Model error estimation: Its application to chemical data assimilation

Richard Ménard*, Yan Yang† and Saroja Polavarapu†

Meteorological Service of Canada
(*) Air Quality Research Branch, Dorval, Québec, Canada
(†) Data Assimilation and Satellite Meteorology Division, Totonto, Canada

Abstract

The estimation of model error, both its bias and error covariance, has been a challenging and somewhat illusive problem of data assimilation. Yet, and especially in new areas of applications such as chemical data assimilation, the observations-minus-model residuals can be as much systematic as random. Estimation of model error is thus a necessary part of the assimilation. The objective of these notes is to introduce a formal method to estimate both the state and the model error, and to expose the problems associated with model error estimation. The resulting algorithm can be included into operational analysis systems by simply adding an additional step. This method is currently being developed for the chemical assimilation analysis scheme at the Meteorological Service of Canada.

1. A simple example of state-model error estimation

We will illustrate the basic mechanism of estimation of state and model error using a simple chemical box model,

$$\frac{\partial x}{\partial t} + L\,x = u \ , \tag{1}$$

where x is the concentration of a specie, L represents the transport operator, and u represent the chemical source, assumed to be constant. In this problem we are interested in estimating two variables; the concentration (i.e. the state) and the source (i.e. the model error), on the basis of observations of concentration. Introducing a state-augmented variable

$$\mathbf{z} = \begin{pmatrix} x \\ u \end{pmatrix}, \tag{2}$$

the state-augmented observation operator is of the form

$$\overline{\mathbf{H}} = \begin{pmatrix} \mathbf{H} & 0 \end{pmatrix}, \tag{3}$$

and this has several implications for the analysis. The results that will follow are easily derived using the Kalman filter equations. The Kalman gain of the state- augmented system can be written as

$$\overline{\mathbf{K}} = \begin{pmatrix} \mathbf{P}^f_{xx}\mathbf{H}^T \\ \mathbf{P}^f_{ux}\mathbf{H}^T \end{pmatrix} \left(\mathbf{H}\,\mathbf{P}^f_{xx}\mathbf{H}^T + \mathbf{R} \right)^{-1}, \tag{4}$$

and his basically composed of two parts, 1 - a concentration (or state estimate) gain matrix,

$$\mathbf{K}_x = \mathbf{P}^f_{xx}\,\mathbf{H}^T \left(\mathbf{H}\,\mathbf{P}^f_{xx}\mathbf{H}^T + \mathbf{R} \right)^{-1} \tag{5}$$

and, 2- a source (or model error) gain matrix

$$\mathbf{K}_u = \mathbf{P}^f_{ux}\mathbf{H}^T \left(\mathbf{H}\,\mathbf{P}^f_{xx}\mathbf{H}^T + \mathbf{R} \right)^{-1} = \mathbf{P}^f_{ux}(\mathbf{P}^f_{xx})^{-1}\mathbf{K}_x \,, \tag{6}$$

and we emphasize, as it is in the case for any unobserved variables, that *it is only from the correlation between the concentration (or state) error and the source (or model) error that observational information can be projected onto the source's (or model error) estimate*. To get some insight on the properties of analysis error it is useful to consider a scalar case – an analysis at a single point. The analysis error is then written as

$$p^a_{xx} = p^f_{xx} - \frac{\left(p^f_{xx}\right)^2}{p^f_{xx} + \sigma^2_o} \tag{7a}$$

$$p^a_{xu} = p^f_{xu} - \frac{p^f_{xx}\, p^f_{xu}}{p^f_{xx} + \sigma^2_o} \tag{7b}$$

$$p^a_{uu} = p^f_{uu} - \frac{\left(p^f_{xu}\right)^2}{p^f_{xx} + \sigma^2_o} \tag{7c}$$

where p^a, p^f denotes respectively the analysis error variance and the forecast error variance, and σ^2_o the observation error variance. A measure of efficiency of the observation is given for instance by the ratios of analysis over forecast error variances, i.e p^a_{xx} / p^f_{xx} , p^a_{xu} / p^f_{xu} , and p^a_{uu} / p^f_{uu}. For the state covariance and the concentration-source error cross-covariance, the reduction of variance is the same,

$$\frac{p^a_{xx}}{p^f_{xx}} = \frac{p^a_{xu}}{p^f_{xu}} = 1 - \frac{1}{1+\beta} \ , \tag{8}$$

where $\beta = \sigma^2_o / p^f_{xx}$, but for the source error variance it is given by

$$\frac{p^a_{uu}}{p^f_{uu}} = 1 - \frac{\rho^2_{xu}}{1+\beta} \ . \tag{9}$$

where ρ_{xu} is the *forecast* concentration-source error correlation. We thus observe the correlation between forecast error concentration and source errors is needed to reduce the uncertainty of the source, and that since $0 \le |\rho_{xu}| \le 1$, the observation has a greater impact in reducing the concentration (i.e. state) error than the source (i.e. model) error. The analysis update, however, changes the cross-correlation. We can show that ratio between the *a posteriori* (analysis) and the *a priori* (forecast) correlation between concentration error and source error, is given by

$$\frac{\left(\rho^+_{xu}\right)^2}{\rho^2_{xu}} = \frac{1}{1 + \frac{1}{\beta}(1-\rho^2_{xu})} \le 1 \ , \tag{10}$$

so *the effect of observing the concentration (i.e. the state) reduces the cross concentration-source error (i.e. state-model error) correlation*. In summary, we note that concentration-source error correlation (i.e. state-model error correlation) is needed to infer information about the source (or model error) from measurements, but the result of observing actually reduces that correlation. Overall, *the efficiency of estimating sources from observations depends on two competing mechanisms; the increase of error correlation between concentration and source due to the forecast, and the reduction of this correlation resulting from the analysis.*

2. Separation of the analysis

The fact that the model error grain matrix can be obtained from the state gain matrix (6), indicates that state and model error estimates can be obtained in sequence (one after the other). Because of the relationship between variational and sequential methods (e.g. Kalman filter), the same separation can be obtained with variational methods. Here we can assume that observation operator is nonlinear. The separation results only from the fact that model error is not an observed variable.

If we were to carry out the 3D Var analysis of the state-augmented variable, we would minimize,

$$J(\mathbf{z}) = \frac{1}{2}\left(\mathbf{z} - \widetilde{\mathbf{z}}^f\right)^T \left(\mathbf{P}^f\right)^{-1}\left(\mathbf{z} - \widetilde{\mathbf{z}}^f\right) + \frac{1}{2}\left(\mathbf{y} - \mathbf{H}(\mathbf{z})\right)^T \mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{H}(\mathbf{z})\right) \tag{11}$$

Here $\widetilde{\mathbf{z}}^f$ is an *unbiased* background state or trial field (tilde indicate an unbiased variables), and

$$\mathbf{P}^f = \begin{pmatrix} \mathbf{P}^f_{xx} & \mathbf{P}^f_{xu} \\ \mathbf{P}^f_{ux} & \mathbf{P}^f_{uu} \end{pmatrix} \tag{12}$$

is the forecast (or background error covariance). The bias is either an a priori, which could be obtained from comparison with climatology, or the model error obtained from the previous analysis. It is convenient to do a co-ordinate transformation, which also acts as a pre-conditioning based on $\mathbf{P}^f$. First define

$$\mathbf{P}^f = \mathbf{S}\mathbf{S}^T \tag{13}$$

There is no unique way to obtain **S**, but one convenient way is

$$\mathbf{S} = \begin{bmatrix} \mathbf{L} & \mathbf{0} \\ \mathbf{P}_{ux}\left(\mathbf{L}^T\right)^{-1} & \mathbf{D}^{-1/2} \end{bmatrix}, \tag{14}$$

where

$$\mathbf{P}_{xx} = \mathbf{L}\mathbf{L}^T \quad , \quad \mathbf{D}^{-1} = \mathbf{P}_{uu} - \mathbf{P}_{ux}\mathbf{P}_{xx}^{-1}\mathbf{P}_{xu}\,. \tag{15}$$

Then with the transformation

$$\mathbf{z} = \mathbf{S}\,\varsigma \quad , \quad \varsigma = \begin{pmatrix} \xi \\ \eta \end{pmatrix} \tag{16}$$

(11) becomes,

$$\begin{aligned} J(\zeta) &= \frac{1}{2}\left(\varsigma - \widetilde{\varsigma}^f\right)^T \left(\varsigma - \widetilde{\varsigma}^f\right) + \frac{1}{2}\left(\mathbf{y} - \mathbf{H}(\mathbf{L}(\xi))\right)^T \mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{H}(\mathbf{L}(\xi))\right) \\ &= \frac{1}{2}\left(\xi - \widetilde{\xi}^f\right)^T \left(\xi - \widetilde{\xi}^f\right) + \frac{1}{2}\left(\eta - \widetilde{\eta}^f\right)^T \left(\eta - \widetilde{\eta}^f\right) + \frac{1}{2}\left(\mathbf{y} - \mathbf{H}(\mathbf{L}(\xi))\right)^T \mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{H}(\mathbf{L}(\xi))\right) \end{aligned} \tag{17}$$

Note that ξ and ζ are independent, and can be solved for sequentially. First minimize:

$$J_1(\xi) = \frac{1}{2}\left(\xi - \widetilde{\xi}^f\right)^T \left(\xi - \xi^f\right) + \frac{1}{2}\left(\mathbf{y} - \mathbf{H}(\xi)\right)^T \mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{H}(\xi)\right) \tag{18}$$

Then minimize

$$J_2(\eta) = \frac{1}{2}\left(\eta - \widetilde{\eta}^f\right)^T \left(\eta - \widetilde{\eta}^f\right) \tag{19}$$

We note that (17) is just the standard 3D Var cost function using **L** as preconditioning, while (18) doesn't need to be minimized, the solution is readily obtained as

$$\boldsymbol{\eta} = \widetilde{\boldsymbol{\eta}}^{f} \tag{20}$$

Since the inverse of **S** is easily seen to be,

$$\mathbf{S}^{-1} = \begin{bmatrix} \mathbf{L}^{-1} & \mathbf{0} \\ -\mathbf{D}^{1/2}\mathbf{P}_{ux}\left(\mathbf{P}_{xx}\right)^{-1} & \mathbf{D}^{1/2} \end{bmatrix}, \tag{21}$$

then we get

$$\mathbf{u}^{a} = \mathbf{u}^{f} + \mathbf{P}_{ux}\left(\mathbf{P}_{xx}\right)^{-1}\left(\mathbf{x}^{a} - \widetilde{\mathbf{x}}^{f}\right), \tag{22}$$

in agreement with using the source gain matrix (6). Thus, for 3D Var, a bias correction can be easily implemented without modification to an existing scheme which minimizes (18), only additional calculation (22) is needed.

This separation of the analysis depends only on the fact that the model error **u** is unobserved, and can actually be used for analysis on unobserved variables. This is, in particular, useful for atmospheric chemistry applications, where generally only a few chemical species are observed while the model state usually carries much more species. The analysis update (22) to unobserved variables can be implemented in parallel, each species on different CPUs or node.

3. Kalman filter formulation

3.1. True system

We first define the evolution of the true augmented state as follows,

$$\mathbf{x}_{k+1}^{t} = M_{k}\left(\mathbf{x}_{k}^{t}\right) + \mathbf{G}_{k}\mathbf{u}_{k}^{t} + \boldsymbol{\varepsilon}_{k}^{q} \tag{23}$$

$$\mathbf{u}_{k+1}^{t} = \mathbf{u}_{k}^{t} + \boldsymbol{\varepsilon}_{k}^{u} \tag{24}$$

$$\mathbf{y}_{k}^{o} = H_{k}\left(\mathbf{x}_{k}^{t}\right) + \boldsymbol{\varepsilon}_{k}^{r}. \tag{25}$$

Note that the state vector has serially correlated model error forcing or colored noise, $\mathbf{u}_{k}$. For simplicity the model error forcing has been assumed constant. $\mathbf{G}_{k}$ is in general a non-square matrix, so that **u** can represent anything from a single parameter to a full state bias (when $\mathbf{G}_{k} = \mathbf{I}$). We will make the usual Kalman filter assumption on the error terms

$$\begin{aligned}
&\left\langle \boldsymbol{\varepsilon}_{k}^{q} \right\rangle = \mathbf{0}, \quad \left\langle \boldsymbol{\varepsilon}_{k}^{q}\left(\boldsymbol{\varepsilon}_{l}^{q}\right)^{T} \right\rangle = \mathbf{Q}_{k}\delta_{l}^{k} \\
&\left\langle \boldsymbol{\varepsilon}_{k}^{r} \right\rangle = \mathbf{0}, \quad \left\langle \boldsymbol{\varepsilon}_{k}^{r}\left(\boldsymbol{\varepsilon}_{l}^{r}\right)^{T} \right\rangle = \mathbf{R}_{k}\delta_{l}^{k} \\
&\left\langle \boldsymbol{\varepsilon}_{k}^{u} \right\rangle = \mathbf{0}, \quad \left\langle \boldsymbol{\varepsilon}_{k}^{u}\left(\boldsymbol{\varepsilon}_{l}^{u}\right)^{T} \right\rangle = \mathbf{S}_{k}\delta_{l}^{k} \\
&\left\langle \boldsymbol{\varepsilon}_{k}^{q}\left(\boldsymbol{\varepsilon}_{l}^{r}\right)^{T} \right\rangle = \left\langle \boldsymbol{\varepsilon}_{k}^{q}\left(\boldsymbol{\varepsilon}_{l}^{u}\right)^{T} \right\rangle = \left\langle \boldsymbol{\varepsilon}_{k}^{u}\left(\boldsymbol{\varepsilon}_{l}^{r}\right)^{T} \right\rangle = \mathbf{0}
\end{aligned} \tag{26}$$

where $\mathbf{Q}_{k}$ is the (state) model error covariance, $\mathbf{R}_{k}$ is the observational error covariance, and $\mathbf{S}_{k}$ is the colored noise model error covariance.

3.2. The forecast step

The forecast equations will take the general form

$$\mathbf{x}_{k+1}^{f} = M_{k}\left(\mathbf{x}_{k}^{a}\right) + (1-\alpha)\mathbf{G}_{k}\mathbf{u}_{k}^{a} \tag{27}$$

$$\mathbf{u}^f_{k+1} = \mathbf{u}^a_k \tag{28}$$

We have introduced here a parameter α that can take either of the two values 0 or 1. For $\alpha = 0$, the forecast is *aware* of the bias and actually takes it into account. For $\alpha = 1$, the forecast is *blind* to the bias, and no correction is made to correct for the forecast bias (such as that assumed in Dee and daSilva 1998).

The evolution of the forecast error is given by

$$\mathbf{P}^f_{xx,k+1} = \mathbf{M}_k \mathbf{P}^a_{xx,k} \mathbf{M}^T_k + \mathbf{Q}_k + \mathbf{M}_k \mathbf{P}^a_{xu,k} \mathbf{G}^T_k + \mathbf{G}_k \mathbf{P}^a_{ux,k} \mathbf{M}^T_k + \mathbf{G}_k \mathbf{P}^a_{uu,k} \mathbf{G}^T_k \tag{29}$$

$$\mathbf{P}^f_{ux,k+1} = \mathbf{P}^a_{ux,k} \mathbf{M}^T_k + \mathbf{P}^a_{uu,k} \mathbf{G}^T_k \tag{30}$$

$$\mathbf{P}^f_{xu,k+1} = \left(\mathbf{P}^f_{ux,k+1}\right)^T \tag{31}$$

$$\mathbf{P}^f_{uu,k+1} = \mathbf{P}^a_{uu,k} + \mathbf{S}_k \tag{32}$$

and is the same for both bias-blind and bias-aware forecast models.

3.3. Analysis step

Consider the adjusted (or de-biased) forecast,

$$\widetilde{\mathbf{x}}^f_k \equiv \mathbf{x}^f_k + \alpha \mathbf{G}_{k-1} \mathbf{u}^a_{k-1} = M_{k-1}\left(\mathbf{x}^a_{k-1}\right) + \mathbf{G}_{k-1} \mathbf{u}^a_{k-1} \tag{33}$$

Its error, $\widetilde{\boldsymbol{\varepsilon}}^f_{x,k} \equiv \widetilde{\mathbf{x}}^f_k - \mathbf{x}^t_k$, also given by

$$\widetilde{\boldsymbol{\varepsilon}}^f_{x,k} = \boldsymbol{\varepsilon}^f_{x,k} + \alpha \mathbf{G}_{k-1} \mathbf{u}^a_{k-1} \tag{34}$$

has zero mean. It is with respect to this quantity (i.e. unbiased quantities) that the analysis step can be formulated, and this yields the familiar form

$$\mathbf{x}^a_k = \widetilde{\mathbf{x}}^f_k + \mathbf{K}^x_k \left[\mathbf{y}^o_k - H_k\left(\widetilde{\mathbf{x}}^f_k\right)\right] \tag{35}$$

$$\mathbf{u}^a_k = \mathbf{u}^f_k + \mathbf{K}^u_k \left[\mathbf{y}^o_k - H_k\left(\widetilde{\mathbf{x}}^f_k\right)\right] = \mathbf{u}^f_k + \mathbf{P}^f_{ux,k} \left(\mathbf{P}^f_{xx,k}\right)^{-1} \left[\mathbf{x}^a_k - \widetilde{\mathbf{x}}^f_k\right], \tag{36}$$

where $\mathbf{K}^x_k$, $\mathbf{K}^u_k$ are given by (5),(6). Note that the forecast error covariance is the one defined using the de-biased forecast, and is related to the bias-blind forecast error $\mathbf{B}^f_{xx,k} \equiv \left\langle \boldsymbol{\varepsilon}^f_{x,k} \left(\boldsymbol{\varepsilon}^f_{x,k}\right)^T \right\rangle$ as follows,

$$\mathbf{P}^f_k = \begin{bmatrix} \mathbf{P}^f_{xx,k} & \mathbf{P}^f_{xu,k} \\ \mathbf{P}^f_{ux,k} & \mathbf{P}^f_{uu,k} \end{bmatrix} = \begin{bmatrix} \mathbf{B}^f_{xx,k} - \alpha^2 \mathbf{G}_{k-1} \mathbf{u}^a_{k-1} \left(\mathbf{u}^a_{k-1}\right)^T \mathbf{G}^T_{k-1} & \mathbf{P}^f_{xu,k} \\ \mathbf{P}^f_{ux,k} & \mathbf{P}^f_{uu,k} \end{bmatrix} \tag{37}$$

The analysis error covariance resulting from the analysis update, is given by

$$\mathbf{P}^a_k = \begin{bmatrix} \left(\mathbf{I} - \mathbf{K}^x_k \mathbf{H}_k\right) \mathbf{P}^f_{xx,k} & \left(\mathbf{I} - \mathbf{K}^x_k \mathbf{H}_k\right) \mathbf{P}^f_{xu,k} \\ -\mathbf{K}^u_k \mathbf{H}_k \mathbf{P}^f_{xx,k} & -\mathbf{K}^u_k \mathbf{H}_k \mathbf{P}^f_{xu,k} \end{bmatrix} \tag{38}$$

Remark: Because of the special form of the observation operator, i.e. only the state is observed, the analysis can be decomposed into the usual state analysis (35), followed by the updating of the model error estimate using the analysis increments (36). Analysis with bias-blind forecast is basically of the same form, except for (33) to define a de-biased forecast, and with appropriate forecast error covariances (37).

4. Lagged innovations diagnostics and its application to the estimation of $\mathbf{P}_{ux}$

Since **u** is not an observed variable, the lag-0 innovations covariance does not provide information about $\mathbf{P}_{ux}$. Dynamics is needed to enable information about $\mathbf{P}_{ux}$ to become "visible" to the observations. In fact, $\mathbf{P}_{ux}$ appears as a model error growth rate. To see this, consider the continuous time equations for error covariances. In its simplest form we have the following equations for the truth,

$$\begin{aligned} \frac{\partial x}{\partial t} &= L\,x + G\,u \\ \frac{\partial u}{\partial t} &= \varepsilon^{u} \end{aligned} \tag{39}$$

and from which we get

$$\frac{\partial P_{xx}}{\partial t} = L\,P_{xx} + P_{xx}L^{T} + G\,P_{ux} + P_{xu}G^{T}\,. \tag{40}$$

The first two terms on the right hand side, represents the usual propagation of analysis error, and GP_{ux} appears as a model error growth rate.

Experience is estimating model error covariance in data assimilation is somewhat limited. The use of χ^2 trend (Ménard and Chang, 2000), or of the curvature of the "fit to observations" (Ménard and Daley, 1996) have been used, but are limited in the number of parameter one can estimate. Using lag-0 and the residual covariances in predictability experiments, Daley (1992b) showed how to estimate the model error covariance. The lag-innovation covariances, introduced by Daley (1992a), turn out to have direct information about $\mathbf{P}_{ux}$.

In this problem, the lag-1 innovation covariance is given by

$$\begin{aligned} \mathbf{C}_{k+1}^{k} \equiv \left\langle \widetilde{\boldsymbol{\varepsilon}}_{x,k+1}^{f} \left(\widetilde{\boldsymbol{\varepsilon}}_{x,k}^{f}\right)^{T} \right\rangle &= \mathbf{H}_{k+1}\,\mathbf{M}_{k}\left[\left(\mathbf{I}-\mathbf{K}_{k}^{x}\,\mathbf{H}_{k}\right)\mathbf{P}_{xx,k}^{f} - \mathbf{K}_{k}^{x}\,\mathbf{R}_{k}\right] \\ &\quad + \mathbf{H}_{k+1}\,\mathbf{G}_{k}\left[\mathbf{P}_{ux,k}^{f}\,\mathbf{H}_{k}^{T} - \mathbf{K}_{k}^{u}\left(\mathbf{R}_{k} + \mathbf{H}_{k}\,\mathbf{P}_{xx,k}^{f}\,\mathbf{H}_{k}^{T}\right)\right] \end{aligned} \tag{41}$$

If $\mathbf{K}_{k}^{x}$ is optimal then the first term in brackets is zero. If the model error is estimated using a correct $\mathbf{K}_{k}^{u}$, then the second term in brackets is also zero. However, when the model error is not estimated, i.e. $\mathbf{K}_{k}^{u} = \mathbf{0}$, then the lag-1 innovation covariance gives,

$$\mathbf{C}_{k+1}^{k} = \mathbf{H}_{k+1}\,\mathbf{G}_{k}\,\mathbf{P}_{ux,k}^{f}\,\mathbf{H}_{k}^{T}\,, \tag{42}$$

In particular when $\mathbf{G}_{k} = \mathbf{H}_{k} = \mathbf{I}$, then

$$\mathbf{C}_{k+1}^{k} = \mathbf{P}_{ux,k}^{f}\,. \tag{43}$$

In general (41) can be written is a quite useful form, as follows. The equation (41) is independent of the fact that $\mathbf{K}_{k}^{x}$, $\mathbf{K}_{k}^{u}$ are optimal, in fact let's suppose that the gain matrices are suboptimal, and denote them by $\widetilde{\mathbf{K}}_{k}^{x}$, $\widetilde{\mathbf{K}}_{k}^{u}$. Since the lag-0 innovation covariance is

$$\mathbf{C}_{k}^{k} = \mathbf{H}_{k}\,\mathbf{P}_{xx,k}^{f}\,\mathbf{H}_{k}^{T} + \mathbf{R}_{k} \tag{44}$$

Then we can write (41) is this useful form

$$\mathbf{C}_{k+1}^{k} = \mathbf{H}_{k+1}\,\mathbf{M}_{k}\left[\mathbf{K}_{k}^{x}-\tilde{\mathbf{K}}_{k}^{x}\right]\mathbf{C}_{k}^{k}+\mathbf{H}_{k+1}\mathbf{G}_{k}\left[\mathbf{K}_{k}^{u}-\tilde{\mathbf{K}}_{k}^{u}\right]\mathbf{C}_{k}^{k} \tag{45}$$

In practice, however, one cannot use $\tilde{\mathbf{K}}_{k}^{u}=\mathbf{0}$, unless $\mathbf{S}_{k}=\mathbf{0}$ because the observation can't control the growth of $\mathbf{P}_{uu}$ causing filter to divergence.

A methodology to extract $\mathbf{P}_{ux}$ from assimilation experiments is currently under investigation, and will be reported subsequently in paper in preparation.

5. A simple experiment

As an illustration of the method, let's consider a simple advection problem on a one-dimensional periodic domain. Let's suppose that the advected field is forced by a constant and uniform forcing to be estimated. Here we have run a Kalman filter both optimal and non-optimal with a $\tilde{\mathbf{K}}_{k}^{u}$ that is ten times smaller than the optimal parameter gain. Here $\mathbf{G}=\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^{T}$, $\mathbf{Q}=\mathbf{0}$, and the error variance of the parameter is chosen to be as large as the forecast error variance. Also we assume to have only one observation point, $\mathbf{H}=\begin{bmatrix}1 & 0 & \cdots & 0\end{bmatrix}$.

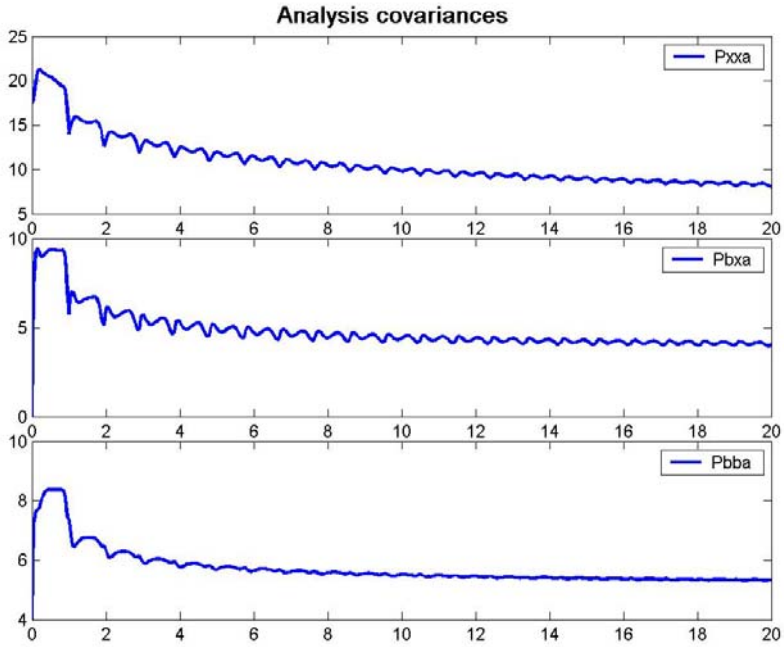

Figure 1 Evolution of trace of the $\mathbf{P}_{xx}^{a}$ (upper panel), the sum of elements of the $\mathbf{P}_{xu}^{a}$ (middle panel), and the parameter error variance $\mathbf{P}_{uu}^{a}$

We observe that although the initial value of the cross-error covariance between state and parameter was initially zero, this terms grows rapidly as a result of the dynamical coupling. The actual value of the parameter is shown next, in figure 2 and 3. Figure 2 shows the result with optimal gain matrices. We note a remarkable track of the true value, in fact we can't distinguish the analysis from the truth in this graphic, as the two curves superposes for the most part except initially where there is a difference of 30 units.

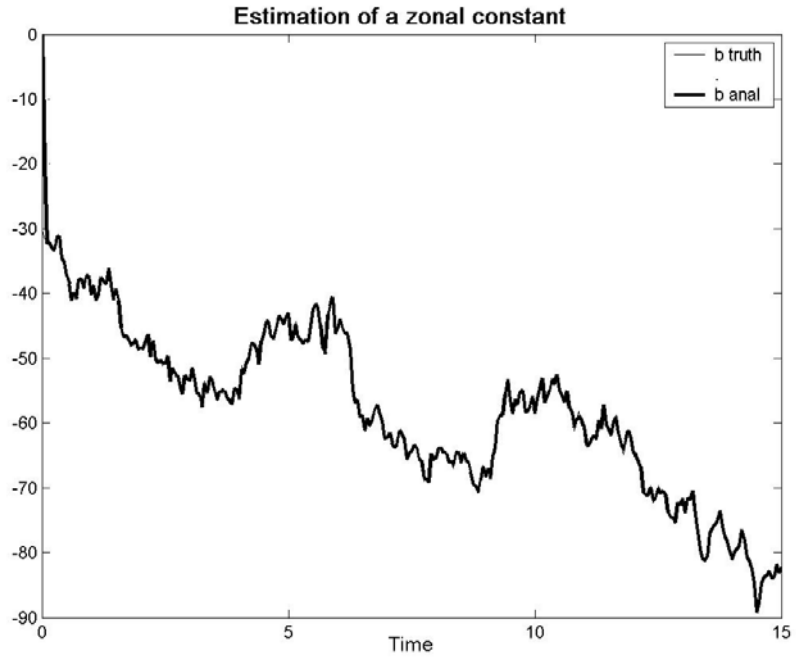

Figure 2 Evolution of u^t , u^f, and u^a , using the optimal Kalman filter.

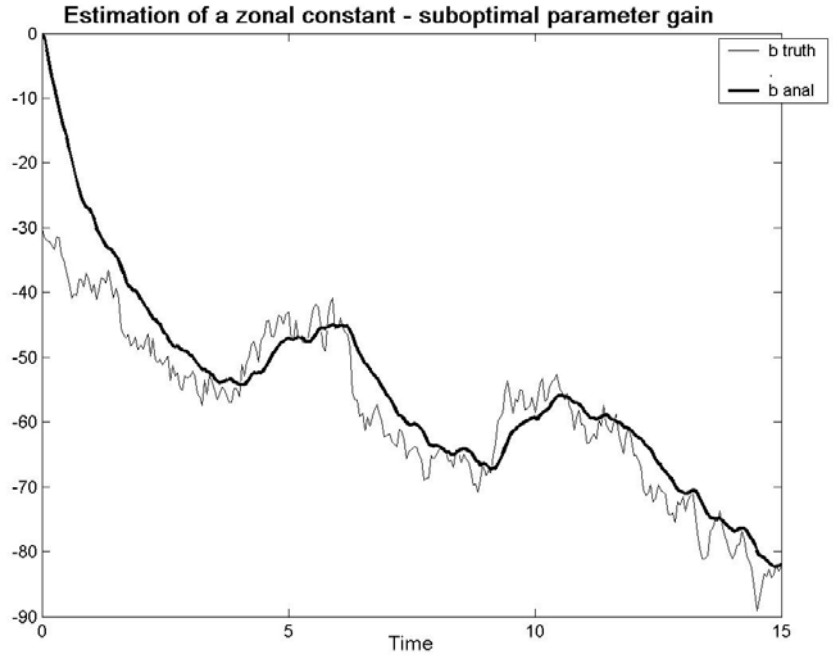

Figure 3 Evolution of u^t , u^f, and u^a , with a sub-optimal parameter filter gain vector $\tilde{\mathbf{K}}_k^u$

We observe somewhat of a good tracking the truth is with sub-optimal parameter gain matrix, that is ten times smaller than the optimal parameter gain matrix. What should be noted from these experiments is that the scheme is able to track a time varying model error. The addition of model error on the ***u*** equation (24) allows precisely to relax the assumption that the true model error should be constant. This an important point since in practice most likely, the time evolution of the parameter is not know a priori.

6. Summary

In numerous applications of data assimilation, not only systematic errors can't be ignored but it may be unclear how to improve the model or observations in a way that can would reduce this error. This is the case for instance in atmospheric chemistry where chemical sources are largely unknown and depend on emission inventories. The problem may also be one of diagnosing the systematic error, when for instance, the systematic error varies over time or space, making it difficult to capture from time means or spatial averaging. Whether it is for estimating, correcting, or for diagnosing purposes, a method that would perform optimal estimation of systematic errors, while at the same time would disentangle the contribution due to the initial errors, is much needed for model improvement and for better use of observations in data assimilation. This is the objective of these notes to establish such a method.

Starting from a state-augmented formulation, we show that the state and model error can be estimated in two steps; first be keeping the analysis step as it is, and second by adding a step for the estimation of model error. This additional step is rather easy to implement and does not involve the minimization of some cost function In fact such separation of the analysis scheme arises solely from having observed and unobserved variables. This splitting can be used for the other problems as well, as for instance in chemical data assimilation. Standard analysis scheme is done with the observed species, and in a second step, the analysis correction to unobserved species can be calculated.

The gain for the model error, i.e. the second step of the analysis, requires the knowledge of the cross-error-covariance between the state and model error. It is shown that lagged-innovations covariances do contain information about the cross-error-covariance between state and model error. Finally, simple experiment of estimation of a single parameter, illustrates, how this estimation procedure, may work, and shows the tracking of a random walk solution.

References

Daley, R., 1992a: The lagged innovation covariance: A performance diagnostic for atmospheric data assimilation. *Mon. Wea. Rev.*, **120**, 178-196.

Daley, R., 1992b: Estimating model-error covariances for application to atmospheric data assimilation. *Mon. Wea. Rev.*, **120**, 1735-1746.

Dee, D.P., and A.M. daSilva, 1998: Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.*, **124**, 269-295.

Ménard, R., and R. Daley, 1996: The application of Kalman smoother theory to the estimation of 4DVAR error statistics. *Tellus*, **48A**, 221-237.

Ménard, R., and L.-P. Chang, 2000: Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part II: -validated results and analysis of variance and correlation dynamics. *Mon. Wea. Rev.*, 128, 2672-2686.