On the robustness of methods to account for background bias in DA to uncertainties in the bias estimates

Alison Fowler (University of Reading, National Centre for Earth Observation)

Introduction

Fundamental to the theory of data assimilation (DA) is that the data are an unbiased estimate of the true state. Often this assumption is far from valid and, without bias correction, the resulting analysis will be biased. Two methods to account for biases in the background that do not require a change to the DA algorithm, are compared: explicit bias correction (BC) and covariance inflation (CI). Both methods rely on an estimate of the background bias. Given the difficulties in estimating the background bias, the robustness of the two methods in producing an unbiased analysis is studied within an idealised linear system

Accounting for background biases

The background biases, β_b^i, may be caused by systematic errors in the numerical model propagating the analysis from the previous assimilation cycle to become the background at the current assimilation cycle, b, as well as biases in the previous analysis, β_a^{i-1}, propagated by the model, $\mathcal{M}$:

$$\beta_\mathrm{b}^i = \mathcal{M}(\beta_\mathrm{a}^{i-1}) + \mathrm{b}.$$

When the biases are known there are two general approaches that could be taken to reducing the bias in the analysis, which both avoid changing the DA algorithm:

- **Explicit bias correction (BC)** removes the bias from the background before assimilation without attempting to correct the source of the bias.
- **Covariance inflation (CI)** increases the weight given to the (presumably) unbiased observations to give the analysis with the smallest root-mean square error [1].

$$\tilde{\mathrm{B}} = \mathrm{B} + \beta_\mathrm{b}\beta_\mathrm{b}^\mathrm{T}.$$

If the biases are known exactly then BC is the most optimal approach to providing an analysis that is unbiased. However, in practice, accurate estimates of the background bias are limited to where high-quality, unbiased observations are available. In order to reduce sampling error, assumptions about ergodicity and homogeneity need to be imposed, which limit the amount of detail that can be provided about how the bias varies in space and time.

Numerical experiments

Idealised model

To illustrate the ability of the BC and CI methods to give an unbiased analysis we set up an idealised linear system in which, without any methods to control the bias, the background error covariance and background bias remain constant as the assimilation system is cycled, i.e. $\beta_\mathrm{b}^{i+1} = \beta_\mathrm{b}^i$ and $trace(\mathrm{B}^{i+1}) = trace(\mathrm{B}^i)$ (1).

In these experiments the specified form of β_b^i and B^i are illustrated in figure 1. Direct observations of each model variable are assimilated with error variance of 5. The linear model and model bias that this implies in order to satisfy (1) are also plotted in figure 1.

Contact information ● Department of Meteorology, University of Reading, Whiteknights, RG6 6AH ●
Email: a.m.fowler@reading.ac.uk ● Web: http://www.met.reading.ac.uk/~gj901587

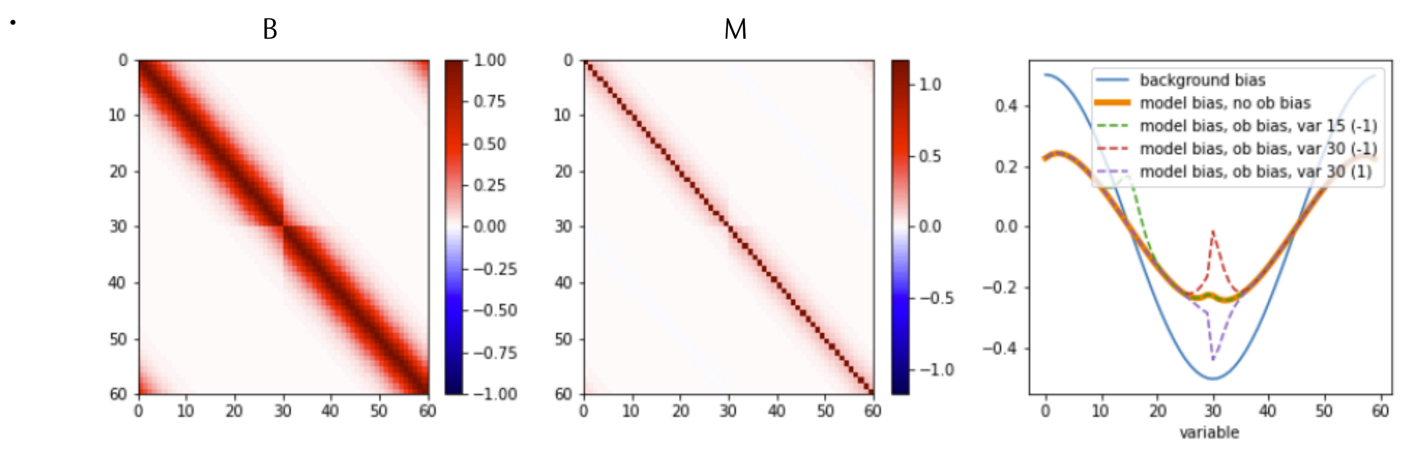

Figure 1 Experiment setup. Left: given B-matrix and Middle: resulting linear model. Right: given background bias (blue) and resulting model bias when; observations are unbiased (orange), one instrument measuring variable 15 has a bias of -1 (dashed green), one instrument measuring variable 30 has a bias of -1 (dashed red), and one instrument measuring variable 30 has a bias of 1 (dashed purple).

Estimating the background bias

The background bias is estimated from a sample of 10000, 100 and 10 innovations (observation minus background) at each grid point. In practice, these sample sizes are unrealistically large. The choice of 10000 is to allow for the performance of the schemes to be compared when the sampling error is negligible. The sample size of 100 and 10 may potentially be obtained from a time-series of observations, with the assumption that the background bias is sufficiently constant over the sample.

To demonstrate the effect of parameterising the bias, estimates of the bias for each grid box are compared to estimating the bias as a constant over 10 grid boxes (reducing the number of parameters from 60 to 6 and increasing the sample size for each bin 10 times). Reducing the number of bins gives a discontinuous estimate of the bias so we also compare to a smoothed version of 6 bins (see fig. 2).

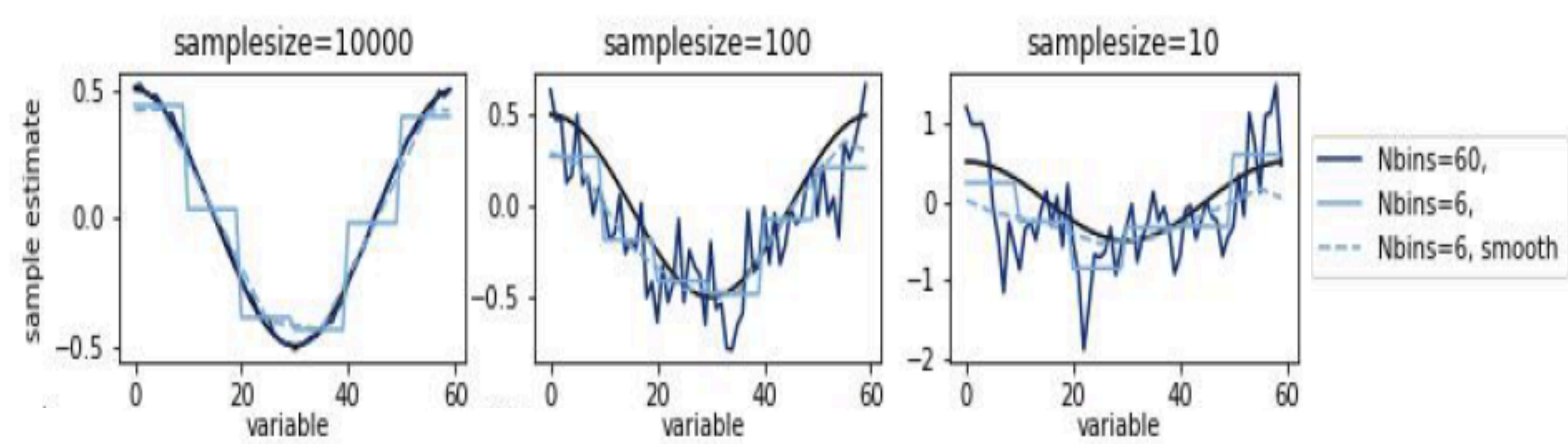

Figure 2: Sample approximations to the true initial background bias (black line) for sample sizes of 10000 (left), 100 (middle) and 10 (right).

Results

Background bias known perfectly but observation bias is uncorrected

Figure 3 shows the analysis bias as a function of variable for the tenth assimilation cycle when different biases in the observations are present but the background bias is known perfectly when applying BC and CI. We can conclude:

- If no observation is present then BC method gives an unbiased analysis
- If an observation is biased then the impact on the analysis depends on the mean innovation. Overall the CI method is most sensitive to the observation biases but does not always result in the largest analysis bias..

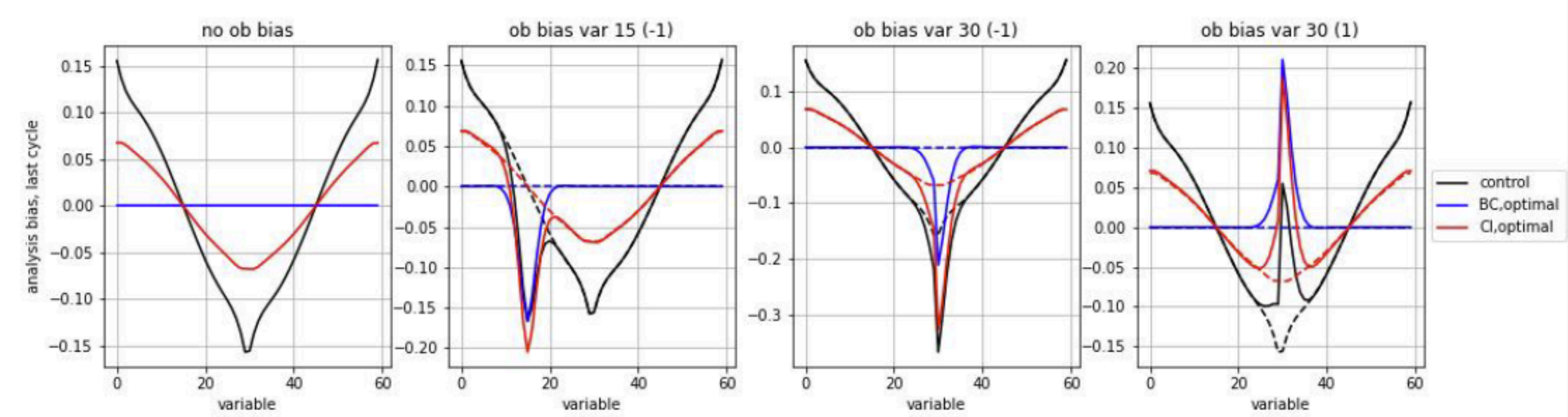

Figure 3: *Background bias perfectly known*: Analysis bias as a function of variable for the tenth assimilation cycle for when the background bias is unaccounted for (black), and accounted for using BC (blue) and CI (red). Panels give results for different observation biases.

Accounting for background bias when the bias is not known perfectly

Figure 4 shows the analysis bias as a function of variable for the tenth assimilation cycle when the observations are unbiased but different approximations to the background are used when applying BC and CI. We can conclude:

- CI is seen to be much more robust to errors in the bias estimate than the BC method, both in terms of sample noise and structural errors in the bias estimate.
- The effect of the noisy estimate of the background bias in the BC method is magnified as the system is cycled. Smoothing the bias estimate is therefore essential for the BC method.

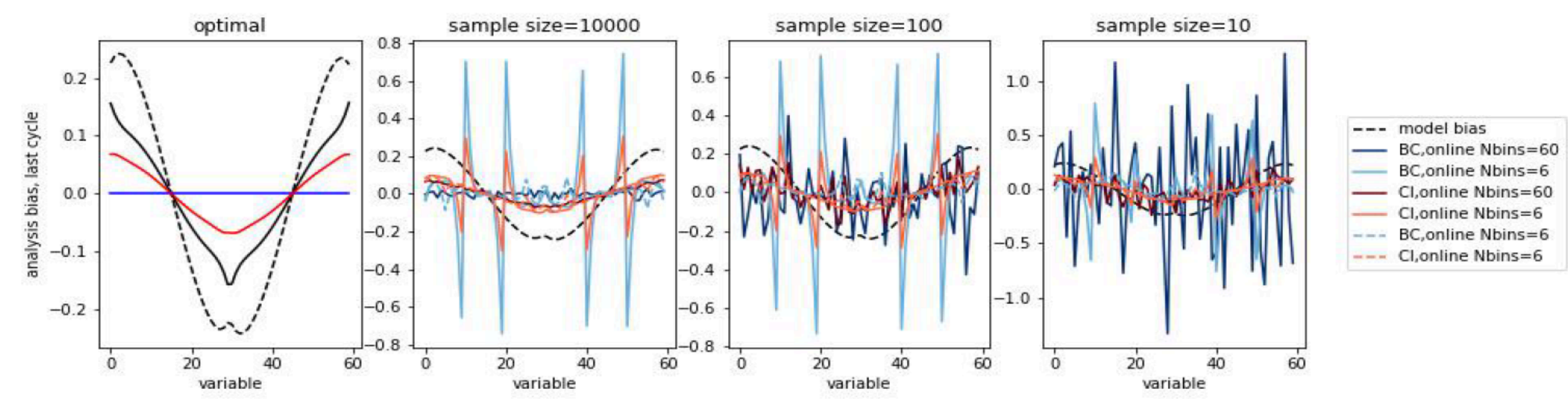

Figure 4: *Background bias estimated from sample of innovations cf. fig 2*: Analysis bias as a function of variable for the tenth assimilation cycle for when the background bias is unaccounted for (black), and accounted for using BC (blue) and CI (red). Panels give results for different sample sizes when estimating the background bias.

Summary

Two methods are used to correct background bias: covariance inflation (CI) and explicit bias correction (BC). CI is found to be more robust in dealing with uncertainty in the background bias by allowing observations to correct for background bias while increasing the analysis uncertainty. However, altering the covariance structure is crucial for the success of this method so it cannot be approximated by variance inflation alone. The CI method is also more sensitive to biased observations.

Another method to correct background/model bias is Weak constraint 4DVar (WC4DVar). WC4DVar acknowledges the uncertainty in the bias estimate but is difficult to implement due to the requirement for a specific DA algorithm and will still be sensitive to assumptions about the structure of the bias [2].

References [1] Dee and Da Silva. QJRMS, 1998. [2] Bonavita and Laloyaux, JAMES, 2020.