Model bias in data assimilation

Yannick Trémolet

ECMWF
Shinfield Park, Reading RG2 9AX, United Kingdom

1 Introduction

Atmospheric models used in data assimilation are not perfect but errors in the model have generally been ignored in variational data assimilation. We present here a short description of weak constraints 4D-Var which explicitely takes model error into account and some preliminary results from its implementation in the ECMWF data assimilation system.

2 Variational Data Assimilation

Variational data assimilation comprises minimising a cost function which measures the discrepancy between the estimated atmospheric state x and the available observations y. In its most general form, as presented by Sasaki (1970), the cost function is written as:

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

where x_b is the background state, B and R are the background and observation error covariance matrices, $\mathcal{H}$ is the observation operator and $\mathcal{F}$ represents the remaining theoretical knowledge after background information has been accounted for, with error covariance matrix C_f. It is important to note that in this very general form, no hypothesis regarding x has been made.

A variety of choices of control vector x are possible. For example, x can be chosen to be the 3D state of the atmosphere at analysis time. This algorithm is known 3 dimensional variational data assimilation or 3D-Var.

Another possible choice is to consider x as the 4D state of the atmosphere during the assimilation window with the additional assumption that the forecast model is perfect. x is only a function of the initial condition x_0 and the size of the control variable reduces to a 3D state using the relation:

$$x_i = \mathcal{M}_i(x_{i-1}) = \mathcal{M}_{0,i}(x_0)$$

where $\mathcal{M}_i$ is the model representing the evolution of the atmospheric state from time t_{i-1} to time t_i and $\mathcal{M}_{0,i}$ represents the model integration from initial time to time t_i. $\mathcal{H}$ is a 4D operator, accounting for the time dimension which makes it possible to use observations at the correct time. 4D-Var implementations were made possible by the use of the adjoint technique introduced by Le Dimet and Talagrand (1986) which allows to evaluate the gradient of the cost function at a reasonable cost by one backward integration of the adjoint model. This approach is known as strong constraint 4D-Var or, simply, 4D-Var. It is used in most operational implementations of 4D-Var.

However, in practice, the model equations are known and solved inexactly, both because of our imperfect knowledge of the atmosphere and because of the discrete representation being used. They can be imposed as a weak constraint in the optimisation problem. In that case, $\mathcal{F}(x)$ can be defined as:

$$\mathcal{F}_i(x) = x_i - \mathcal{M}_i(x_{i-1}).$$

C_f is the model error covariance matrix denoted Q. The weak constraint 4D-Var cost function becomes:

$$\begin{aligned} J(x) &= \frac{1}{2}(x_0 - x_b)^T B^{-1}(x_0 - x_b) + \frac{1}{2}[\mathcal{H}(x) - y]^T R^{-1}[\mathcal{H}(x) - y] \\ &+ \frac{1}{2}\sum_{i=1}^{n}[x_i - \mathcal{M}_i(x_{i-1})]^T Q_i^{-1}[x_i - \mathcal{M}_i(x_{i-1})]. \end{aligned}$$

The problem defined above can be solved using x as the control variable but other choices are possible. Model error η, defined as the error at each time step by:

$$\eta_i = x_i - \mathcal{M}_i(x_{i-1})$$

can be used as the control variable. It is also possible to define model error as the error relative to the model solution integrated from the initial condition x_0. It is denoted β satisfying:

$$\beta_i = x_i - \mathcal{M}_{0,i}(x_0).$$

The choices of control vector $(x_i)_{i=0,\ldots,n}$, $(x_0, \{\eta_i\}_{i=1,\ldots,n})$ or $(x_0, \{\beta_i\}_{i=1,\ldots,n})$ are equivalent in the sense that they all represent the full 4D atmospheric state. A description of the properties of 4D-Var with these three possible control variables is given by Trémolet (2005).

3 Weak constraint 4D-Var in practice

3.1 Size of the problem

At the resolution currently used operationally at ECMWF, the size of the model error covariance matrix is such that it would take 9 million years to gather as many observations as there are parameters in Q. To gather meaningful statistics, it would require orders of magnitude more data. Estimating Q would also require being able to separate model error from other sources of error. There is not enough information available to define this problem without important simplifications. Assuming a simplifying model can be found for the model error covariance matrix, as is already the case for the background error covariance matrix, the number of parameters defining Q would be reduced and could be determined. However, the number of degrees of freedom in the control variable is still orders of magnitude larger than the number of observations available in each assimilation cycle. This problem is still highly under-determined and additional hypotheses are required for a proper estimation of x to be possible.

3.2 Model error control variable

Model error comes from several sources, some of which are constant (errors related to orography) while others can be almost periodic (errors related to diurnal cycle) or flow dependent (errors in physical processes). Discretisation and numerical errors may be more random. The simplest possible approximation is to consider that model error is constant over the assimilation period, or follows a predefined time dependency as proposed by Derber (1989) or Zupanski (1993). Other choices have been proposed such as the use of a Markov chain generated by a variable at a coarser resolution than η by Zupanski (1997). Another possible choice would be a Fourier series expansion which would represent the diurnal cycle as proposed by Griffith and Nichols (1998). It is also possible to include a vanishing term to represent model spin-up.

In the example presented below, model error is assumed constant in time over the assimilation window, keeping in mind that, other, less restrictive assumptions will have to be considered in the future.

In the case where a constant forcing η is chosen as the control variable, the model state x verifies:

$$x_i = \mathcal{M}_i(x_{i-1}) + \eta$$

and the cost function becomes:

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

This constant model error forcing makes weak constraint 4D-Var affordable and η is a good representation of systematic model error.

When β is chosen as the control variable, the approximation chosen is $\beta_i = \beta$, the model state x verifies:

$$x_i = \mathcal{M}_{0,i}(x_0) + \beta$$

and the cost function becomes:

$$J(x) = \frac{1}{2}(x_0 - x_b)^T B^{-1}(x_0 - x_b) + \frac{1}{2}[\mathcal{H}(\mathcal{M}(x_0) + \beta) - y]^T R^{-1}[\mathcal{H}(\mathcal{M}(x_0) + \beta) - y] + \frac{1}{2}\beta^T Q_\beta^{-1}\beta.$$

With this choice of control variable, the state of the model is not directly perturbed by β, thus avoiding spin-up and balance issues. β is a good representation of the global model bias against the mean of all observations. It does not correct for bias of one subset of observations against another subset of observations.

The choice of $(x_i)_{i=0,\ldots,n}$, $(x_0, \{\eta_i\}_{i=1,\ldots,n})$ or $(x_0, \{\beta_i\}_{i=1,\ldots,n})$ as a control variable are equivalent, but the choice of the approximation $\eta_i = \eta$ or $\beta_i = \beta$ are not. The table below summarises the possible choices of control variable and simplifying assumptions which can be made in variational data assimilation:

x_0	x_0	x_0	x_0	x_0
x_i	$\eta_i = \eta$	$\beta_i = \beta$	$\eta_i = 0$	$\eta_i = 0$
$x_i \approx \mathcal{M}_i(x_{i-1})$	$x_i = \mathcal{M}_i(x_{i-1}) + \eta$	$x_i = \mathcal{M}_{0,i}(x_0) + \beta$	$x_i = \mathcal{M}_i(x_{i-1})$	$x_i = x_0$
$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$
Weak constraint 4D-Var	Systematic model error	Model bias	4D-Var	3D-Var

3.3 Model Error Covariance Matrix

The model error covariance matrix can be approximated in many ways. Statistics for model error could be defined from the model's implementation, taking into account uncertainties in each aspect of the model (numerics and all physical processes). However, a more practical approach, used by most authors, is an approximation based on the background error covariance matrix B such as $Q = \alpha B$ where α is a scalar empirically determined.

In the incremental formulation of 4D-Var, the perturbation evolves according to:

$$\delta x_n = M_n \ldots M_1 \delta x_0 + \sum_{i=1}^{n} M_n \ldots M_{i+1}\eta_i.$$

This shows that δx_0 can be identified with η_0 (more details are given by Trémolet (2003)). Furthermore, the solution of the analysis equation satisfies:

$$\begin{aligned} \delta x_0 &= BH^T(R + HBH^T)^{-1}(\mathcal{H}(x_b) - y) \\ \eta &= QH^T(R + HQH^T)^{-1}(\mathcal{H}(x_b) - y) \end{aligned}$$

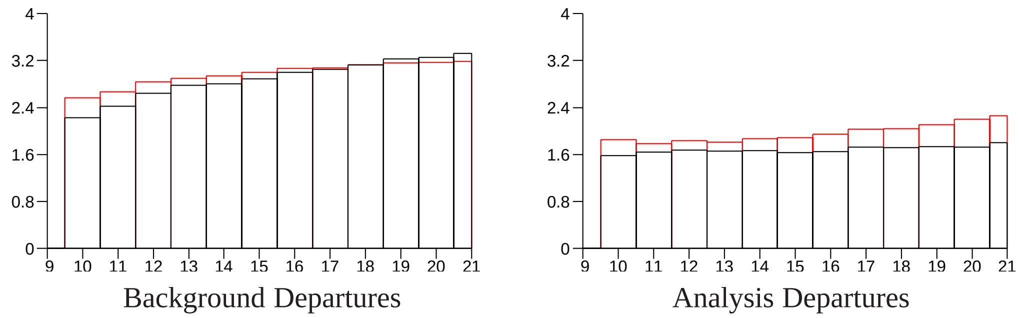

Figure 1: Profiler wind-speed standard deviation for North America for each hour bin in the assimilation window. Control is in red, model error experiment in black.

If Q and B are proportional, δx_0 and η are projected on the same subspace defined by the range of B. This means that δx_0 and η both predominantly retrieve the same information and the assumption that $Q = \alpha B$ is too limiting.

In the current ECMWF data assimilation system, the background error covariance matrix B is estimated from an ensemble of 4D-Var assimilations (Fisher (2003)). At a given time, the atmospheric states from the forecasts run from each or the 4D-Var members is supposed to be a representation of 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 and the differences between these tendencies can be interpreted as possible uncertainties in the model or realisations of model error. This gives a basis to estimate Q by applying the statistical model used for B to tendencies instead of analysis increments. This generates a covariance matrix which spans different directions than B. The vertical and horizontal correlations obtained from that method are generally narrower in both the vertical and horizontal dimensions and the average standard deviations have smaller amplitude than for B.

4 Experimental results

4.1 Identification of model error

Figure 1 shows that the analysis fit to observations is more uniform over the assimilation window when model error is taken into account as expected from Dee (1995). The fit of the background to observations is improved only at the start of the window which raises the question of the validity of the constant forcing outside the assimilation window.

Figure 2 shows the mean model error forcing and the mean initial condition increments for the month of July 2004. Taking model error into account greatly reduces the initial condition increment over Antarctica. This is confirmed by the mean observation departure for AMSU-A on NOAA-16 for the same period shown on figure 3. The zonal mean of the initial condition increment also shows that the oscillations in the polar temperature increment in the stratosphere have reduced (figure 4). Furthermore, the bias with respect to AMSU-A is more uniform over both hemispheres and, in the southern hemisphere, the background departures standard deviation is reduced and more data is used when model error is taken into account. This is confirmed by the statistics of the fit to radiosonde data over Antarctica (not shown) which show that the vertical oscillations in bias are reduced and standard deviation is reduced above 50 hPa for both the background and analysis. The temperature profiles of the analysis, initial condition increment and model error forcing for the South Pole are are shown on figure 5. They confirm that the oscillations have reduced and show that the vertical structure of the initial

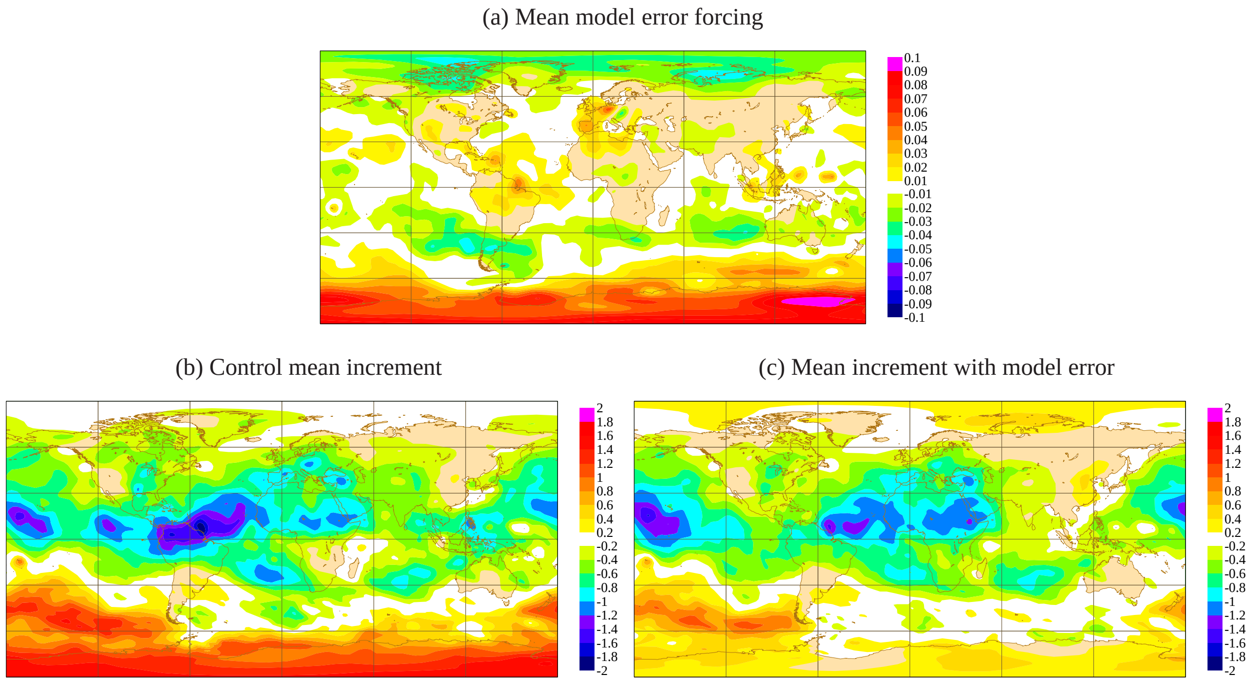

Figure 2: Average temperature model error forcing and initial condition increments for the month of July 2004 at model level 11 (≈5hPa).

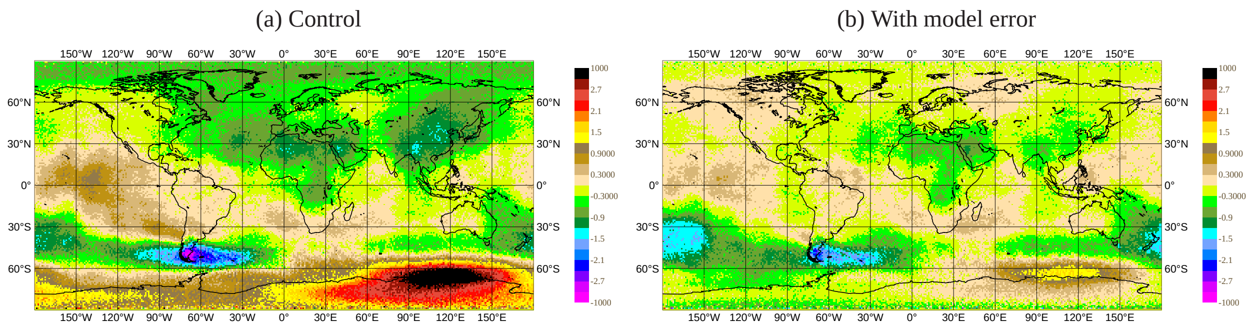

Figure 3: NOAA-16 AMSU-A channel 14 mean first guess departures for July 2004, in strong and weak constraint 4D-Var.

condition increments and of the forcing are very different. This gives the analysis more freedom to fit the data without causing spurious oscillations and shows the importance of having different structure functions for the initial condition and model error control variables.

These results show that the systematic model error in the winter stratosphere has been well captured and corrected by the model error forcing term.

4.2 Model bias and observation bias

Figure 6a shows the mean temperature forcing at the lowest model level over North America. Several bulls-eyes are visible on the map. Observation statistics show that, with model error, the bias for aircraft low level

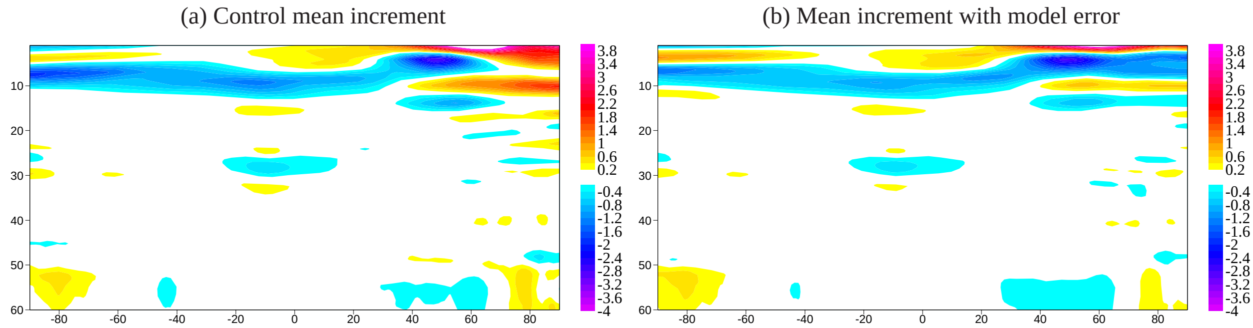

Figure 4: Zonal mean of temperature increments for the month of July 2004.

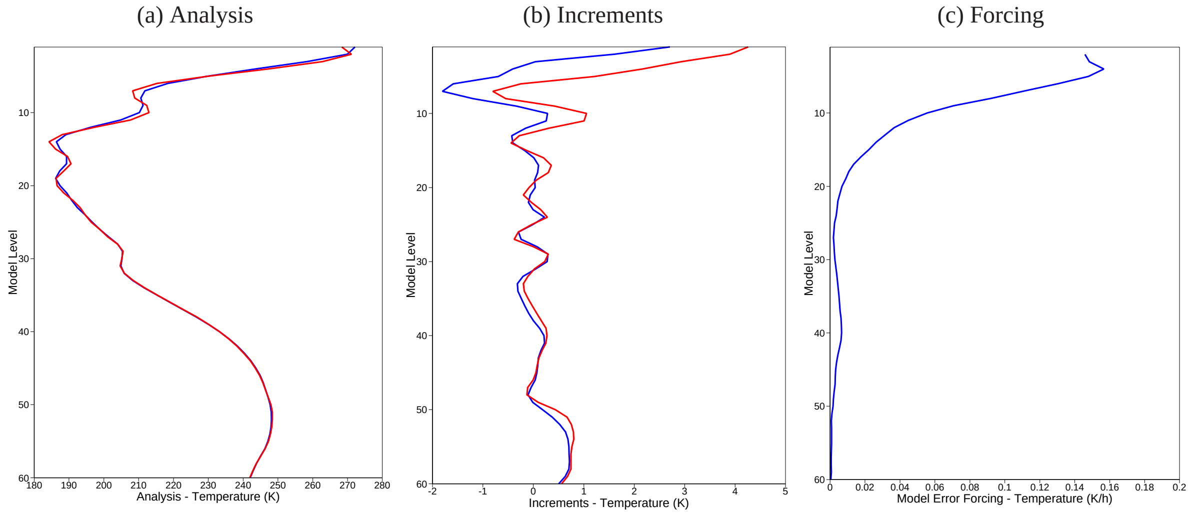

Figure 5: Average temperature profiles for the South Pole area on 05/07/2004.

temperature observations was reduced and removing aircraft data in the box shown in the figure removes the bulls-eye (figure 6b). The pattern in the forcing was generated by the observations from aircrafts near Denver airport. One characteristic of that type of data is that it is very frequent in time as aircraft traffic is very dense around major airports. In strong constraint 4D-Var, the initial condition increment generated by such observation departure pattern will be very diffuse since it is the evolved increment, over a period of time, that has to match the data. With the constant model error forcing term used in this experiment, it is very easy to fit these observations by adding a constant forcing at the airport location. In that case, it seems that the model error forcing term has captured observation bias rather than model error. An independent study (Lars Isaksen, personal communication) as shown that aircraft temperature observations are biased in the ascending and descending periods with respect to the cruise level observations as shown in figure7. This example shows that, in addition to capturing model bias, weak constraint 4D-Var can capture observation bias.

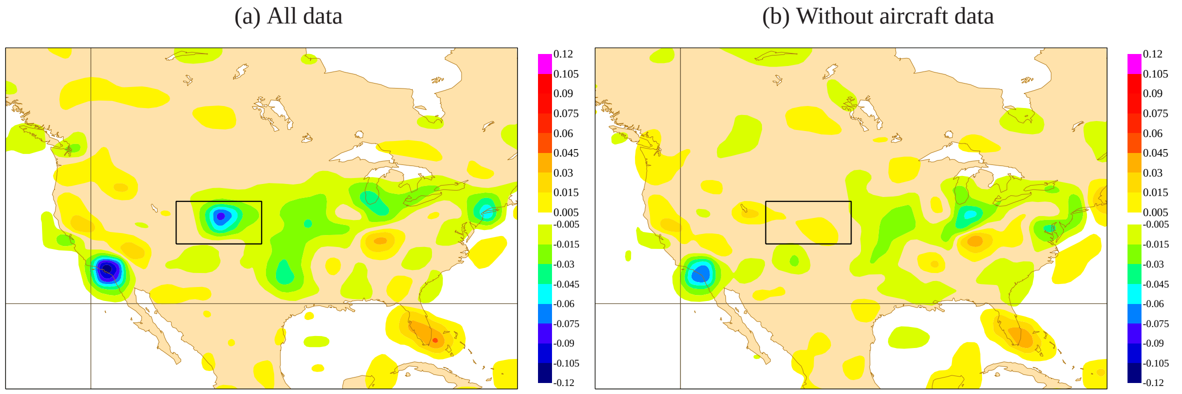

Figure 6: Average temperature forcing at the lowest model level over North America. Removing aircraft data in the area shown on the picture eliminates the spurious forcing.

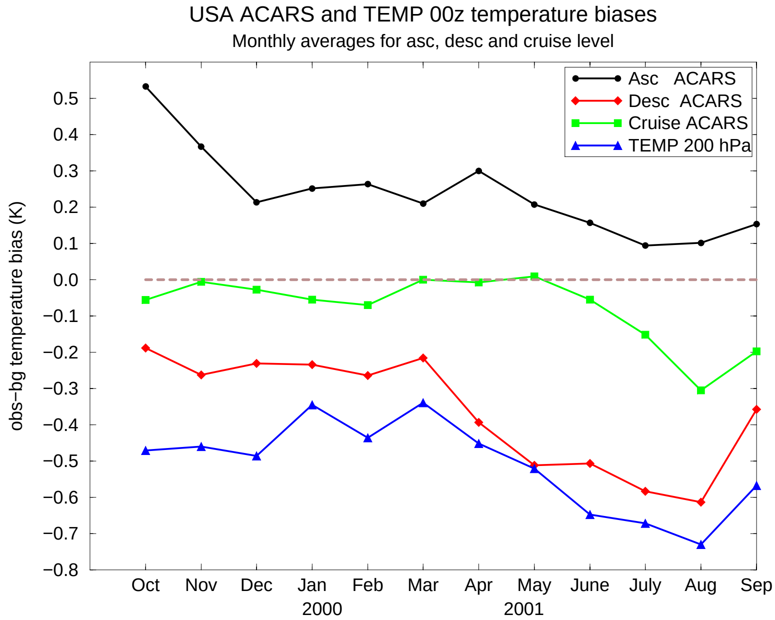

Figure 7: Aircraft temperature observations are biased. Figure from Lars Isaksen, ECMWF.

5 Conclusion

Weak constraints 4D-Var has been implemented in the IFS. It fits observations more uniformly over the assimilation window and captures some known model errors such as the temperature bias in the winter stratosphere. It also captures some observation bias such as the temperature bias in ascending/descending aircraft observations.

4D-Var only sees the difference between model and observations: the distinction between model bias and observation bias can only be achieved through the prior knowledge built into the formulation of 4D-Var. Two tools are available to define the errors we wish to capture: the model error covariance matrix (here geared towards the small scales) which can be tuned to selectively capture certain space and time scales we think are representative of certain types of error and the model for model error (here a constant 3D field) which can be more or less relevant to various forms of error. More research is still needed in both areas to exploit the full potential of weak constraint 4D-Var. The most appropriate formulation of weak constraint 4D-Var in terms of the control variable η, β or x will have to be determined and a choice will have to be made as to what type of model error (random or systematic or bias) we want to estimate. Interactions with variational observation bias correction (Derber and Wu (1998), Dee (2004)) will also have to be studied and taken into account.

In addition to lifting the questionable assumption that the model is perfect, model error is valuable information which can be used in several ways. It can be added as forcing in the forecast model, or at the post-processing stage if a model bias was determined, to improve the forecast. It should also help identify model deficiencies and improve the model.

References

D. Dee. Testing the perfect model assumption in variational data assimilation. In *WMO Symposium on Assimilation of Observations in Meteorology and Oceanography*, Tokyo, March 1995.

D. Dee. Variational bias correction of radiance data in the ECMWF system. In *Workshop on Assimilation of high spectral resolution sounders in NWP*, ECMWF, pages 97–112, 2004.

J. Derber. A variational continuous assimilation technique. *Mon. Wea. Rev.*, 117:2437–2446, 1989.

J. Derber and W.-S. Wu. The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, 126:2287–2299, 1998.

M. Fisher. Background error covariance modelling. In *Workshop on Recent developments in data assimilation for atmosphere and ocean*, ECMWF, pages 45–63, September 2003.

A. Griffith and N. Nichols. Adjoint methods for treating model error in data assimilation. In *Numerical Methods for Fluid Dynamics*, volume VI of *Institute for CFD, Oxford*, pages 335–344, 1998.

F.-X. Le Dimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: Theorical aspects. *Tellus*, 38A:97–110, 1986.

Y. Sasaki. Some basic formalisms in numerical variational analysis. *Mon. Wea. Rev.*, 98:875–883, 1970.

Y. Trémolet. Model error in variational data assimilation. In *Seminar on Recent developments in data assimilation for atmosphere and ocean*. ECMWF, September 2003.

Y. Trémolet. Accounting for an imperfect model in 4D-Var. Tech. Memo. 477, ECMWF, November 2005.

D. Zupanski. A general weak constraint applicable to operational 4DVAR data assimilation systems. *Mon. Wea. Rev.*, 125:2274–2291, 1997.

M. Zupanski. Regional four-dimensional variational data assimilation in a quasi-operational forecasting environment. *Mon. Wea. Rev.*, 121:2396–2408, 1993.