Bias correction of satellite radiance observations in the Met Office NWP system

Ruth Taylor

December 20, 2019

The Met Office has used a variational bias correction scheme, known as VarBC, for all satellite radiances assimilated operationally since March 2016. VarBC, described in detail by Cameron and Bell (2018), is an adaptive scheme which follows the incremental formulation described by Auligné et al. (2007). A bias correction is modelled for each observation as a linear combination of a set of bias predictors, which typically depend on atmospheric state $\mathbf{x}$ and on other circumstances of the observation, with a coefficient for each predictor. A vector of m observations $\mathbf{y^o}$ will have a corresponding vector of m bias corrections $\mathbf{b^o}(\mathbf{x}, \boldsymbol{\beta})$, in which the kth element is given by

$$b_k = s_k + \sum_{i=1}^{I_k} \beta_i p_{k,i}(\mathbf{x}) \tag{1}$$

where the subscript i indexes the I_k predictors $p_{k,i}$ and the corresponding β_i form a vector of bias correction coefficients $\boldsymbol{\beta}$. The s_k allow the bias correction to retain a static (invariant) component, for reasons discussed below.

In VarBC the predictor coefficients $\boldsymbol{\beta}$ are included in the 4D-Var control vector and allowed to vary in the course of minimisation. They appear in the cost function J both in the observation term, via the correction given by the b_k of Equation 1, and in an additional term penalising departures of the coefficients from their background values $\boldsymbol{\beta}^\mathbf{b}$, so that

$$\begin{aligned} J(\mathbf{x}, \boldsymbol{\beta}) = \frac{1}{2}(\mathbf{x} - \mathbf{x^b})^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x^b}) \\ + \frac{1}{2}(\mathbf{y^o} - \mathbf{b^o}(\mathbf{x}, \boldsymbol{\beta}) - H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y^o} - \mathbf{b^o}(\mathbf{x}, \boldsymbol{\beta}) - H(\mathbf{x})) \\ + \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\beta}^\mathbf{b})^T \mathbf{B}_\beta^{-1}(\boldsymbol{\beta} - \boldsymbol{\beta}^\mathbf{b}) \end{aligned} \tag{2}$$

where	$\mathbf{x^b}$	is the background value of the atmospheric state
	$\mathbf{B}$	the error covariance of the background
	$H(\mathbf{x})$	the vector of forward-modelled observations equivalent to the state $\mathbf{x}$
	$\mathbf{R}$	the error covariance of the observations (and their operator)
and	$\mathbf{B}_\beta$	the error covariance of the prior estimate of the bias predictor coefficients.

Note that Equation 1 is given in a general form, allowing each observation to have its own unique set of bias predictors. In the Met Office implementation, observations are grouped by channel, with a set of predictors for each channel on each instrument. The vector $\boldsymbol{\beta}$ contains a set of coefficients for each channel.

The formulation given for J in Equation 2 neglects any correlations between the background errors for the bias coefficients $\boldsymbol{\beta}$ and for the state $\mathbf{x}$. Further, $\mathbf{B}_\beta$ itself is taken to be diagonal, with error variances V_{β_i} derived by considering those elements of J's Hessian matrix which correspond to the bias coefficients. Note that in estimating $\mathbf{B}_\beta$ as described below we use the diagonal elements of $\mathbf{R}$, but the assimilation itself may still use a full $\mathbf{R}$ to take into account inter-channel correlations. The second and third terms of the penalty J can be rewritten with explicit summation over the m observations. For observations in the channel corrected using coefficients β_i and β_j, the contribution to J is

$$\frac{1}{2}\sum_k \big(y_k^o - (s_k + \sum_{i=1}^{I_k}\beta_i p_{k,i}) - H(\mathbf{x})\big) R_k^{-1} \big(y_k^o - (s_k + \sum_{j=1}^{I_k}\beta_j p_{k,j}) - H(\mathbf{x})\big) + \frac{1}{2}\sum_{i=1}^{I_k} V_{\beta_i}^{-1}(\beta_i - {\beta^b}_i)^2 \quad (3)$$

Double differentiation of (3) then gives the corresponding elements of the Hessian as

$$\frac{\partial^2 J}{\partial\beta_i \partial\beta_j} = V_{\beta_i}^{-1}\delta_{i,j} + \sum_{k=1}^{m} p_{k,i} R_k^{-1} p_{k,j} \quad (4)$$

where $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise, from which the diagonal elements of the Hessian can be obtained:

$$\frac{\partial^2 J}{\partial\beta_i^2} = V_{\beta_i}^{-1} + \sum_{k=1}^{m} R_k^{-1} p_{k,i}^2 \quad (5)$$

The ratio of the two terms on the right-hand side of Equation 5 determines the sensitivity of the bias coefficients to the observations in the current assimilation cycle. In the Met Office implementation of VarBC, the $V_{\beta_i}^{-1}$ are set for each cycle to a multiple of the second term of Equation 5, using a scaling factor N_{bgerr}/m:

$$V_{\beta_i}^{-1} = \frac{N_{bgerr}}{m}\sum_{k=1}^{m} R_k^{-1} p_{k,i}^2 \quad (6)$$

Note the summations in Equations 4 - 6 are over those observations with a bias correction including the ith bias predictor, that is, over observations in a given channel. N_{bgerr} is channel-specific, and is calculated as

$$N_{bgerr} = \mathtt{MAX}(m_{avg}, M_{min})\left(\frac{1}{2^{\frac{1}{n}} - 1}\right) \quad (7)$$

where m_{avg} is the number of contributing observations expected per data assimilation cycle, and M_{min} ensures that N_{bgerr} has a lower limit. n is a tunable dimensionless parameter: N_{bgerr}, and thus $V_{\beta_i}^{-1}$, increase with n, increasing the penalty for varying $\boldsymbol{\beta}$ from the prior $\boldsymbol{\beta}^{\mathbf{b}}$, and decreasing the sensitivity of the bias coefficients to observations in the cycle. With the dependence formulated in Equation 7, n (expressed as a multiple of the data assimilation cycle length) is referred to as a "bias halving time", because of the way it characterises the response of the data assimilation system to an externally-imposed change in bias. Cameron and Bell (2018) give a derivation of Equation 7, as well as some illustrations of its operation.

At the Met Office, the m_{avg} are based on recent statistics for the channel in question, and are carried forward between cycles using the same mechanism as is used to carry the background bias coefficients analysed for one cycle forward to become background (and first-guess) values for the next. The bias halving time n should be chosen so as to strike a balance between responsiveness, via adaptation rate, on the one hand, and proper sampling of geographical and diurnal variations on the other. To date, the same value of n has been used for all instruments, but the code does allow different adaptation rates to be set for different predictors and for different observation types.

For efficient minimisation, a preconditioning transformation is applied to the control variables with the aim of bringing the form of the Hessian of J closer to the identity matrix I. With transformed variables $\mathbf{v}^{\beta}$, the transformation $\mathbf{U}_{\beta}$ can be defined by

$$\boldsymbol{\beta}' = \mathbf{U}_{\beta}\mathbf{v}^{\beta} \tag{8}$$

where $\boldsymbol{\beta}' = \boldsymbol{\beta} - \boldsymbol{\beta}^{\mathbf{b}}$, that is, a correction to the prior value. The conventional approach, using a transformation of the form $\mathbf{U}\mathbf{U}^T = \mathbf{B}$ (that is, based only on the background term of J) is ineffective here because observation terms can no longer be neglected when bias correction coefficients are considered. $\mathbf{B}_{\beta}$ also has a different structure from $\mathbf{B}$. The Met Office has followed the approach of Dee (2004), grouping observations by channel and deriving the symmetric square root of the Hessian for the bias correction coefficients, which in this case can be computed prior to minimisation. Summing over observations in a given channel to form a covariance matrix $\mathbf{C}$ with elements

$$C_{i,j} = \frac{1}{m}\sum_{k=1}^{m} p_{k,i}p_{k,j} \tag{9}$$

we construct a transformation

$$\mathbf{U}_{\beta} = \left[\mathbf{B}_{\beta}^{-1} + \frac{m}{R}\mathbf{C}\right]^{-1/2} \tag{10}$$

where $\mathbf{B}_{\beta}$ is the background matrix (with diagonal elements V_{β_i}), and the R_k are assumed to have the same value R for every observation in the group. It can be shown that with this $\mathbf{U}_{\beta}$, $(\mathbf{U}_{\beta}\mathbf{U}_{\beta}^T)^{-1}$ is equal to the Hessian of J for the observation group (Equation 4) as required. As for the derivation of $\mathbf{B}_{\beta}$ above, this use of the diagonal elements of $\mathbf{R}$ does not require 4D-Var to be performed with a diagonal $\mathbf{R}$.

Specifics of implementation

The Met Office's Global and limited-area UKV suites' use of VarBC is very similar, including (currently) the choice of two days as a bias halving time in both systems. As the UKV model has both a smaller domain and a shorter assimilation window than the Global model, observation counts are typically lower and M_{min} (Equation 7) is set to a value of 150 for most instruments, as opposed to 1000 in the Global model. Coefficients evolve independently for the two systems.

The predictors used by Met Office VarBC consist of

- A constant, with no spatial dependence. The evolution of this term's coefficient allows its contribution to vary in time.
- Two air masses, 850-300 hPa thickness and 200-50 hPa thickness. These are common to the scheme's static predecessor, described by Harris and Kelly (2001).
- A number of Legendre polynomials in scan position, describing the spatial variation in bias across a scan. Typically four are used, though up to six are available. The variation of the coefficient used with each polynomial allows the bias pattern to adapt with time. These biases are added to a "static" (non-evolving) component for each scan position, the s_k in Equation 1 above, representing residual variations across the scan which cannot be modelled accurately with the relatively small number of Legendre polynomials. Values for the s_k are obtained using time-averaged O-B statistics (see below).
- A set of up to ten Fourier series functions, allowing for variation of bias with orbital angle. This formulation was developed (Booton et al. (2015)) to describe complex residual biases

exhibited by the Special Sensor Microwave Imager/Sounder (SSMIS), and has since also been applied to the Microwave Radiation Imager (MWRI) aboard FY-3C.

Sets of static bias correction coefficients, calculated using O-B and predictor statistics collected over a number of DA cycles, are useful in various circumstances: as well as application of the static component of scan bias mentioned above, a bias correction may be needed for channels that are used in pre-assimilation observation processing (e.g. for quality control, or to obtain surface quantities using 1D-Var) but are not subsequently assimilated, and so cannot have evolving biases. Accurate initial values for VarBC coefficients also help to reduce any spin-up period needed when introducing a new instrument to the assimilation system. Lastly, the Met Office VarBC scheme allows for a subset of channels on a given instrument to have their bias corrections omitted from 4D-Var minimisation, but still to have a static bias correction applied in observation processing. Table 1 lists several window channels on each of the AIRS, IASI and CrIS instruments which have a fixed correction in the Met Office system, preventing slow drifts in the bias correction that were observed during initial testing and that have been attributed to an interaction with the fitting of surface skin temperature (Cameron and Bell (2018)). Similarly, AMSU-A Channel 14 and ATMS Channel 15 (the highest-peaking temperature-sounding channel on each instrument) are not bias corrected at all, acting instead as anchors for the model stratosphere, where there are few other observations. Where static bias corrections are required, coefficients are usually calculated using at least ten days' worth of data and biases are routinely monitored so that they can be updated as required.

Implementation of a method to suppress drifts using a constrained bias correction (for example, by including a term in J penalising departures from a fixed background bias as well as from the previous cycle's background, following Han and Bormann (2016a), Han and Bormann (2016b)) is planned. A further planned development is the selection of observations that contribute to the bias correction; currently everything that is assimilated, including an increasing number of cloud-affected observations, is used to calculate a correction.

Table 1: Table of hyperspectral sounder window channels, with wavenumbers ν, to which fixed bias corrections are applied.

AIRS		IASI		CrIS	
Channel	ν (cm^{-1})	Channel	ν (cm^{-1})	Channel	ν (cm^{-1})
475	801.10	662	810.25	404	901.875
497	809.18	668	811.75	427	916.250
528	820.83	756	833.75	447	928.750
587	843.91	867	861.50	464	939.375
787	917.31	921	875.00	473	945.000
791	918.75	1027	901.50	482	950.625
914	965.43	1133	928.00	484	951.875
950	979.13	1194	943.25	501	962.500
		1271	962.50	529	980.000

References

Auligné, T. et al. (2007). Adaptive bias correction for satellite data in a numerical weather prediction system. *Q. J. R. Meteorol. Soc.*, 133:631–642. doi: 10.1002/qj.56.

Booton, A., Bell, W., and Pavelin, E. (2015). An Improved Bias Correction for SSMIS: Assimilation Assessment. Proceedings of the 20th International TOVS Study Conference, Lake Geneva, Wisconsin, U.S.A.

Cameron, J. and Bell, W. (2018). The testing and implementation of variational bias correction (VarBC) in the Met Office global NWP system. Met Office Forecasting Research Technical Report 631.

Dee, D. (2004). Variational bias correction of radiance data in the ECMWF system. In *Proceedings of ECMWF Workshop on Assimilation of high spectral resolution sounders in NWP*. URL: https://www.ecmwf.int/en/elibrary/8930-variational-bias-correction-radiance-data-ecmwf-system.

Han, W. and Bormann, N. (2016a). Constrained adaptive bias correction for satellite radiance assimilation in the ECMWF 4D-Var system. EUMETSAT NWP-SAF Visiting Scientist report NWPSAF-EC-VS-028. URL: https://www.nwpsaf.eu/vs_reports/nwpsaf-ec-vs-028.pdf.

Han, W. and Bormann, N. (2016b). Constrained adaptive bias correction for satellite radiance assimilation in the ECMWF 4D-Var system. ECMWF Technical Memorandum 783.

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