On the diagnosis of model error statistics using weak-constraint data assimilation

N. E. Bowler, Met Office, FitzRoy Road, Exeter, EX1 3PB, UK

May 13, 2016

Abstract

Outputs from a data assimilation system may be used to diagnose observation and background error statistics, as has been demonstrated by previous researchers. In this study, that technique is extended to diagnose model-error statistics using a weak-constraint data assimilation. It deals with a set of observations over a time window and uses the temporal distribution to separate model errors from errors in the background forecast. In idealised tests this method is shown to be able to successfully distinguish between model, background and observation errors. The success of this method depends on the prior assumptions included in the weak-constraint data assimilation and how well these describe the true nature of the system being modelled.

1 Introduction

It has long been recognised that computer models of complex physical processes are imperfect. What has been less clear is how to estimate the magnitude and structure of these imperfections. In particular, how does one differentiate errors in the numerical model from those in the observations and from any chaotic growth of small errors intrinsic to the system being modelled?

Recently Todling (2015) introduced a method to diagnose the model-error covariance from a pair of data assimilations — one of which is a filter and the other a smoother (able to use future observations). This system is described as sequential, since it is devised for a set of observations which are available at discrete times, rather than being spread over a given time window.

2 Weak-constraint data assimilation

To find the analysis in a system which is affected by model error one can use weak-constraint data assimilation. At a set of times we allow the analysis trajectory to depart from the solution given by the nonlinear model according to

$$\mathbf{x}_i = M_i(\mathbf{x}_{i-1}) + \boldsymbol{\eta}_i \tag{1}$$

where $\mathbf{x}_i$ is the model state at time i and M_i is the nonlinear model propagator from time $i-1$ to time i. At each time we are permitting a modification of the model state of $\boldsymbol{\eta}_i$.

If we use this perturbed model to define a four-dimensional state $\underline{\mathbf{x}}$, then we can write the weak-constraint cost function as (Trémolet, 2006)

$$\begin{aligned} J_{\text{weak}} = \frac{1}{2}(\mathbf{x}-\mathbf{x}_\text{b})^\text{T}\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_\text{b}) + \frac{1}{2}\sum_{i=1}^{n}\boldsymbol{\eta}_i^\text{T}\mathbf{Q}_i\boldsymbol{\eta}_i + \\ \frac{1}{2}(\mathbf{y}-H(\underline{\mathbf{x}}))^\text{T}\mathbf{R}^{-1}(\mathbf{y}-H(\underline{\mathbf{x}})). \end{aligned} \tag{2}$$

where $\mathbf{Q}_i$ is the model-error covariance for time i and n is the total number of times at which a modification is allowed. If we assume that the modification at each time is the same, then we may write the total effect of the modifications on the trajectory as

$$\underline{\mathbf{N}}\boldsymbol{\eta} = \begin{pmatrix} 0 \\ \mathbf{M}_{1\leftarrow 1} \\ \mathbf{M}_{2\leftarrow 2} + \mathbf{M}_{2\leftarrow 1} \\ \dots \\ \sum_{i=1}^{t}\mathbf{M}_{t\leftarrow i} \end{pmatrix} \boldsymbol{\eta} \tag{3}$$

where $\mathbf{M}_{j\leftarrow i}$ is the linear propagator of the numerical model from time i to time j. The final term in equation (2) then becomes

$$\begin{aligned} J_o = \frac{1}{2}(\mathbf{y} - H(\underline{M}(\mathbf{x}_\text{b})) - \mathbf{H}\underline{\mathbf{M}}\mathbf{d}_\text{b}^\text{a} - \mathbf{H}\underline{\mathbf{N}}\boldsymbol{\eta})^\text{T}\mathbf{R}^{-1} \\ (\mathbf{y} - H(\underline{M}(\mathbf{x}_\text{b})) - \mathbf{H}\underline{\mathbf{M}}\mathbf{d}_\text{b}^\text{a} - \mathbf{H}\underline{\mathbf{N}}\boldsymbol{\eta}) \end{aligned} \tag{4}$$

where $\mathbf{d}_\text{b}^\text{a}$ is the increment applied to the initial condition. From this we can derive an expression for the model forcing term as

$$\boldsymbol{\eta} = \mathbf{K}^\text{q}\mathbf{d}_\text{b}^\text{o} \tag{5}$$

where

$$\mathbf{K}^\text{q} = \mathbf{Q}\underline{\mathbf{N}}^\text{T}\mathbf{H}^\text{T}\left(\mathbf{H}\underline{\mathbf{M}}\mathbf{B}\underline{\mathbf{M}}^\text{T}\mathbf{H}^\text{T} + \mathbf{R} + \mathbf{H}\underline{\mathbf{N}}\mathbf{Q}\underline{\mathbf{N}}^\text{T}\mathbf{H}^\text{T}\right)^{-1} \tag{6}$$

and $\mathbf{d}_\text{b}^\text{o}$ is the innovation (the difference between the observations and the background trajectory).

2.1 Diagnosis using weak-constraint DA

Desroziers *et al.* (2005) introduced a method to diagnose the observation-, background- and analysis-error covariance matrices from data assimilation statistics. To extend this technique to model errors, we first need to calculate the covariance of the innovations. Following the assumption we made earlier we take the model errors to be constant during the DA window, but uncorrelated with background and observation errors. In this case the innovation covariance is

$$E((\mathbf{d}_\text{b}^\text{o})\,(\mathbf{d}_\text{b}^\text{o})^\text{T}) = \mathbf{R}^\text{o} + \mathbf{H}\underline{\mathbf{M}}\mathbf{B}^\text{o}\underline{\mathbf{M}}^\text{T}\mathbf{H}^\text{T} + \mathbf{H}\underline{\mathbf{N}}\mathbf{Q}^\text{o}\underline{\mathbf{N}}^\text{T}\mathbf{H}^\text{T} \tag{7}$$

where $\mathbf{Q}^\text{o}$ is the observed model-error covariance. For observations at the end of the window the last term is proportional to the number of time-steps squared, n^2, since each $\underline{\mathbf{N}}$ contains a summation of n terms.

Thus, the cross-covariance between the model forcing term and the innovation will be

$$\begin{aligned} E(\mathbf{H}\underline{\mathbf{N}}\boldsymbol{\eta}(\mathbf{d}_\text{b}^\text{o})^\text{T}) &= \mathbf{H}\underline{\mathbf{N}}\mathbf{K}^\text{q}E(\mathbf{d}_\text{b}^\text{o}(\mathbf{d}_\text{b}^\text{o})^\text{T}) \\ &= \mathbf{H}\underline{\mathbf{N}}\mathbf{Q}\underline{\mathbf{N}}^\text{T}\mathbf{H}^\text{T}\Delta\mathbf{K}^\text{w} \end{aligned} \tag{8}$$

where $\Delta\mathbf{K}^{\mathrm{w}}$ is given by

$$\Delta\mathbf{K}^{\mathrm{w}} = \left(\mathbf{H}\underline{\mathbf{M}}\mathbf{B}\underline{\mathbf{M}}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} + \mathbf{R} + \mathbf{H}\underline{\mathbf{N}}\mathbf{Q}\underline{\mathbf{N}}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\right)^{-1} \left(\mathbf{H}\underline{\mathbf{M}}\mathbf{B}^{\mathrm{o}}\underline{\mathbf{M}}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} + \mathbf{R}^{\mathrm{o}} + \mathbf{H}\underline{\mathbf{N}}\mathbf{Q}^{\mathrm{o}}\underline{\mathbf{N}}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\right). \tag{9}$$

To simplify the estimating procedure we use only observations from the first time in the data assimilation window, since the above expression will not then include the tangent linear model.

3 Experimental setup

To investigate the behaviour of the diagnostics tests were completed using the model of Lorenz (1995) which is based on the idea of waves propagating around a latitude circle. This circle is divided into 40 grid-points, and at each time step the grid-points are updated according to

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F \tag{10}$$

where the variables x_i, $i = 1, 2, ..., N$, are defined on a cyclic chain such that $x_{-1} = x_{N-1}$, $x_0 = x_N$ and $x_1 = x_{N+1}$. These experiments use a forcing term $F = 8$ which is within the chaotic regime. The Runge-Kutta 4^{th} order method was used to perform the time stepping, for intervals of $\delta t = 0.05$.

To create a model which is affected by model error, we follow an approach similar to that of Todling (2015). For each time-step in the truth run a random term is added to equation (10) of the following form

$$\delta\mathbf{r} = \mathbf{G}^{1/2}\delta\mathbf{p} \tag{11}$$

where $\mathbf{G}^{1/2}$ is the symmetric square-root of the covariance matrix $\mathbf{G}$. For the first half of the domain $\mathbf{G}$ takes values given by a Gaussian function of the distance between the points, using a length-scale of 5 grid-points. For the second half of the domain all the elements are zero, meaning that only the first half of the model is perturbed.

The data assimilation was run using weak-constraint 4DVar. This was given observations every time-step, and the data-assimilation window used observations from three times. Observations were produced by perturbing the truth run with errors sampled from $N(0, 0.1^2)$. By choosing small observation errors we ensure that the analysis errors are small, and the tangent-linear approximation used by 4DVar is valid.

4 Results

Figure 1 shows the estimates of the single-step model-error covariance matrix for the Lorenz '95 system. The initial input to the data assimilation (top-left) is an homogenous and nearly-diagonal covariance matrix. This is taken from the background-error covariance matrix estimated from an experiment using the Lorenz '95 model without model error and scaled to give reasonable results. As an approximation to the true model-error covariance matrix (top-right) it is quite poor.

The diagnostic estimate of the model-error covariance is shown in the bottom-left. The second half of the domain does not experience model error, and the estimated model error covariance is much reduced in this region. There is still an imprint of the initial model-error covariance in the estimated matrix,

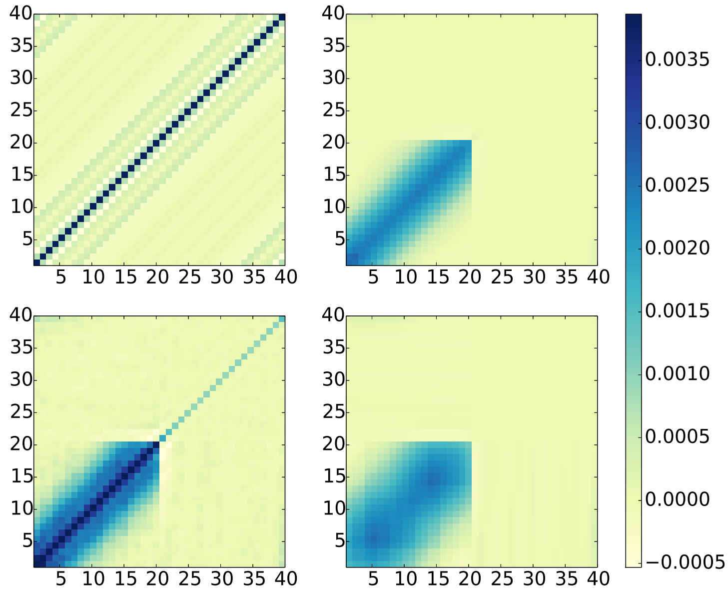

Figure 1: Single-step model error covariance matrix for the Lorenz '95 model. In these graphs the top-left graph shows the scaled input provided to the data assimilation, the top-right shows the true covariance matrix which is the target for the estimation. The bottom-left graph shows the estimate from the first run of the DA, and the bottom-right graph shows the estimate from the tenth run.

but the magnitude is much reduced. The diagonal elements in the first part of the domain are also reduced. However the off-diagonal elements are increased, reflecting the correlations in the true error covariance matrices. This is iterated by placing the diagnosed **B**, **Q** and **R** matrices as input in the next run of the data assimilation. After 10 iterations the diagnosed **Q** matrix (bottom right) is very close to the true **Q** matrix (top right).

Acknowledgement

This work was developed following a series of discussions with Mike Cullen. The author would like to thank him for his very helpful input.

References

Desroziers G, Berre L, Chapnik B, Poli P. 2005. Diagnosis of observation, background and analysis-error statistics in observation space. *QJR Meteorol Soc* **131**: 3385–3396. doi: 10.1256/qj.05.108.

Lorenz EN. 1995. Predictability: a problem partly solved. In: *Proceedings of the seminar on predictability*, vol. I. ECMWF: Reading, Berkshire, UK, pp. 1–18.

Todling R. 2015. A lag-1 smoother approach to system-error estimation: sequential method. *QJR Meteorol Soc* **141**: 1502–1513. doi: 10.1002/qj.2460.

Trémolet Y. 2006. Accounting for an imperfect model in 4D-Var. *QJR Meteorol Soc* **132**: 2483–2504. doi: 10.1256/qj.05.224.