information (imperfect model), more control.

Single time-step sub-windows:

- Each assimilation problem is instantaneous = 3D-Var,
- Equivalent to a string of 3D-Var problems coupled together and solved as a single minimisation problem,
- Approximation can be extended to non instantaneous sub-windows.
(1) Weak constraint 4D-Var

This implementation is an approximation of weak constraint 4D-Var with an assimilation window that extends indefinitely in the past... which is equivalent to a Kalman smoother that has been running indefinitely.
Weak Constraint 4D-Var: Sliding Window

(1) Weak constraint 4D-Var

(2) Extended window
Weak Constraint 4D-Var: Sliding Window

1. Weak constraint 4D-Var

2. Extended window

3. Initial term has converged

This implementation is an approximation of weak constraint 4D-Var with an assimilation window that extends indefinitely in the past... which is equivalent to a Kalman smoother that has been running indefinitely.
Weak Constraint 4D-Var: Sliding Window

(1) Weak constraint 4D-Var

(2) Extended window

(3) Initial term has converged

(4) Assimilation window is moved forward
This implementation is an approximation of weak constraint 4D-Var with an assimilation window that extends indefinitely in the past...

...which is equivalent to a Kalman smoother that has been running indefinitely.
Outline

1 Introduction

2 Theoretical Maximum Likelihood Formulation

3 Practical 4D Variational Data Assimilation
   • Model Error Forcing Control Variable
   • 4D State Control Variable

4 Covariance Matrix

5 Results
   • Constant Model Error Forcing
   • Systematic Model Error
   • Is it model error?

6 Towards a long assimilation window

7 Summary and Discussion
In the forcing formulation of weak constraint 4D-Var:
- Background term to address systematic error,
- Interactions with variational observation bias correction,
- 24h assimilation window,
- Extend model error to the troposphere and to other variables (humidity).

Weak constraint 4D-Var with a 4D state control variable:
- Four dimensional problem with a coupling term between sub-windows and can be interpreted as a smoother over assimilation cycles.
- Can we extend the incremental formulation?

The two weak constraint 4D-Var approaches are mathematically equivalent (for linear problems) but lead to very different minimization problems.
- Can we combine the benefits of treating sub-windows in parallel with efficient minimization?
- 4D-Var scales well up to 1,000s of processors, can it scale to 100,000s of processors in the future?
Weak Constraint 4D-Var: Open Questions

- Weak Constraint 4D-Var allows the perfect model assumption to be removed and the use of longer assimilation windows.
  - How much benefit can we expect from long window 4D-Var?

- Weak Constraint 4D-Var requires knowledge of the statistical properties of model error (covariance matrix).
  - The forecast model is such an important component of the data assimilation system. It is surprising how little we know about its error characteristics.
  - How can we access realistic samples of model error? How can observations be used?
  - 4D-Var can handle time-correlated model error. What type of correlation model should be used?
  - Can we distinguish model error from observation bias or other errors? Is there a need to anchor the system?

- The statistical description of model error is one of the main current challenges in data assimilation.