Detection and correction of model bias during data assimilation

Dick P. Dee*

ECMWF, Shinfield Park, Reading
RG2 9AX, United Kingdom
dee@ecmwf.int

1 Introduction

This paper concerns the ubiquitous *bias problem* in data assimilation, which is caused by the presence of systematic (as opposed to random) errors in both models and observations. Systematic model errors can arise, for example, from inaccurate land-surface forcing, poor resolution of the boundary layer, simplified representations of moist physics and clouds, and various other model imperfections. Observation biases for different types of instruments are often caused by specific aspects of the measurement process, or—in the context of data assimilation—by approximations in the observation operators that are used to represent the relationship between the observables and the model variables. It takes a great deal of work to screen the observations and correct them for known biases in order to extract the information that is actually useful to assimilate (e.g. McNally *et al.* 2000). However, we will assume in what follows that all this has been done, and that any remaining bias in the observations may be neglected by the analysis scheme.

Data assimilation in the presence of model bias can generate spurious signals and trends in the assimilated fields, and this problem is exacerbated by changes in the observing system. Figure (1) shows a simple illustration of an assimilation of unbiased observations with a biased model. The visible tendency for the model to drift from the true state produces a positive mean error in the assimilation, whose size depends on the accuracy as well as the frequency of the observations. As a result, a change in the characteristics of the observing system causes an apparent change in climate.

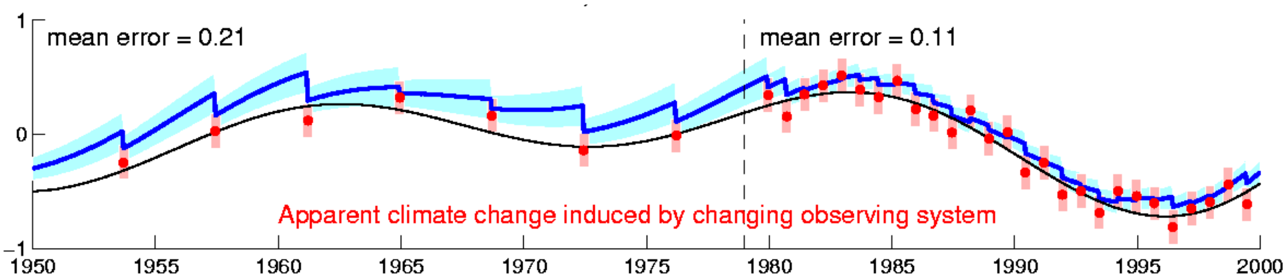

Figure 1: Assimilation of unbiased observations in a biased model, and the effect of a change in observing frequency on the apparent climate. Observations and their error bars are indicated in red; the assimilation trajectory and its error bars in blue. The black curve represents the truth.

In this paper we consider methods for detecting, estimating, and correcting model bias during data assimilation. The techniques are strictly statistical and do not resolve the underlying cause of model bias. Naturally it is always preferable to remove the cause once it has been identified, but that is not an easy task. Hopefully the statistical estimation of model bias (or other systematic model errors) will be beneficial to model developers.

*On leave from NASA Global Modeling and Assimilation Office, Goddard Space Flight Center, Greenbelt, MD 20771, USA

2 Systematic errors and data assimilation

All analysis methods are of the form

$$\mathbf{x}^a - \mathbf{x}^f = \mathbf{k}[\mathbf{y} - \mathbf{h}(\mathbf{x}^f)] \tag{1}$$

where $\mathbf{x}^{a,f}$ are the analysis and background estimates, respectively, of the true state $\mathbf{x}$, $\mathbf{h}(\cdot)$ is the observation operator, the vector $\mathbf{y}$ contains the current observations, and the operator $\mathbf{k}(\cdot)$ represents the analysis algorithm. We use the superscript f rather than b for the background state to avoid confusion later on—it may be helpful to think of *forecast* or *first guess* to associate $\mathbf{x}^f$ with the background state. The input to the analysis consists of the observed-minus-background residuals $\mathbf{y} - \mathbf{h}(\mathbf{x}^f)$, also known as the innovations, data departures, etc.

Note that both $\mathbf{h}$ and $\mathbf{k}$ may be nonlinear, but we can define their linearized versions:

$$\mathbf{H} = \left.\frac{\partial \mathbf{h}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}^f}, \qquad \mathbf{K} = \left.\frac{\partial \mathbf{k}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}^f} \tag{2}$$

Here $\mathbf{K}$ is the familiar *gain matrix* from linear estimation theory. We also define the analysis/background/observation errors by

$$\mathbf{e}^a = \mathbf{x}^a - \mathbf{x}, \qquad \mathbf{e}^f = \mathbf{x}^f - \mathbf{x}, \qquad \mathbf{e}^o = \mathbf{y} - \mathbf{h}(\mathbf{x}) \tag{3}$$

We then obtain the following three diagnostic relations:

$$\left\langle \mathbf{y} - \mathbf{h}(\mathbf{x}^f) \right\rangle \approx \langle \mathbf{e}^o \rangle - \left\langle \mathbf{H}\mathbf{e}^f \right\rangle \tag{4}$$

$$\left\langle \mathbf{x}^a - \mathbf{x}^f \right\rangle \approx \langle \mathbf{K}\mathbf{e}^o \rangle - \left\langle \mathbf{K}\mathbf{H}\mathbf{e}^f \right\rangle \tag{5}$$

$$\langle \mathbf{e}^a \rangle \approx \langle \mathbf{K}\mathbf{e}^o \rangle + \left\langle [\mathbf{I} - \mathbf{K}\mathbf{H}]\, \mathbf{e}^f \right\rangle \tag{6}$$

where $\langle \cdot \rangle$ may represent any linear averaging method, although when referring to *bias* we usually mean an Eulerian time average. However it is useful to keep in mind that much of what follows can be interpreted different ways depending on the averaging method, i.e., on the choice of ensemble.

The first of these relations, (4), simply shows that the mean observed-minus-background residuals depend on the mean errors in observations and background. Obviously, if a certain observing instrument is biased, then this will be reflected by the residual statistics. Alternatively, non-zero mean departures may indicate biases in the first guess, possibly caused by systematic errors in the assimilating model. In the absence of a true reference, there is no general method for separating observation biases from model-generated biases without introducing independent hypotheses about the model and the data. In some cases it is reasonable to suppose that biases in the departures are caused by the model; e.g., when departures for different and independent observing systems show similar large-scale biases that can be explained by known deficiencies in the model.

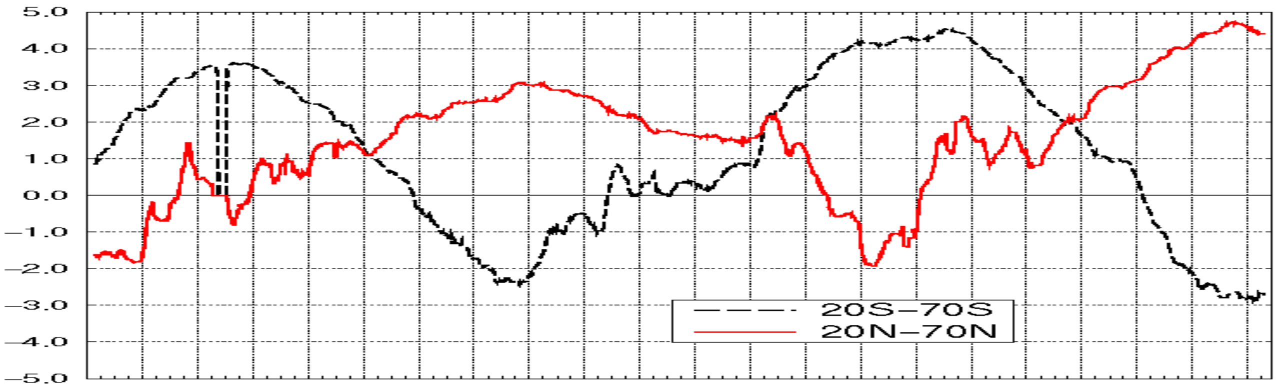

Figure 2: November 1998 – June 2000 monthly mean observed-minus-background residuals for AMSU-A channel 14, northern hemisphere (red) and southern hemisphere (black). Figure provided by Tony McNally, ECMWF.

To illustrate, Fig. (2) shows two timeseries of monthly mean observed-minus-background residuals for Advanced Microwave Sounding Unit (AMSU-A) channel 14 brightness temperatures, averaged over northern hemisphere (20N–70N) and southern hemisphere (20S–70S) locations. Both curves clearly show an annual cycle with a change in the mean of about 5–7 K. This particular channel is most sensitive to temperatures at around 1 hPa. In this case the culprit is almost certainly the model, since it is known to exhibit large, seasonally dependent temperature biases in the stratosphere, and some independent observational evidence from other instruments confirms that this AMSU channel is accurate (McNally, *pers. comm.*).

The second diagnostic equation, (5), states that model and/or observation biases also show up in the mean corrections to the background that are produced by the analysis, i.e., in the analysis increments. In the absence of biases the mean analysis increments should be negligible compared to the typical magnitude of any single increment (e.g., as measured by the root-mean-square). Due to the presence of $\mathbf{K}$ on the right-hand side of (5), it is not a simple matter to infer the magnitude of model and/or observation biases from the magnitude of the mean analysis increments. In a multivariate analysis system, for example, model temperature biases can cause non-zero mean wind analysis increments even in the absence of wind observations. Mean departures are therefore more informative than mean analysis increments for identifying and estimating biases.

Figure (3) shows monthly mean temperature analysis increments during a three-year period from an assimilation made at the ECMWF in preparation for the ERA-40 reanalysis production. During this period the major source of observational information at levels higher than 10 hPa consisted of radiance data from the stratospheric sounding unit (SSU) on board the NOAA-14 satellite. In this preliminary experiment, the SSU data were not used during a three-month period in the beginning of 1995. The corresponding jump of about 0.5 K in the mean temperature increments is clearly evident; this jump is strictly artificial and caused by the presence of a persistent warm bias in the model at levels around 2–5 hPa. The increments at 1 hPa and higher are clearly related to this as well, and are probably due to the vertical structure functions used in the analysis.

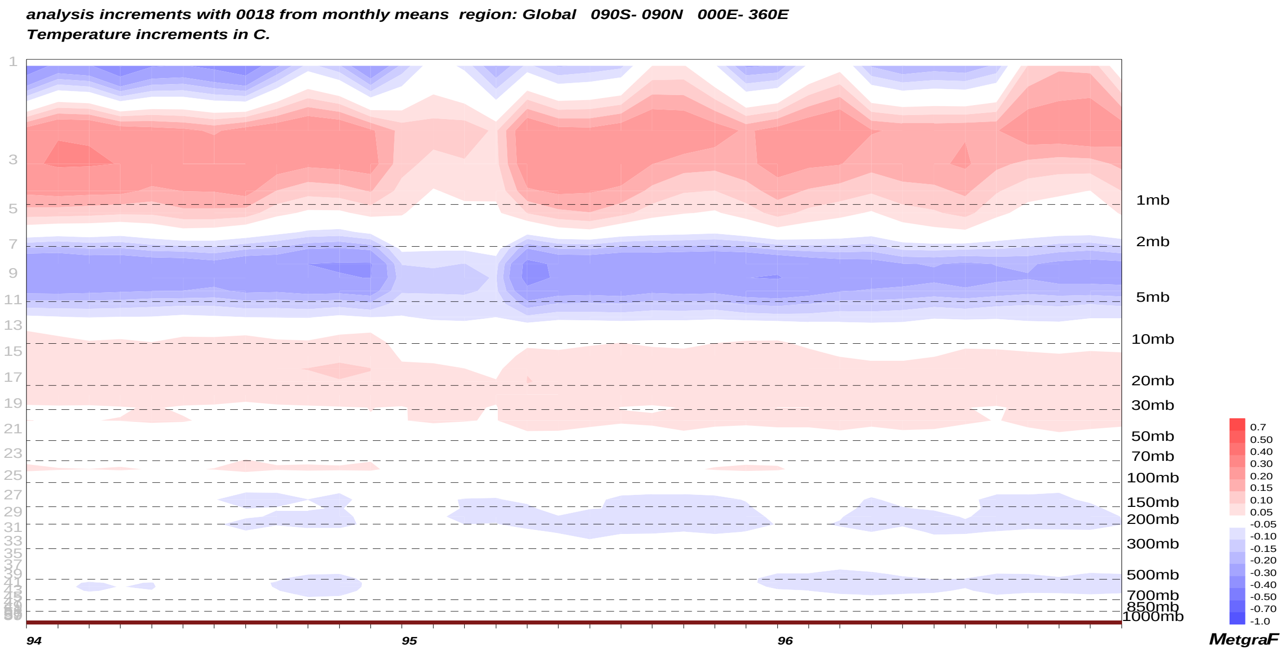

Figure 3: Globally averaged monthly mean temperature analysis increments during 1994–1996, obtained from a preliminary ERA-40 run. No SSU data were used during a three-month period in the beginning of 1995. Figure obtained from Per Kållberg, ECMWF.

Finally, (6) states that the presence of bias in the first guess and/or the observations generally results in biased analyses. This is true for any linear analysis scheme of the form (1); there is no reason to believe that nonlinear implementations of (1) change this fact unless they have been specifically designed to do so. Therefore the analysis method must be modified to account for bias explicitly in order to produce an unbiased assimilation.

3 Correcting persistent errors in the first guess

In some cases (e.g. upper stratospheric temperature, tropical tropospheric humidity) short-term model forecasts contain relatively large, persistent components of error. Instead of zero-mean background errors it is then more appropriate to assume

$$\mathbf{e}^f = \mathbf{b} + \tilde{\mathbf{e}}^f \qquad \text{with} \qquad \langle \tilde{\mathbf{e}}^f \rangle = 0 \tag{7}$$

where the constant (but otherwise unknown) vector $\mathbf{b}$ represents the effect of model bias. It is not difficult to estimate this component from the available observations. This is accomplished by the following simple modification of the standard linear analysis equations:

$$\hat{\mathbf{b}}_k = \hat{\mathbf{b}}_{k-1} - \mathbf{L}_k \left[\mathbf{y}_k - \mathbf{H}_k(\mathbf{x}_k^f - \hat{\mathbf{b}}_{k-1})\right] \tag{8}$$

$$\mathbf{x}_k^a = (\mathbf{x}_k^f - \hat{\mathbf{b}}_k) + \mathbf{K}_k \left[\mathbf{y}_k - \mathbf{H}_k(\mathbf{x}_k^f - \hat{\mathbf{b}}_k)\right] \tag{9}$$

where the subscript k denotes time. The first equation updates a prior bias estimate $\hat{\mathbf{b}}_{k-1}$ based on the latest observations $\mathbf{y}_k$, while the second produces an analysis based on the bias-corrected background $(\mathbf{x}_k^f - \hat{\mathbf{b}}_k)$. Note that omission of $\hat{\mathbf{b}}$ (i.e. setting $\hat{\mathbf{b}} = 0$) gives the standard *bias-blind* analysis equations. It is also worth pointing out that the bias correction (9) is optional; i.e., the bias estimation in (8) can be performed independently as a diagnostic, computed either during the data assimilation or in post-processing.

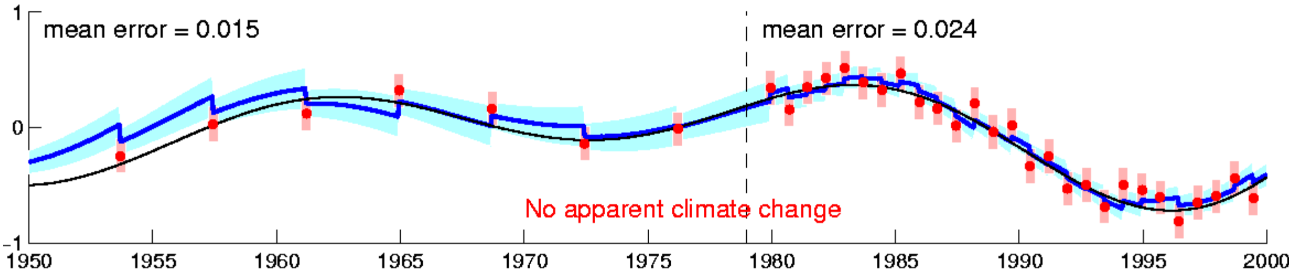

Figure 4: Adaptive model bias correction applied during the assimilation of Fig. (1).

As a trivial illustration we show in Fig. (4) the earlier example of Fig. (1), but now including bias correction on the background. The algorithm learns, after the first few analyses, that the model forecast consistently over-estimates the observation by a fixed amount. It then uses this information to adjust subsequent model predictions. As a result, the mean errors rapidly approach zero and become independent of the observing frequency. The bias correction is adaptive: if the bias were to change, then the algorithm would slowly adjust the estimates based on current observations. Thus, the analysis attempts to make better use of the observations by incorporating the recent history of observed-minus-forecast residuals.

This modified analysis algorithm was developed by Dee and da Silva (1998), who also explained its connection with the more general *separate bias estimator* developed by Friedland (1969), which is based on state augmentation in a Kalman filter framework. Dee and Todling (2000) showed that the algorithm is optimal when

$$\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k\right]^{-1} \tag{10}$$

$$\mathbf{L}_k = \mathbf{P}_k^b \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{P}_k^b \mathbf{H}_k^T + \mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k\right]^{-1} \tag{11}$$

where

$$\mathbf{P}_k^f = \langle \tilde{\mathbf{e}}_k^f (\tilde{\mathbf{e}}_k^f)^T \rangle \tag{12}$$

$$\mathbf{P}_k^b = \left\langle \mathbf{e}_k^b (\mathbf{e}_k^b)^T \right\rangle, \qquad \mathbf{e}_k^b = \hat{\mathbf{b}}_{k-1} - \mathbf{b} \tag{13}$$

Note that $\hat{\mathbf{b}}_{k-1}$ in this scheme represents a prior estimate of the bias at time t_k, and $\mathbf{e}_k^b$ is defined as the error in this estimate. The required covariances are not known; however, a carefully developed model for the background error covariance $\mathbf{P}^f$ is always available in a skillful data assimilation system. In modeling the error covariances $\mathbf{P}^b$ associated with the bias estimates $\hat{\mathbf{b}}$ one can take advantage of this, and use, for example,

$$\mathbf{P}_k^b = \gamma \mathbf{P}_k^f \tag{14}$$

with γ a calibration parameter. As discussed by Dee and Todling (2000), this parameter controls the variation in time of the bias estimates $\hat{\mathbf{b}}$. A small value of γ implies that the estimates are updated slowly, and they will accordingly converge to the long-term time-averaged background error. Large values of γ would result in very noisy bias estimates. An appropriate value can be estimated from observed-minus-background residuals by imposing a constraint on the response function associated with the estimator (8), e.g., one can require that the time sequence of bias-corrected residuals be as close to white as possible. See Dee and Todling (2000) for further details.

Dee and Todling (2000) implemented (8–9) for the moisture component of a global atmospheric data assimilation system, with good success. They used (14) with a value for γ such that the bias estimates varied on a time scale of about 7-10 days. The moisture analysis bias with respect to radiosonde observations was almost completely eliminated. In addition, the bias in the 6-hour forecast was reduced by a factor of two, indicating that the bias correction procedure actually improved the moisture analyses.

4 A simplified version of the algorithm

The extra cost of the bias update in (8) is considerable, since it requires the computation of an additional analysis. The cost may be reduced by using only a subset of the observations for estimating the bias—as was done by Dee and Todling (2000)—or by expressing the bias in terms of a relatively small number of parameters, as we will discuss in the next section. Alternatively, with some approximations the algorithm can be simplified considerably to remove almost all the extra cost of the bias estimation, as follows.

If the model (14) is used with γ small (i.e., slowly varying bias estimates), then the bias updates will be small as well. It may then be acceptable to replace $\hat{\mathbf{b}}_k$ in (9) by $\hat{\mathbf{b}}_{k-1}$. In that case the terms in brackets on the right-hand sides of (8) and (9) are identical. Furthermore, using (14) in (10, 11) implies

$$\mathbf{L}_k = \gamma \mathbf{P}^f \mathbf{H}_k^T \left[(1+\gamma)\mathbf{H}_k \mathbf{P}^f \mathbf{H}_k^T + \mathbf{R}_k\right]^{-1} \tag{15}$$

$$\approx \gamma \mathbf{K}_k \tag{16}$$

when $\gamma \ll 1$. This approximation seems reasonable in view of the many uncertainties that enter into the specification of $\mathbf{K}$ in any realistic setting. We can then reverse the order of (8–9) and obtain the following:

$$\mathbf{x}_k^a = (\mathbf{x}_k^f - \hat{\mathbf{b}}_{k-1}) + \mathbf{K}_k \left[\mathbf{y}_k - \mathbf{H}_k(\mathbf{x}_k^f - \hat{\mathbf{b}}_{k-1})\right] \tag{17}$$

$$\hat{\mathbf{b}}_k = \hat{\mathbf{b}}_{k-1} - \gamma \mathbf{K}_k \left[\mathbf{y}_k - \mathbf{H}_k(\mathbf{x}_k^f - \hat{\mathbf{b}}_{k-1})\right] \tag{18}$$

The bias update equation (18) is now trivial to compute, as it can make use of a previous calculation made in (17). This simplification of the algorithm was first suggested by A. da Silva, and is briefly described in Radakovich *et. al* (2001).

Equation (18) shows that the bias estimates are computed in this algorithm as a recursive time average of the analysis increments. It is important to note that this works only because the analysis in (17) is computed using bias-corrected background estimates. Therefore the two components of this form of the algorithm are interdependent, unlike the original (8–9). In fact, it is clear from (5) discussed earlier that the mean analysis increments in a bias-blind analysis scheme (i.e., a scheme that does not explicitly correct the background bias) are *not* equal to the background bias; see Dee and da Silva 1998, section 2(c), for further discussion.

Any existing data assimilation system can be easily modified to incorporate this algorithm. This becomes clear if we write (17–18) as follows:

$$\tilde{\mathbf{x}} = \mathbf{x}_k^f - \hat{\mathbf{b}}_{k-1} \tag{19}$$

$$\begin{cases} \mathbf{dy} = \mathbf{y}_k - \mathbf{h}_k(\tilde{\mathbf{x}}) \\ \mathbf{dx} = \mathbf{k}_k(\mathbf{dy}) \\ \mathbf{x}_k^a = \tilde{\mathbf{x}} + \mathbf{dx} \end{cases} \tag{20}$$

$$\hat{\mathbf{b}}_k = \hat{\mathbf{b}}_{k-1} - \gamma\mathbf{dx} \tag{21}$$

The central section, (20), represents the original bias-blind analysis scheme. The modifications needed to handle bias estimation and correction are external and can be implemented at a high level in the code—perhaps even at script level.

The simplified algorithm assumes the covariance model (14). This implies that the bias updates (i.e., the updates to the bias estimates) are restricted to the same subspace as are the analysis increments. If this subspace is stationary (which is not exactly true, in practice) then the bias estimates will be in this subspace as well, and therefore the spatial and multivariate structure of the analysis corrections with and without bias correction will be similar. This may be advantageous, especially if the covariances for the original bias-blind analysis system have been carefully constructed to generate balanced analysis increments. However, if it is known *a priori* that the structure of the persistent errors in the system is different than that of the random errors, then this algorithm is not appropriate.

Different refinements and extensions to the algorithm are possible. In particular, instead of applying the bias correction directly to the background as in (19), one can attempt to use the latest bias estimate to adjust the model forcing during the integration, in order to guide it to an unbiased state. This would render the step (19) unnecessary. The simplest way to accomplish this is with a linear incremental updating scheme (as formulated by Bloom *et al.* 1996), in which a fraction of the bias estimate $\hat{\mathbf{b}}$ is removed from the model output after each model time step. Specifically, the biased model integration from t_{k-1} to t_k in N time steps

$$\mathbf{x}_k^f = \mathbf{m}_{k,k-1}(\mathbf{x}_{k-1}^a) \tag{22}$$

$$= \mathbf{m}_{k,k-\frac{1}{N}}(\mathbf{m}_{k-\frac{1}{N},k-\frac{2}{N}}(\quad \cdots \quad (\mathbf{m}_{k-\frac{N-1}{N},k-1}(\mathbf{x}_{k-1}^a)) \quad \cdots \quad)) \tag{23}$$

is replaced with

$$\tilde{\mathbf{x}}_k^f = \widetilde{\mathbf{m}}_{k,k-1}(\mathbf{x}_{k-1}^a) \tag{24}$$

$$= \mathbf{m}_{k,k-\frac{1}{N}}(\mathbf{m}_{k-\frac{1}{N},k-\frac{2}{N}}(\quad \cdots \quad (\mathbf{m}_{k-\frac{N-1}{N},k-1}(\mathbf{x}_{k-1}^a) - \frac{1}{N}\hat{\mathbf{b}}_{k-1}) \quad \cdots \quad) - \frac{1}{N}\hat{\mathbf{b}}_{k-1}) - \frac{1}{N}\hat{\mathbf{b}}_{k-1} \tag{25}$$

This *incremental bias correction* replaces (19). See Radakovich *et al.* (2001) for an application of such a scheme in a land-data assimilation study. This is a first step toward using the bias estimates to modify the model itself, rather than applying a correction to the model-generated background. In a complex multivariate system a more sophisticated approach may be needed, such as proposed by Bell *et al.* (2001) in the context of an ocean model. They combined a statistical bias estimation scheme similar to ours with hypotheses about the physical mechanisms that caused the bias, and then adjusted specific model terms to suppress the bias generation.

5 Parameterized deterministic error models

The term *bias* usually means an error which is constant in time. More generally we might consider any deterministic component of error that can be described in terms of a set of parameters β that are constant (or slowly varying) in time. For example, it is not unusual for temperature errors in a short-term forecast near the surface to depend systematically on the local time of day, due to the model's difficulty in representing the diurnal cycle. In that case it would be appropriate to assume a specific temporal dependence of the background errors. For other, more practical reasons, it may be desirable to reduce the number of degrees of freedom in a constant bias term by presuming, for example, a fixed spatial structure, or by expressing it in terms of a limited number of empirical orthogonal functions or in some other appropriate basis.

In any case, let us now assume the following model for the background errors:

$$\mathbf{e}^f = \mathbf{b}(\beta) + \tilde{\mathbf{e}}^f \qquad \text{with} \qquad \left\langle \tilde{\mathbf{e}}^f \right\rangle = 0 \tag{26}$$

where $\mathbf{b}$ is a known function of a vector β of unknown parameters. This approach becomes especially interesting when the number of bias parameters is much smaller than the dimension of $\mathbf{b}$. In general, $\mathbf{b}$ may depend on location and/or time, or even on the state $\mathbf{x}$.

The usual technique for deriving an estimation algorithm is to relate the unknowns to the available data. Omitting time subscripts for the moment, we have

$$\mathbf{y} - \mathbf{h}(\mathbf{x}^f) \approx \mathbf{e}^o - \mathbf{H}\mathbf{e}^f \tag{27}$$

$$= \mathbf{e}^o - \mathbf{H}\mathbf{b}(\beta) - \mathbf{H}\tilde{\mathbf{e}}^f \tag{28}$$

If we define

$$\mathbf{g}(\beta) = \mathbf{H}\mathbf{b}(\beta) \qquad \text{and} \qquad \tilde{\mathbf{e}} = \mathbf{e}^o - \mathbf{H}\tilde{\mathbf{e}}^f \tag{29}$$

then

$$\mathbf{dy} = \mathbf{g}(\beta) + \tilde{\mathbf{e}} \qquad \text{with} \qquad \begin{cases} \langle \tilde{\mathbf{e}} \rangle \approx 0 \\ \left\langle \tilde{\mathbf{e}}\tilde{\mathbf{e}}^T \right\rangle \approx \mathbf{H}\mathbf{P}^f\mathbf{H} + \mathbf{R} \end{cases} \tag{30}$$

This defines a *measurement model* for β, which describes the information about the parameters β that is implicit in the observations.

There are many ways to derive estimation algorithms for β based on (30). For example, the bias parameters can be added to the control vector in a variational analysis scheme and estimated simultaneously with the state of the system, as is sometimes done for parameters that describe observation bias (Derber and Wu 1998). Similarly, they can be added to the state vector in a Kalman filter (Friedland 1969). If the number of parameters is sufficiently small, then they can simply be estimated from the data by nonlinear least squares, i.e., by solving

$$\hat{\beta}_k = \arg\min_{\beta} ||\mathbf{dy}_k - \mathbf{g}_k(\beta)||^2 \tag{31}$$

prior to the analysis at time t_k. The optimal estimate of β, given the observations at t_k, would require replacing the L^2 norm in (31) by a covariance-weighted norm using the covariances specified in (30), but this is expensive and would probably not change the result very much.

During data assimilation, the parameter estimates may be used to correct the deterministic component of background error $\mathbf{b}(\beta)$. For example, if $\mathbf{b} = \mathbf{b}(\mathbf{x}; \beta)$ then $\mathbf{b}$ can be evaluated at any given time and location by using the latest estimates for $\mathbf{x}$ and β. In the special case when $\mathbf{b}$ is linear in β, i.e.,

$$\mathbf{b}(\beta) = \mathbf{B}_k\beta \tag{32}$$

with $\mathbf{B}_k$ a matrix with known (possibly state-dependent) coefficients, then the analogue of (8–9) is

$$\hat{\beta}_k = \hat{\beta}_{k-1} - \mathbf{L}_k^\beta \left[\mathbf{y}_k - \mathbf{H}_k(\mathbf{x}_k^f - \mathbf{B}_k\hat{\beta}_{k-1})\right] \tag{33}$$

$$\mathbf{x}_k^a = (\mathbf{x}_k^f - \mathbf{B}_k\hat{\beta}_k) + \mathbf{K}_k \left[\mathbf{y}_k - \mathbf{H}_k(\mathbf{x}_k^f - \mathbf{B}_k\hat{\beta}_k)\right] \tag{34}$$

This estimator is optimal when $\mathbf{K}_k$ is as given by (10) and

$$\mathbf{L}_k^\beta = \mathbf{P}_{k-1}^\beta \mathbf{B}_k^T \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{B}_k \mathbf{P}_{k-1}^\beta \mathbf{B}_k^T \mathbf{H}_k^T + \mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k\right]^{-1} \tag{35}$$

where

$$\mathbf{P}_k^\beta = \left\langle \mathbf{e}_k^\beta (\mathbf{e}_k^\beta)^T \right\rangle, \qquad \mathbf{e}_k^\beta = \hat{\beta}_{k-1} - \beta \tag{36}$$

The gain matrix $\mathbf{L}_k^\beta$ for the bias update is low-rank: it has one row for each bias parameter and as many columns as there are observations. Since $\mathbf{B}$ is known, only the parameter estimation covariances in $\mathbf{P}^\beta$ (in addition to $\mathbf{P}$ and $\mathbf{R}$) need to be specified in order to implement this algorithm. If the number of parameters is much smaller than the number of available observations then the specification of $\mathbf{P}^\beta$ becomes less important.

Chepurin *et al.* (2003) have used a version of this algorithm for bias correction in a tropical ocean model. They expressed the bias in the ten-day forecast as a linear combination of a constant field, an annual cycle component, and two empirical orthogonal basis functions. They were then able to estimate the coefficients of these components and correct the forecast bias during the data assimilation.

6 Some pitfalls

The Achilles' heel of this approach is that it requires unbiased observations—or, rather, that any remaining observation bias must be small compared to the bias in the model. As we pointed out, in practice it takes a great deal of work to justify this position. Prior to analysis, observations are thoroughly screened and bias-corrected using a variety of methods and techniques. The vast majority of available observations are actually removed during pre-processing and never make it to the analysis. If and when an observation is finally used, it is presumed to be sufficiently accurate to have a positive impact on the assimilation. The analysis increments in a bias-blind analysis scheme are in fact treated as corrections to the model. It therefore seems reasonable, or at least logically consistent, to view any remaining non-random component of the analysis increments as resulting from model errors. In practice, of course, it is necessary to constantly reassess the underlying assumption of unbiased observations by looking for indications that observation bias is wrongly attributed to the model.

One such indication would be that the bias correction causes a deterioration of forecast skill. It is also useful to compare the magnitude of the bias estimates produced by the algorithm with estimates of background bias in a bias-blind control experiment. The latter can be directly estimated from the observed-minus-background residuals, or by performing an off-line diagnostic bias estimation using (8). If the estimated bias is indeed due to the background and not to the observations, then its removal from the background should result in a better analysis, and hence improve the subsequent forecast. Therefore the bias estimates in an assimilation with bias correction should be smaller than the estimated background bias in the original, bias-blind assimilation. If this

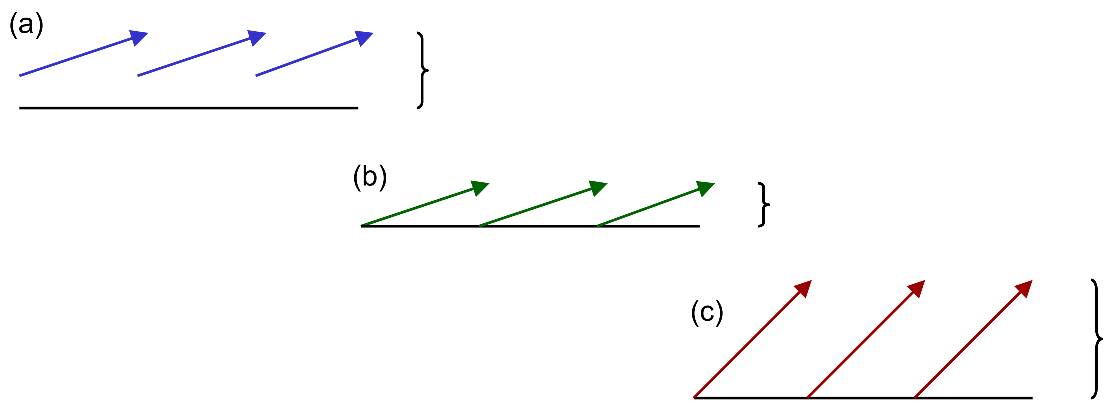

Figure 5: Possible scenarios for the outcome of applying model bias correction during data assimilation: (a) No bias correction; (b) Bias correction resulting in reduced forecast bias; (c) Bias correction resulting in increased forecast bias.

is not the case, then the bias correction actually leads to a worse analysis, perhaps because the bias should be attributed to the observations, rather than to the background. These possible scenarios are summarized in Fig. (6).

There may be reasons other than the presence of observation bias that cause the algorithm to fail. Even when the estimated bias is rightly attributed to the background, the algorithm still does not remove the source of the bias. For a model which is not perfect there is no guarantee that an unbiased analysis leads to a better forecast. There may be mechanisms in the model that rapidly force it to a preferred (biased) trajectory. Intermittent bias correction may provide the model with an ‘unfriendly’ initial state, i.e. one that causes rapid spinup/spindown. Possibly this can be prevented by an incremental correction scheme during the model integration, but it may be that only internal changes to the model can solve this problem.

7 Conclusion

We have presented some techniques for estimating and correcting systematic components of background errors during data assimilation. The approach is strictly statistical, i.e., it tries to make the best use of the available observations to correct errors in the background fields, without addressing the mechanisms that cause the errors in the first place. Clearly it would be preferable to remove the source of the errors rather than merely treating symptoms, but this can be said about any analysis system. It seems only logical that an analysis algorithm be designed to correct all error components that can be identified from the available data.

References

Bell, M. J., M. J. Martin, and N. K. Nichols, 2003: Assimilation of data into an ocean model with systematic errors near the equator. *Q. J. R. Meteorol. Soc.*, in press.

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271.

Chepurin, G. A., J. Carton, and D. Dee, 2003: Forecast model bias correction in ocean data assimilation. *Mon. Wea. Rev.*, submitted.

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

Dee, D. P., and R. Todling, 2000: Data assimilation in the presence of forecast bias: The GEOS moisture analysis. *Mon. Wea. Rev.*, **128**, 3268–3282.

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

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

Friedland, B., 1969: Treatment of bias in recursive filtering. *IEEE Trans. Autom. Contr.*, **AC-14**, 359–367.

Harris, B. A., and G. Kelly, 2001: A satellite radiance-bias correction scheme for data assimilation. *Q. J. R. Meteorol. Soc.*, **127**, 1453–1468.

McNally, A. P., J. C. Derber, W.-S. Wu, and B. B. Katz, 2000: The use of TOVS level-1b radiances in the NCEP SSI analysis system. *Q. J. R. Meteorol. Soc.*, **126**, 689–724.

Radakovich, J. D., P. R. Houser, A. M. da Silva, and M. G. Bosilovitch, 2001: Results from global land-surface data assimilation methods. *5th Symposium on Integrated Observing Systems, Albuquerque, New Mexico*.