The importance of the timing of anchor observations in 4D variational bias correction: theory and idealised experiments

Received: 13 August 2024 | Revised: 21 May 2025 | Accepted: 4 June 2025

DOI: 10.1002/qj.5043

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

The importance of the timing of anchor observations in 4D variational bias correction: Theory and idealised experiments

Alison M. Fowler[1,2] | **Devon J. Francis**[1,2] | **Amos S. Lawless**[1,2] | **J. R. Eyre**[3] | **Stefano Migliorini**[3]

[1]School of Mathematical, Physical and Computational Sciences, University of Reading, Reading, UK

[2]National Centre for Earth Observation, Reading, UK

[3]Met Office, Exeter, UK

Correspondence
Alison M. Fowler, Department of Meteorology, University of Reading, Reading, RG6 6ET, UK.
Email: a.m.fowler@reading.ac.uk

Funding information
Natural Environment Research Council, Grant/Award Number: PR140015; UK Research and Innovation, Grant/Award Number: NE/S007261/1

Abstract

Variational bias correction (VarBC), by correcting for significant biases in satellite radiances, is a key component of many modern numerical weather prediction (NWP) systems. However, there is a risk that VarBC may be contaminated by biases present in the assimilating model, inducing a bias in the analysis and in turn reducing forecast skill. Due to the limited reliability of metrics for assessing the value of anchor observations in NWP, this article instead takes the approach of developing and exemplifying new theory to understand how to optimise the impact of anchor observations (assimilated observations with negligible bias) to minimise the contamination of model bias in VarBC. This is important because the number and variety of satellite radiances assimilated is expected to increase in the future. Therefore, the new theory presented can guide how the anchor observation network should also be developed. The new insight may also be crucial in optimising the anchoring effect of historically sparse observations in the context of reanalyses. We present this new theory and theory-driven examples to show that the timing of the anchor observations can impact the accuracy of VarBC substantially. Anchor observations towards the end of the assimilation window provide more information about the accumulated model bias and offer a stronger constraint on VarBC, as such VarBC is more successful at quantifying and correcting the radiance biases only. However, precise anchor observations at the end of the window can increase the contamination of the initial state analysis by model bias. The interaction between the model bias contamination of VarBC and the initial state analysis is studied in idealised cycled data assimilation experiments using the Lorenz 96 model, highlighting the importance of VarBC for an accurate analysis of the state.

KEYWORDS

4DVar, NWP model bias correction, observation network design, satellite radiances, theory of anchor observations

1 | INTRODUCTION

The assimilation of satellite radiances has been demonstrated to have the greatest impact of all observation types on global numerical weather prediction (NWP) skill, due to their unique coverage of key meteorological variables (Cardinali, 2009; Eyre *et al.*, 2022). However, satellite radiances often exhibit significant biases that need to be corrected for them to be beneficial to NWP. Many operational meteorological centres use variational bias correction (VarBC) to correct for biases in the satellite radiances as part of the data assimilation system (Auligné *et al.*, 2007; Dee, 2005; Derber & Wu, 1998). By estimating and correcting for these biases within the data assimilation system, VarBC can make use of both the model and other uncorrected observations as a reference of the unbiased "true" state. However, if the model is also biased and the model bias is uncorrected, then this can contaminate the estimate of the radiance bias. Thus, the corrected radiance observations will be ineffective in counteracting the model bias and unable to draw the analysis towards the truth. Therefore, to ensure that VarBC identifies radiance bias and not model bias, it is known that a good network of observations with negligible bias is important (Eyre, 2016). These are referred to as anchor observations to emphasise the role they play in providing an unbiased anchoring reference for the assimilation system.

The most important anchor observations identified are radiosonde and Global Navigation Satellite System–Radio Occultation (GNSS-RO) observations (Radnóti *et al.*, 2012). The horizontal sampling of radiosonde profile data is relatively sparse in time and space and is mostly confined to populated regions. However, these observations provide a good vertical resolution of information about the temperature, pressure, humidity, and winds up to a height of approximately 20 hPa (Ingleby *et al.*, 2016). They also have the benefit of providing a long record, having been used in NWP for over seven decades (Sun *et al.*, 2013). GNSS-RO observations, introduced in significant numbers in 2006, on the other hand, give good global coverage of temperature information in the stratosphere and upper to mid-lower troposphere, and humidity in the lower troposphere (Buontempo *et al.*, 2008; Cardinali & Healy, 2014). GNSS-RO observations have proven to be particularly beneficial for constraining the large stratospheric temperature biases present in the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS: Poli *et al.*, 2010; Laloyaux *et al.*, 2020b). However, in the context of reanalysis pre-2006, the lack of GNSS-RO data means that the anchoring role of radiosonde data is more important (Hersbach *et al.*, 2020). It should be noted that radiosondes may also have biases, which are usually corrected offline (Haimberger *et al.*, 2012). GNSS-RO also has small biases when assimilated as refractive index or accumulated delay.

Going forward, the observing network is evolving quickly, with the number of satellites and the spectral resolution of the instruments increasing. The variety of information on variables that can be inferred from satellite instruments is also increasing (World Meteorological Organization, 2020). Therefore, as the proportion and variety of satellite radiances assimilated is expected to grow, it is important to understand the requirements of the network of anchor observations that should also be developed.

The development of the observing network is influenced by evidence of the current value of its observations in NWP (e.g., Bouttier & Kelly, 2001; Kull *et al.*, 2021; WMO, 2024). A very useful metric of the value of the observations is known as the forecast sensitivity to observation-based impact (FSOI), which uses the model adjoint to quantify the sensitivity of the forecast skill to the assimilated observations (Langland & Baker, 2004). When a model adjoint is unavailable, the FSOI can be modified to use sensitivity information derived from an ensemble, referred to as the ensemble FSOI (EFSOI: Kalnay *et al.*, 2012). However, the reliability of the FSOI/EFSOI, as well as other metrics for assessing the value of anchor observations, is limited (Candy *et al.*, 2021). This article therefore aims to develop theory to understand what influences the value of anchor observations in NWP.

We present theory-driven examples to show how anchor observations can be used to control the contamination of model bias when the model bias itself is not corrected. We focus on the 4D implementation of VarBC, that is, the observations are distributed in both time and space. The results demonstrate the importance of the precision (defined in terms of error variance), relative timing, and position of the anchor observations. This builds on the previously published work of Francis *et al.* (2023), in which the role of anchor observations was studied in 3D (i.e., the observations are distributed in space only) VarBC and hence the timing of the anchor observations was not considered. This current work also highlights how the timing of the anchor observations will impact the contamination of the model bias in the analysis of the state variables and VarBC in different ways.

In Section 2 we present new theory on how model bias can contaminate the estimate of the radiance bias-correction coefficients in 4D VarBC. In Section 3 we then consider three simplified observing strategies to highlight the importance of the timing of the anchor observations. In Section 4 we illustrate how these theoretical results extend to a cycled assimilation system using numerical experiments with the Lorenz 96 model (Lorenz, 1995) and demonstrate that the FSOI does not quantify the value of the anchor observations reliably

in the presence of uncorrected model bias. Finally, in Sections 5 and 6 we summarise the results and discuss the implications for the development of the network of anchor observations.

Within this current work, we do not consider the case in which the model bias is also explicitly corrected for, for example, using weak-constraint 4DVar (Laloyaux *et al.*, 2020a; Trémolet, 2006) or an ensemble framework (e.g., Li *et al.*, 2009). This is primarily because very few centres have been able to apply this approach successfully, due to the complexity of the model biases, which makes it extremely difficult to provide a prior probability. We will return to this point in the discussion section.

2 | VARIATIONAL BIAS CORRECTION IN THE PRESENCE OF MODEL BIAS

VarBC allows for the online correction of radiance biases by augmenting the state vector, $\mathbf{x}_0 \in \mathbb{R}^n$ (the model state at the beginning of the assimilation window), in 4DVar with bias-correction coefficients, $\boldsymbol{\beta} \in \mathbb{R}^r$ (Auligné *et al.*, 2007). For an introduction to 4DVar, see, for example, Rabier (2005). As in Francis *et al.* (2023) (hereafter referred to as F23), we define the augmented state vector as $\mathbf{v} = (\mathbf{x}_0^{\mathrm{T}}, \boldsymbol{\beta}^{\mathrm{T}})^{\mathrm{T}}$.

The analysis of $\mathbf{v}$ is then found by minimising the 4D VarBC cost function, which finds the best fit to a prior estimate of $\mathbf{v}$ and the available observations:

$$J(\mathrm{v}) = \frac{1}{2}\left(\mathrm{v}-\mathrm{v}^{\mathrm{b}}\right)^{\mathrm{T}}\mathrm{B}_{\mathrm{v}}^{-1}\left(\mathrm{v}-\mathrm{v}^{\mathrm{b}}\right) + \sum_{t=1}^{p_{\mathrm{T}}}\left(\mathrm{y}_t - h_t(m_{0\to t}(\mathrm{x}_0),\boldsymbol{\beta})\right)^{\mathrm{T}}\mathrm{R}_t^{-1}\left(\mathrm{y}_t - h_t(m_{0\to t}(\mathrm{x}_0),\boldsymbol{\beta})\right), \tag{1}$$

where $\mathbf{v}^{\mathrm{b}}$ is the background (first-guess/prior) estimate of the augmented state vector. The background estimate for $\mathbf{x}_0$ is typically given by a forecast initialised by the previous analysis, whereas the background estimate for $\boldsymbol{\beta}$ assumes the persistence of the radiance bias from the previous analysis.

$\mathbf{y}_t$ are the observations at time t, with p_{T} observation times throughout the window. $\mathbf{R}_t$ and

$$\mathbf{B}_{\mathrm{v}} = \begin{pmatrix} \mathbf{B}_{\mathbf{x}} & 0 \\ 0 & \mathbf{B}_{\beta} \end{pmatrix}$$

are the observation and background-error covariances, respectively; that is, it is assumed that the error in the background estimate of the state and $\boldsymbol{\beta}$ are uncorrelated.

To compare the augmented state with the observations, the state variables at the initial time are first propagated to the time of the observations using the dynamical model, $m_{0\to t}$. h_t is then the mapping from the state variables given at time t and $\boldsymbol{\beta}$ to the observation variables at time t. If the k^{th} observation is to be bias-corrected, h_t will include an additive correction term of the form

$$c_k(\mathbf{x}^{\mathrm{b}}, \boldsymbol{\beta}) = s_k + \sum_{i=1}^{r_k}\beta_{k,i}\rho_{k,i}(\mathbf{x}^{\mathrm{b}}), \tag{2}$$

where s_k is a constant, $\rho_{k,i}$ are the r_k predictors, and $\beta_{k,i}$ are the corresponding bias-correction coefficients, which are to be updated at each cycle (Cameron & Bell, 2018). The choice of predictors depends on the sensor and channel to be corrected. These may be as simple as a fixed global correction to remove scan-dependent biases or may involve multiple air-mass-dependent predictors to characterise and remove more complex bias patterns (Cameron & Bell, 2018). The sensitivity of the update of the bias-correction coefficients during each cycle depends on the choice of $\mathbf{B}_{\beta}$. At the Met Office, $\mathbf{B}_{\beta}$ is diagonal and a function of a pre-defined "bias-halving time" parameter (Cameron & Bell, 2018).

An analytical expression for the analysis of the state and $\boldsymbol{\beta}$ that minimises Equation (1) can be found by linearising the observation operator around the best estimate of the augmented state:

$$\mathbf{v}^{\mathrm{a}} = \mathbf{v}^{\mathrm{b}} + \mathbf{K}_{\mathrm{v}}(\hat{\mathbf{y}} - \hat{h}(\mathbf{v}^{\mathrm{b}})), \tag{3}$$

where the hat overbars indicate that the observations and observation operator are four-dimensional, so that we can drop the summation over time notation used in Equation (1) (see Lorenc, 1986).

$\mathbf{K}_{\mathrm{v}}$ is referred to as the Kalman gain matrix and is given as a function of the background-error covariance matrix and the four-dimensional observation-error covariance matrix $\hat{\mathbf{R}}$:

$$\mathbf{K}_{\mathrm{v}} = \mathbf{B}_{\mathrm{v}}\hat{\mathbf{H}}_{\mathrm{v}}^{\mathrm{T}}(\hat{\mathbf{H}}_{\mathrm{v}}\mathbf{B}_{\mathrm{v}}\hat{\mathbf{H}}_{\mathrm{v}}^{\mathrm{T}} + \hat{\mathbf{R}})^{-1}, \tag{4}$$

where $\hat{\mathbf{H}}_{\mathrm{v}}$ is the Jacobian of the four-dimensional observation operator.

The 4D observation operators, $\hat{h}$, include the model used to assimilate the data, m. This may include a significant bias. In this work, we define the error in the model over timestep i as

$$\boldsymbol{\eta}_{i+1} = m_{i\to i+1}(\mathbf{x}_i) - m^{\dagger}_{i\to i+1}(\mathbf{x}_i), \tag{5}$$

where $m^{\dagger}_{i\to i+1}$ is the true discretised model propagating the state at time i forward one model timestep. The model bias is given by the expectation $\langle\boldsymbol{\eta}_i\rangle$.

In Appendix A, the error in the analysis is derived for the augmented state, $\mathbf{v}$. The analysis error for the bias-correction coefficient only, $\boldsymbol{\epsilon}^{\mathrm{a}}_{\beta}$, is then isolated (see Section A.1). To understand how the bias-corrected (BC) observations (i.e., the radiances) and anchor observations contribute to the analysis error, we then separate out the terms associated with each. The resulting expression for the analysis error for the bias-correction coefficient is then

$$\boldsymbol{\epsilon}^{\mathrm{a}}_{\beta} = \boldsymbol{\epsilon}^{\mathrm{b}}_{\beta} + \mathbf{K}_{\beta,\mathrm{BC}}\Big(\widehat{\boldsymbol{\epsilon}}_{\mathrm{BC}} - \widehat{\boldsymbol{\eta}}_{\mathrm{BC}} - \widehat{\mathbf{C}}_{\beta}\boldsymbol{\epsilon}^{\mathrm{b}}_{\beta} - \widehat{\mathbf{H}}_{\mathrm{BC}}\mathbf{D}(\widehat{\boldsymbol{\epsilon}}_{\downarrow} - \widehat{\boldsymbol{\eta}}_{\downarrow}) - \widehat{\mathbf{H}}_{\mathrm{BC}}(\mathbf{I}_n - \mathbf{D}\widehat{\mathbf{H}}_{\downarrow})\boldsymbol{\epsilon}^{\mathrm{b}}_{x}\Big) \tag{6}$$

(Equation A.6 in Section A.1), where

- $\boldsymbol{\epsilon}^{\mathrm{b}}_{\beta} \in \mathbb{R}^{r}$ is the error in the background estimate of the bias-correction coefficient,
- $\widehat{\boldsymbol{\epsilon}}_{\mathrm{BC}} \in \mathbb{R}^{p_{\mathrm{BC}}}$ and $\widehat{\boldsymbol{\epsilon}}_{\downarrow} \in \mathbb{R}^{p_{\downarrow}}$ are the 4D instrument errors for the p_{BC} bias-corrected and $p_{\downarrow}$ anchor observations, respectively (including the effect of the bias correction),
- $\widehat{\boldsymbol{\eta}}_{\mathrm{BC}} \in \mathbb{R}^{p_{\mathrm{BC}}}$ and $\widehat{\boldsymbol{\eta}}_{\downarrow} \in \mathbb{R}^{p_{\downarrow}}$ are the model errors "seen" by the BC and anchor observations, respectively (see Section A.3 for definition), and
- $\boldsymbol{\epsilon}^{\mathrm{b}}_{x} \in \mathbb{R}^{n}$ is the background error in the state.

$\widehat{\mathbf{H}}_{\mathrm{BC}} \in \mathbb{R}^{p_{\mathrm{BC}}\times n}$ and $\widehat{\mathbf{H}}_{\downarrow} \in \mathbb{R}^{p_{\downarrow}\times n}$ are the Jacobians of the 4D observation operators with respect to the state for the BC and anchor observations, respectively. $\widehat{\mathbf{C}}_{\beta} \in \mathbb{R}^{p_{\mathrm{BC}}\times r}$ is the Jacobian of the bias-correction term with respect to the bias-correction coefficients. This has the hat overbar to represent the correction made to the 4D BC observations.

$\mathbf{D}$ in Equation (6) depends on the error covariance of the anchor observations relative to that of the background state in the space of the anchor observations:

$$\mathbf{D} = \mathbf{B}_x\widehat{\mathbf{H}}^{\mathrm{T}}_{\downarrow}(\widehat{\mathbf{H}}_{\downarrow}\mathbf{B}_x\widehat{\mathbf{H}}^{\mathrm{T}}_{\downarrow} + \widehat{\mathbf{R}}_{\downarrow})^{-1}. \tag{7}$$

$\mathbf{K}_{\beta,\mathrm{BC}}$ in Equation (6) is the block of the Kalman gain matrix, Equation (4), describing the sensitivity of the analysis of $\boldsymbol{\beta}$ to the BC observations:

$$\begin{aligned}\mathrm{K}_{\beta,\mathrm{BC}} &= \mathrm{B}_{\beta}\hat{\mathrm{C}}_{\beta}\left[(\hat{\mathrm{H}}_{\mathrm{v}}\mathrm{B}_{\mathrm{v}}\hat{\mathrm{H}}^{\mathrm{T}}_{\mathrm{v}} + \hat{\mathrm{R}})^{-1}\right]_{\mathrm{BC,BC}} \\ &= \mathrm{B}_{\beta}\hat{\mathrm{C}}_{\beta}\left[\begin{pmatrix}\hat{\mathrm{H}}_{\mathrm{BC}}\mathrm{B}_x\hat{\mathrm{H}}^{\mathrm{T}}_{\mathrm{BC}} + \hat{\mathrm{C}}_{\beta}\mathrm{B}_{\beta}\hat{\mathrm{C}}^{\mathrm{T}}_{\beta} + \hat{\mathrm{R}}_{\mathrm{BC}} & \hat{\mathrm{H}}_{\mathrm{BC}}\mathrm{B}_x\hat{\mathrm{H}}^{\mathrm{T}}_{\downarrow} \\ \hat{\mathrm{H}}_{\downarrow}\mathrm{B}_x\hat{\mathrm{H}}^{\mathrm{T}}_{\mathrm{BC}} & \hat{\mathrm{H}}_{\downarrow}\mathrm{B}_x\hat{\mathrm{H}}^{\mathrm{T}}_{\downarrow} + \hat{\mathrm{R}}_{\downarrow}\end{pmatrix}^{-1}\right]_{\mathrm{BC,BC}}.\end{aligned} \tag{8}$$

Note that in Equation (8) we assume that the observation vector, $\hat{\mathbf{y}}$, is partitioned as $(\hat{\mathbf{y}}^{\mathrm{T}}_{\mathrm{BC}}, \hat{\mathbf{y}}^{\mathrm{T}}_{\downarrow})^{\mathrm{T}}$.

From Equation (8) we see how $\mathbf{K}_{\beta,\mathrm{BC}}$ depends on the full observing system and the information the observations provide about the state, as well as $\boldsymbol{\beta}$. This is illustrated in Figure 1 for the case in which $\mathbf{x}$ and $\boldsymbol{\beta}$ are scalars and one BC and one anchor observation are assimilated, with the observation operator for each equal to 1 and the bias correction, $c(\beta)$, equal to $\boldsymbol{\beta}$. For this scalar example, we therefore have

$$\widehat{\mathbf{H}}_{\mathrm{v}} = \begin{pmatrix}\widehat{\mathbf{H}}_{\mathrm{BC}} & \widehat{\mathbf{C}}_{\beta} \\ \widehat{\mathbf{H}}_{\downarrow} & 0\end{pmatrix} = \begin{pmatrix}1 & 1 \\ 1 & 0\end{pmatrix},$$

resulting in

$$\begin{aligned}K_{\beta,\mathrm{BC}} &= \frac{B_{\beta}(B_x + R_{\downarrow})}{(B_x + B_{\beta} + R_{\mathrm{BC}})(B_x + R_{\downarrow}) - B_x^2} \\ &= \frac{B_{\beta}(B_x + R_{\downarrow})}{B_xR_{\downarrow} + (B_x + R_{\downarrow})(B_{\beta} + R_{\mathrm{BC}})}.\end{aligned}$$

In the first panel, we see, as expected, that the sensitivity of the analysis of $\boldsymbol{\beta}$ to the BC observations (quantified by $\mathbf{K}_{\beta,\mathrm{BC}}$) is a balance between the information in the BC observations and the background estimate of $\boldsymbol{\beta}$. That is, $\mathbf{K}_{\beta,\mathrm{BC}}$ increases as the BC observations become more precise (R_{BC} reduces) and the background bias-correction coefficient becomes less precise (B_{β} increases).

In the second two panels, we see that, when R_{BC} is small, $\mathbf{K}_{\beta,\mathrm{BC}}$ also increases as the anchor observation and background state become more precise ($R_{\downarrow}$ and B_x reduce). This is due to their role as a reference state for estimating the bias-correction coefficients. Therefore, the more precise either the anchor observation or background state is, the more information the BC observations can provide about $\boldsymbol{\beta}$. In fact, in this simple scalar experiment, one can see that the last two panels are identical, as the two references are treated identically by VarBC.

This is different from the behaviour that we would expect for $\mathbf{K}_{x,\mathrm{BC}}$ (the sensitivity of the analysis of the state to the BC observations), which increases as the BC observations become more precise (R_{BC} reduces) but reduces as the anchor observation and background state become more precise ($R_{\downarrow}$ and B_x reduce). This is because, unlike the bias-correction coefficients, the state becomes less constrained by the BC observations when there is more information in the background estimate of the state and the anchor observations (as controlled by the relative magnitudes of R_{BC}, $R_{\downarrow}$, and B_x).

The different sensitivity of the analysis of state and bias-correction coefficients to the BC observations can be illuminated further by comparing the equations for their error covariances. In Appendix B.1, the analysis-error covariance of the bias-correction coefficient is derived to be

$$\mathbf{A}_{\beta} = (\mathbf{I}_r - \mathbf{K}_{\beta,\mathrm{BC}}\widehat{\mathbf{C}}_{\beta})\mathbf{B}_{\beta} \tag{9}$$

(Equation B.1 in Section B.1). Note that this relationship assumes that the model error is a bias with no random

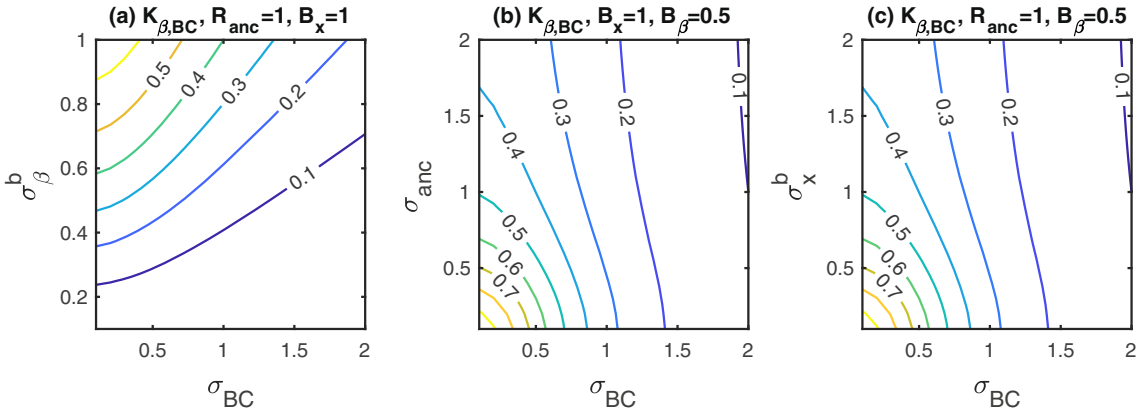

FIGURE 1 Scalar example of how the sensitivity of the analysis of the bias coefficients to bias-corrected observations, $\mathbf{K}_{\beta,\mathrm{BC}}$ (Equation 8), depends on the standard deviation of the errors in the BC observations (x-axis in each panel) and the standard deviation of the errors in the (a) background bias-correction coefficient, (b) anchor observations, and (c) background state. The values of the fixed parameters are given in the title of each plot. Note that panels (b) and (c) are identical, because, in this scalar example, the two references for computing the bias (the anchor observations and the background state) are interchangeable. [Colour figure can be viewed at wileyonlinelibrary.com]

component. A derivation of $\mathbf{A}_\beta$ modified for the case in which the model error has a covariance can also be found in Section B.1. This can be compared with the analysis-error covariances of the state (also derived in Section B.1):

$$\mathbf{A}_x = (\mathbf{I}_n - \mathbf{K}_{x,\mathrm{BC}}\hat{\mathbf{H}}_{\mathrm{BC}})(\mathbf{I}_n - \mathbf{D}\hat{\mathbf{H}}_{\perp})\mathbf{B}_x. \tag{10}$$

In Equation (9), we see that the greater the sensitivity of the analysis of $\boldsymbol{\beta}$ to the BC observations (i.e., the larger the magnitude of the terms in $\mathbf{K}_{\beta,\mathrm{BC}}$), the greater the reduction in the trace of $\mathbf{A}_\beta$ compared with the trace of $\mathbf{B}_\beta$. However, this relationship between $\mathbf{A}_\beta$ and $\mathbf{K}_{\beta,\mathrm{BC}}$ in Equation (9) is not mirrored in the relationship between $\mathbf{K}_{x,\mathrm{BC}}$ and $\mathbf{A}_x$ in Equation (10).

Studying the terms of $\mathbf{K}$ illustrates what influences the weighting of the different sources of information when analysing the state and $\boldsymbol{\beta}$. However, in the presence of unaccounted-for model bias, the assumptions on which the data assimilation algorithms are based break down. In this case, the information will not necessarily be used in the correct way. In the following, we are interested in how the anchor observations can mitigate the analysed bias-correction coefficients, $\boldsymbol{\beta}^{\mathrm{a}}$, from becoming contaminated with the bias in the background state and model ($\langle\boldsymbol{\epsilon}_x^{\mathrm{b}}\rangle$ and $\langle\boldsymbol{\eta}\rangle$ respectively). We therefore derive the expected error in the analysed bias-correction coefficient by assuming that VarBC was applied correctly in the previous cycle, so that the effect of the model and background bias can be isolated in the current assimilation cycle (in Section 4, we will relax this assumption in cycled numerical experiments). The assumption that VarBC was applied correctly in the previous cycle translates to the following assumptions:

- there is no bias, only random error, in the background estimate of the bias-correction coefficient ($\langle\boldsymbol{\epsilon}_\beta^{\mathrm{b}}\rangle = 0$);
- there is no bias, only random error, in the bias-corrected instrument error ($\langle\hat{\boldsymbol{\epsilon}}_{\mathrm{BC}}\rangle = 0$), i.e. the bias correction is performed accurately;
- by definition, there is no bias, only random error, in the instrument error of the anchor observations ($\langle\hat{\boldsymbol{\epsilon}}_{\perp}\rangle = 0$).

This allows us then to derive from Equation (6) the following expression for the expected error in the analysed bias-correction coefficient:

$$\langle\boldsymbol{\epsilon}_\beta^{\mathrm{a}}\rangle = -\mathbf{K}_{\beta,\mathrm{BC}}\Big(\langle\hat{\boldsymbol{\eta}}_{\mathrm{BC}}\rangle - \hat{\mathbf{H}}_{\mathrm{BC}}\mathbf{D}\langle\hat{\boldsymbol{\eta}}_{\perp}\rangle + \hat{\mathbf{H}}_{\mathrm{BC}}(\mathbf{I}_n - \mathbf{D}\hat{\mathbf{H}}_{\perp})\langle\boldsymbol{\epsilon}_x^{\mathrm{b}}\rangle\Big). \tag{11}$$

Equation (11) can be compared with eq. (24) of F23, in which only the background bias was considered in a 3DVar implementation; note the slight difference in the definition of $\mathbf{D}$.

In F23 it was shown how the position of the anchor observations relative to the BC observations is important for their ability to reduce the contamination of model bias in VarBC. When the anchor and BC observations observe different variables and locations, the information that the anchor observation provides about the model bias in VarBC depends on $\mathbf{H}_{\mathrm{BC}}\mathbf{B}_x\mathbf{H}_{\perp}^{\mathrm{T}}$ (i.e., the strength of the background-error correlations between the BC and anchor observed variables) and the overlap in the background bias observed by the BC and anchor observations ($\mathbf{H}_{\mathrm{BC}}\langle\boldsymbol{\epsilon}_x^{\mathrm{b}}\rangle$ and $\mathbf{H}_{\perp}\langle\boldsymbol{\epsilon}_x^{\mathrm{b}}\rangle$, respectively). When $\mathbf{H}_{\mathrm{BC}}\mathbf{B}_x\mathbf{H}_{\perp}^{\mathrm{T}}$ contains strong correlations, the anchor observations will have a large impact on VarBC. In this case, if the anchor

and BC observations observe similar background bias then the anchor observations can reduce the effect of the background bias in VarBC. Alternatively, if the anchor and BC observations observe different background biases, then the anchor observations can act as a medium for the background bias seen only by the anchor observations to contaminate VarBC.

Extending these results to a 4DVar implementation means that we now need to include the effect of model bias within the assimilation window. We also need to consider the timing of the anchor observation, with the assimilation model becoming another medium for the information in the observations to be spread spatially from one timestep to another.

The role of the anchor observations in mitigating the analysed observation bias-correction coefficients, $\boldsymbol{\beta}^{\mathrm{a}}$, from becoming contaminated with the bias in the background state and model can be revealed by looking at the limits of Equation (11) as the anchor observations become more precise (their error variance reduces). The more precise the anchor observations, the greater the weight they will have in the assimilation system. In an optimal system, better (i.e., more accurate and precise) data will provide a better analysis. However, this is not necessarily true when we have unaccounted-for model bias. If the system is not optimal, then attempting to fit strongly to precise observations will highlight the sub-optimalities.

The theory presented up to this point is completely general in terms of the meaning of the observation operators, the sampling of the observations, and the distribution of the model bias. In the following subsections and in the next section, we will analyse Equation (11) for a series of different simplified/idealised observation configurations to highlight cases when the anchor observations reduce or increase the damaging effects of model bias. The observations can continue to be interpreted in a general way, but we do not take explicitly into account the complex operators of the radiances and anchor observations, such as GNSS-RO.

The theory presented assumes that the anchor observations (i.e., those not bias-corrected) have negligible bias. The validity of this assumption is questionable. When biases are present in the "anchor" observations, an extra bias term will appear in Equation (11). We will discuss the implications of this further in Section 6.

2.1 | Anchor and bias-corrected observations completely overlap

Before focusing in on the timing of the anchor observations, let us first consider the extreme case when the anchor and BC observations completely overlap. That is, they are not only at the same times but also have the same variables and are at the same location. In this case, the observation operators for the anchor and BC observations will be the same, which we will denote by $\widehat{\mathbf{H}}$, that is, $\widehat{\mathbf{H}}_{\mathrm{BC}} = \widehat{\mathbf{H}}_{\perp} = \widehat{\mathbf{H}}$. The model bias "seen" by the anchor and BC observations will also be the same, which we will denote by $\widehat{\boldsymbol{\eta}}$, that is, $\widehat{\boldsymbol{\eta}}_{\mathrm{BC}} = \widehat{\boldsymbol{\eta}}_{\perp} = \widehat{\boldsymbol{\eta}}$.

In this case, Equation (11) then simplifies to

$$\langle \boldsymbol{\epsilon}^{\mathrm{a}}_{\beta} \rangle = -\mathbf{K}_{\beta,\mathrm{BC}} \Big((\mathbf{I}_p - \widehat{\mathbf{H}}\mathbf{D}) \langle \widehat{\boldsymbol{\eta}} \rangle + (\mathbf{I}_p - \widehat{\mathbf{H}}\mathbf{D}) \widehat{\mathbf{H}} \langle \boldsymbol{\epsilon}^{\mathrm{b}}_{x} \rangle \Big), \quad (12)$$

where p is the number of anchor (or equivalently BC) observations.

From the definition of $\mathbf{D}$ (Equation 7), we see that, in this case, as the anchor observations become more precise ($\widehat{\mathbf{R}}_{\perp} \to \mathbf{0}$), $\widehat{\mathbf{H}}\mathbf{D} \to \mathbf{I}_p$ (assuming that $\widehat{\mathbf{H}}\mathbf{B}_x\widehat{\mathbf{H}}^{\mathrm{T}}$ is invertible). Therefore, in this special case, infinitely precise anchor observations can eliminate the contamination of background and model bias in the analysis of $\boldsymbol{\beta}$. However, in this same case the bias in the analysis of the state (see Section A.2, Equation A.10) will not be eliminated by infinitely precise anchor observations.

2.2 | Anchor and bias-corrected observations have no overlap

We can also consider the other extreme case, where the anchor and BC observation share no information. That is, $\widehat{\mathbf{H}}_{\mathrm{BC}}\mathbf{B}_x\widehat{\mathbf{H}}^{\mathrm{T}}_{\perp} = \mathbf{0}$, in other words there is no correlation between the background errors transformed to the space of the BC observations and the background errors transformed to the space of the anchor observations. In this case, $\widehat{\mathbf{H}}_{\mathrm{BC}}\mathbf{D} = \mathbf{0}$ and so Equation (11) simplifies to

$$\langle \boldsymbol{\epsilon}^{\mathrm{a}}_{\beta} \rangle = -\mathbf{K}_{\beta,\mathrm{BC}} \Big(\langle \widehat{\boldsymbol{\eta}}_{\mathrm{BC}} \rangle + \widehat{\mathbf{H}}_{\mathrm{BC}} \langle \boldsymbol{\epsilon}^{\mathrm{b}}_{x} \rangle \Big). \quad (13)$$

In this case, the anchor observations play no role in computing the analysis of $\boldsymbol{\beta}$. Therefore, from Equation (13) we see that there is no contribution from the model bias "seen" by the anchor observations to the expected value of the error in the analysis of the bias-correction coefficients. We also see that, compared with the case when we had complete overlap, there is now no route for the anchor observations to control the contamination of the background bias and the model bias "seen" by the BC observations in the expected value of the error in the analysis of the bias-correction coefficients. As $\mathbf{K}_{\beta,\mathrm{BC}}$ (Equation 8) has no dependence on the anchor observations, the anchor observations will also not reduce the error variance of the analysis of the bias-correction coefficients.

3 | THE IMPORTANCE OF THE TIMING OF ANCHOR OBSERVATIONS

To understand the importance of the timing of the anchor observations, in the following we will study Equation (11) for three different simplified observation configurations: (i) anchor and BC observations at the same times (Section 3.1), (ii) anchor observations after the BC observations (Section 3.2), and (iii) anchor observations before the BC observations (Section 3.3). These examples fall in between the two extreme cases presented in Sections 2.1 and 2.2.

3.1 | Case I: Anchor and bias-corrected observations at the same time

Let us assume that both the BC and anchor observations are at the same time, j, but do not necessarily observe the same state variables. In this case, the 4D observation operator Jacobians become $\widehat{\mathbf{H}}_{\mathrm{BC}} = \mathbf{H}_{\mathrm{BC}}\mathbf{M}_{0\to j}$ and $\widehat{\mathbf{H}}_{\downarrow} = \mathbf{H}_{\downarrow}\mathbf{M}_{0\to j}$, where $\mathbf{M}_{0\to j}$ is the Jacobian of the dynamical model propagating the state from time 0 to time j and $\mathbf{H}_{\mathrm{BC}}$ and $\mathbf{H}_{\downarrow}$ are the Jacobians of the observation operators mapping from the state at time j to the BC observations and anchor observations at the same time, respectively. The model bias "seen" by the BC and anchor observations becomes $\langle\widehat{\boldsymbol{\eta}}_{\mathrm{BC}}\rangle = \mathbf{H}_{\mathrm{BC}}\sum_{k=1}^{j}\mathbf{M}_{k\to j}\langle\boldsymbol{\eta}_k\rangle$and $\langle\widehat{\boldsymbol{\eta}}_{\downarrow}\rangle = \mathbf{H}_{\downarrow}\sum_{k=1}^{j}\mathbf{M}_{k\to j}\langle\boldsymbol{\eta}_k\rangle$.

Substituting these expressions into Equation (11) and letting the accumulated model bias up to time j be represented by $\langle\boldsymbol{\xi}_{0\to j}\rangle = \sum_{k=1}^{j}\mathbf{M}_{k\to j}\langle\boldsymbol{\eta}_k\rangle$, we obtain

$$\begin{aligned}\langle\boldsymbol{\epsilon}^{\mathrm{a}}_{\beta}\rangle = -\mathbf{K}_{\beta,\mathrm{BC}}\mathbf{H}_{\mathrm{BC}}\big((\mathbf{I}_n - \mathbf{M}_{0\to j}\mathbf{D}\mathbf{H}_{\downarrow})\langle\boldsymbol{\xi}_{0\to j}\rangle \\ +(\mathbf{I}_n - \mathbf{M}_{0\to j}\mathbf{D}\mathbf{H}_{\downarrow})\mathbf{M}_{0\to j}\langle\boldsymbol{\epsilon}^{\mathrm{b}}_{x}\rangle\big).\end{aligned} \tag{14}$$

As illustrated in Sections 2.1 and 2.2, the ability of precise anchor observations to reduce the effect of the model and background bias depends on the overlap in the observations described by $\widehat{\mathbf{H}}_{\mathrm{BC}}\mathbf{B}_x\widehat{\mathbf{H}}_{\downarrow}^{\mathrm{T}}$, which in this case is $\mathbf{H}_{\mathrm{BC}}\mathbf{M}_{0\to j}\mathbf{B}_x\mathbf{M}_{0\to j}^{\mathrm{T}}\mathbf{H}_{\downarrow}^{\mathrm{T}}$. This is analogous to the 3D case studied in F23, in which the importance of correlations in $\mathbf{B}_x$ was discussed. The difference compared with the 3D case is that now the model can also advect information and so it is the structure of the evolved error covariance matrix, $\mathbf{M}_{0\to j}\mathbf{B}_x\mathbf{M}_{0\to j}^{\mathrm{T}}$, that is important for spreading information about the model and background biases provided by the anchor observations.

3.2 | Case II: Anchor observations after bias-corrected observations

In the second example, we consider the case in which the anchor observations are made at one time, j, that is later than the BC observations made at time i ($i < j$). In this case, the 4D observation operator Jacobians become $\widehat{\mathbf{H}}_{\mathrm{BC}} = \mathbf{H}_{\mathrm{BC}}\mathbf{M}_{0\to i}$ and $\widehat{\mathbf{H}}_{\downarrow} = \mathbf{H}_{\downarrow}\mathbf{M}_{0\to j}$. The model bias "seen" by the BC observations becomes

$$\langle\widehat{\boldsymbol{\eta}}_{\mathrm{BC}}\rangle = \mathbf{H}_{\mathrm{BC}}\sum_{k=1}^{i}\mathbf{M}_{k\to i}\langle\boldsymbol{\eta}_k\rangle = \mathbf{H}_{\mathrm{BC}}\langle\boldsymbol{\xi}_{0\to i}\rangle. \tag{15}$$

Because the anchor observations are at a later time, they will see the model bias accumulated up to the timestep of the BC observations (propagated up to the time of the anchor observations), but also additional model bias not seen by the BC observations:

$$\begin{aligned}\langle\widehat{\boldsymbol{\eta}}_{\downarrow}\rangle &= \mathbf{H}_{\downarrow}\sum_{k=1}^{j}\mathbf{M}_{k\to j}\langle\boldsymbol{\eta}_k\rangle \\ &= \mathbf{H}_{\downarrow}\left(\mathbf{M}_{i+1\to j}\langle\boldsymbol{\xi}_{0\to i}\rangle + \sum_{k=i+1}^{j}\mathbf{M}_{k\to j}\langle\boldsymbol{\eta}_k\rangle\right) \\ &= \mathbf{H}_{\downarrow}\big(\mathbf{M}_{i+1\to j}\langle\boldsymbol{\xi}_{0\to i}\rangle + \langle\boldsymbol{\xi}_{i+1\to j}\rangle\big).\end{aligned} \tag{16}$$

Substituting these expressions into Equation (11), we obtain

$$\begin{aligned}\langle\boldsymbol{\epsilon}^{\mathrm{a}}_{\beta}\rangle &= -\mathbf{K}_{\beta,\mathrm{BC}}\big(\mathbf{H}_{\mathrm{BC}}\langle\boldsymbol{\xi}_{0\to i}\rangle - \mathbf{H}_{\mathrm{BC}}\mathbf{M}_{0\to i}\mathbf{D}\mathbf{H}_{\downarrow}(\mathbf{M}_{i+1\to j}\langle\boldsymbol{\xi}_{0\to i}\rangle \\ &\quad +\langle\boldsymbol{\xi}_{i+1\to j}\rangle)) + \mathbf{H}_{\mathrm{BC}}\mathbf{M}_{0\to i}(\mathbf{I}_n - \mathbf{D}\mathbf{H}_{\downarrow}\mathbf{M}_{0\to j})\langle\boldsymbol{\epsilon}^{\mathrm{b}}_{x}\rangle\big) \\ &= -\mathbf{K}_{\beta,\mathrm{BC}}\mathbf{H}_{\mathrm{BC}}\big(\big(\mathbf{I}_n - \mathbf{M}_{0\to i}\mathbf{D}\mathbf{H}_{\downarrow}\mathbf{M}_{i+1\to j}\big) \\ &\quad \times \big(\langle\boldsymbol{\xi}_{0\to i}\rangle + \mathbf{M}_{0\to i}\langle\boldsymbol{\epsilon}^{\mathrm{b}}_{x}\rangle\big) - \mathbf{M}_{0\to i}\mathbf{D}\mathbf{H}_{\downarrow}\langle\boldsymbol{\xi}_{i+1\to j}\rangle\big).\end{aligned} \tag{17}$$

In this case, the overlap in the observations is described by $\mathbf{H}_{\mathrm{BC}}\mathbf{M}_{0\to i}\mathbf{B}_x\mathbf{M}_{0\to j}^{\mathrm{T}}\mathbf{H}_{\downarrow}^{\mathrm{T}}$. Therefore, the temporal error correlations in the propagated background errors between times i and j are now important. This overlap appears at the beginning of the $\mathbf{H}_{\mathrm{BC}}\mathbf{M}_{0\to i}\mathbf{D}$ term (see definition of $\mathbf{D}$, Equation 7) and is therefore important for reducing the effect of $\langle\boldsymbol{\xi}_{0\to i}\rangle$ (the model bias seen by both the anchor and BC obs) and the effect of $\langle\boldsymbol{\epsilon}^{\mathrm{b}}_{x}\rangle$. However, unlike the previous case, the fact that the anchor observations will see more accumulated model bias than the BC observations is also important. This can be shown by assuming that the anchor observations measure the same variables as the BC observations, such that $\mathbf{H}_{\mathrm{BC}} = \mathbf{H}_{\downarrow} = \mathbf{H}$, and the model propagating the biases from the time of the BC observations to the time of the anchor observations is the

identity ($\mathbf{M}_{i+1\to j}=\mathbf{I}_n$). In this case, we have perfect overlap in $\widehat{\mathbf{H}}_{\mathrm{BC}}\mathbf{B}_x^{1/2}$ and $\widehat{\mathbf{H}}_{\mathbb{1}}\mathbf{B}_x^{1/2}$, as $\widehat{\mathbf{H}}_{\mathbb{1}}=\widehat{\mathbf{H}}_{\mathrm{BC}}$. However, the additional model bias seen by the anchor observations is still $\langle\boldsymbol{\xi}_{i+1\to j}\rangle$ and

$$\begin{aligned}\langle\boldsymbol{\epsilon}_{\beta}^{\mathrm{a}}\rangle=-\mathbf{K}_{\beta,\mathrm{BC}}\big((\mathbf{I}_p-\mathbf{H}\mathbf{M}_{0\to i}\mathbf{D})\mathbf{H}(\langle\boldsymbol{\xi}_{0\to i}\rangle+\mathbf{M}_{0\to i}\langle\boldsymbol{\epsilon}_x^{\mathrm{b}}\rangle)\\-\mathbf{H}\mathbf{M}_{0\to i}\mathbf{D}\mathbf{H}\langle\boldsymbol{\xi}_{i+1\to j}\rangle\big).\end{aligned}\tag{18}$$

In this special case, the terms multiplying $\langle\boldsymbol{\xi}_{0\to i}\rangle$ and $\langle\boldsymbol{\epsilon}_x^{\mathrm{b}}\rangle$ will tend to $\mathbf{0}$ as the anchor observations become more precise ($\widehat{\mathbf{R}}_{\mathbb{1}}\to\mathbf{0}$ and $\mathbf{H}\mathbf{M}_{0\to i}\mathbf{D}\to\mathbf{I}_p$). However, the additional model error seen only by the anchor observations $\langle\boldsymbol{\xi}_{i+1\to j}\rangle$ will not be eliminated and $\mathbf{H}\mathbf{M}_{0\to i}\mathbf{D}\mathbf{H}\langle\boldsymbol{\xi}_{i+1\to j}\rangle$ will tend to an upper limit of $\mathbf{H}\langle\boldsymbol{\xi}_{i+1\to j}\rangle$ as $\widehat{\mathbf{R}}_{\mathbb{1}}\to\mathbf{0}$.

3.3 | Case III: Anchor observations before bias-corrected observations

In the third and final example, we consider the case in which the anchor observations are made at one time i, which is earlier than the BC observations made at one time j ($j>i$). In this case, the 4D observation operator Jacobians become $\widehat{\mathbf{H}}_{\mathrm{BC}}=\mathbf{H}_{\mathrm{BC}}\mathbf{M}_{0\to j}$ and $\widehat{\mathbf{H}}_{\mathbb{1}}=\mathbf{H}_{\mathbb{1}}\mathbf{M}_{0\to i}$. The model bias "seen" by the BC and anchor observations becomes $\langle\widehat{\boldsymbol{\eta}}_{\mathrm{BC}}\rangle=\mathbf{H}_{\mathrm{BC}}\sum_{k=1}^{j}\mathbf{M}_{k\to j}\langle\boldsymbol{\eta}_k\rangle=\mathbf{H}_{\mathrm{BC}}(\mathbf{M}_{i+1\to j}\langle\boldsymbol{\xi}_{0\to i}\rangle+\langle\boldsymbol{\xi}_{i+1\to j}\rangle)$ and $\langle\widehat{\boldsymbol{\eta}}_{\mathbb{1}}\rangle=\mathbf{H}_{\mathbb{1}}\sum_{k=1}^{i}\mathbf{M}_{k\to i}\langle\boldsymbol{\eta}_k\rangle=\mathbf{H}_{\mathbb{1}}\langle\boldsymbol{\xi}_{0\to i}\rangle$.

Substituting these expressions into Equation (11), we obtain

$$\begin{aligned}\langle\boldsymbol{\epsilon}_{\beta}^{\mathrm{a}}\rangle&=-\mathbf{K}_{\beta,\mathrm{BC}}\big(\mathbf{H}_{\mathrm{BC}}(\mathbf{M}_{i+1\to j}\langle\boldsymbol{\xi}_{0\to i}\rangle+\langle\boldsymbol{\xi}_{i+1\to j}\rangle)\\&\quad-\mathbf{H}_{\mathrm{BC}}\mathbf{M}_{0\to j}\mathbf{D}\mathbf{H}_{\mathbb{1}}\langle\boldsymbol{\xi}_{0\to i}\rangle\\&\quad+\mathbf{H}_{\mathrm{BC}}\mathbf{M}_{0\to j}(\mathbf{I}_n-\mathbf{D}\mathbf{H}_{\mathbb{1}}\mathbf{M}_{0\to i})\langle\boldsymbol{\epsilon}_x^{\mathrm{b}}\rangle\big),\\&=-\mathbf{K}_{\beta,\mathrm{BC}}\mathbf{H}_{\mathrm{BC}}\big((\mathbf{M}_{i+1\to j}-\mathbf{M}_{0\to j}\mathbf{D}\mathbf{H}_{\mathbb{1}})\langle\boldsymbol{\xi}_{0\to i}\rangle\\&\quad+\langle\boldsymbol{\xi}_{i+1\to j}\rangle+\mathbf{M}_{0\to j}(\mathbf{I}_n-\mathbf{D}\mathbf{H}_{\mathbb{1}}\mathbf{M}_{0\to i})\langle\boldsymbol{\epsilon}_x^{\mathrm{b}}\rangle\big)\\&=-\mathbf{K}_{\beta,\mathrm{BC}}\mathbf{H}_{\mathrm{BC}}\big((\mathbf{M}_{i+1\to j}-\mathbf{M}_{i+1\to j}\mathbf{M}_{0\to i}\mathbf{D}\mathbf{H}_{\mathbb{1}})\langle\boldsymbol{\xi}_{0\to i}\rangle\\&\quad+\langle\boldsymbol{\xi}_{i+1\to j}\rangle+\mathbf{M}_{i+1\to j}\mathbf{M}_{0\to i}(\mathbf{I}_n-\mathbf{D}\mathbf{H}_{\mathbb{1}}\mathbf{M}_{0\to i})\langle\boldsymbol{\epsilon}_x^{\mathrm{b}}\rangle\big)\\&=-\mathbf{K}_{\beta,\mathrm{BC}}\mathbf{H}_{\mathrm{BC}}\big(\mathbf{M}_{i+1\to j}(\mathbf{I}_n-\mathbf{M}_{0\to i}\mathbf{D}\mathbf{H}_{\mathbb{1}})\big(\langle\boldsymbol{\xi}_{0\to i}\rangle\\&\quad+\mathbf{M}_{0\to i}\langle\boldsymbol{\epsilon}_x^{\mathrm{b}}\rangle\big)+\langle\boldsymbol{\xi}_{i+1\to j}\rangle\big).\end{aligned}\tag{19}$$

In this case, the overlap in the observations is described by $\mathbf{H}_{\mathrm{BC}}\mathbf{M}_{0\to j}\mathbf{B}_x\mathbf{M}_{0\to i}^{\mathrm{T}}\mathbf{H}_{\mathbb{1}}^{\mathrm{T}}$ and so the temporal error correlations in the propagated background errors between time i and j are still important. However, in this case the additional model error only seen by the BC observations $\langle\boldsymbol{\xi}_{i+1\to j}\rangle$ can only be controlled by the precision of the anchor observations through $\mathbf{K}_{\beta,\mathrm{BC}}$. In Section 2 we saw that $\mathbf{K}_{\beta,\mathrm{BC}}$ increases as the anchor observations become more precise. Therefore, more precise anchor observations will exacerbate the contamination of VarBC by the model bias seen only by the BC observations.

3.4 | Scalar illustration

To illustrate the contributions of the different bias sources to the total expected error in the analysis of $\boldsymbol{\beta}$ (Equation 11) for the different observation configuration cases, we set up a simple scalar experiment. In each case, we assume we have an assimilation window of 10 timesteps and one each of BC and anchor observations is made during the window observing the state variable directly.

- Case I: both the anchor and BC observation are at the end of the window.
- Case II: the anchor observation is at the end of the window and the BC observation is halfway through the window.
- Case III: the anchor observation is halfway through the window and the BC observation is at the end of the window.

As in the reference values for Figure 1, $B_x=R_{\mathrm{BC}}=1$, $B_\beta=0.5$, and $R_{\mathbb{1}}$ is allowed to vary. Unlike in Figure 1, we now consider an assimilation window. The scalar model is given by $m=1.025$, chosen so that, in the special case when $R_{\mathbb{1}}=1$ and the observations are available at the beginning of the window, the assimilation system is static (the contraction of the uncertainty by the assimilation is balanced by the growth in uncertainty due to the model). The model bias at each timestep is set to a constant such that $\langle\eta\rangle=0.01$. The background bias is set to be the accumulation of the evolved model bias over the previous assimilation window; this gives a value of $\langle\epsilon_x^{\mathrm{b}}\rangle=\sum_{i=0}^{9}m^i\langle\eta\rangle=0.112$, where m^i is the ith power of m. Note that these exact values are not too important, as we only wish to understand the qualitative difference between the different terms in Equation (11). However, the relative magnitudes of the two biases ($\langle\epsilon_x^{\mathrm{b}}\rangle/\langle\eta\rangle\approx 11$) and the fact that they are the same sign are based on reasonable assumptions and will have implications for the interpretation of these results. The true value of β does not appear in Equation (11) and does therefore not need to be defined here.

The total value of $\langle\epsilon_\beta^{\mathrm{a}}\rangle$ is plotted as a function of $R_{\mathbb{1}}$ for the three observation configuration cases with black solid lines in Figure 2. The contribution due to the background bias ($-\mathbf{K}_{\beta,\mathrm{BC}}\widehat{\mathbf{H}}_{\mathrm{BC}}(\mathbf{I}_n-\mathbf{D}\widehat{\mathbf{H}}_{\mathbb{1}})\langle\boldsymbol{\epsilon}_x^{\mathrm{b}}\rangle$ from Equation 11) is plotted with red triangles, the contribution due to the model bias seen by both the BC and anchor observations is plotted with yellow circles, and the contribution due to model

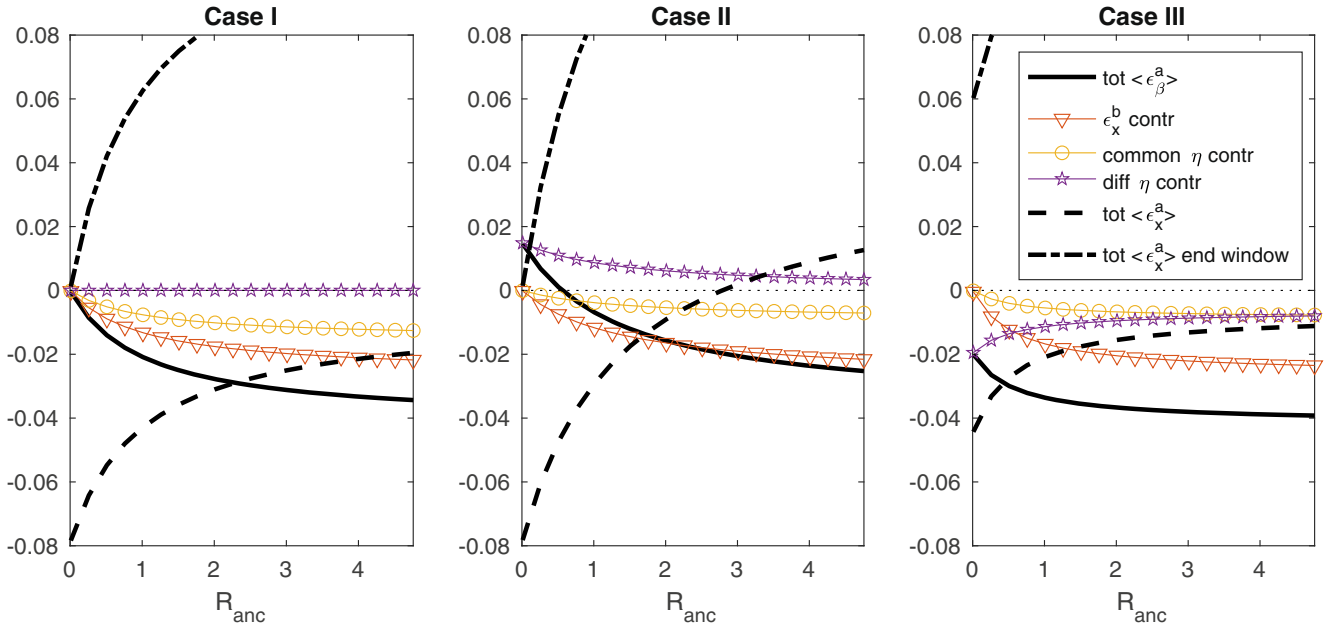

FIGURE 2 Scalar illustration of Equation (11), the contribution of the different bias terms to the total expected error in the analysis of $\boldsymbol{\beta}$ (black solid line) as a function of the anchor observation-error variance (R_{anc}): background bias term (red triangles), model bias seen by both anchor and BC observations (yellow circles), and model bias seen by only the anchor observations or BC observations in Cases II and III respectively (purple stars). For comparison, the expected error in the analysis of $\mathbf{x}$ at the beginning of the window (Equation A.10: black dashed line) and at the end of the window (Equation A.11: black dash–dotted line) as a function of R_{anc} is also plotted. Each panel gives results for the different observing configurations described in Section 3.4. [Colour figure can be viewed at wileyonlinelibrary.com]

bias seen by only one of the observations is plotted with purple stars.

For the three cases, we see that, as the precision of the anchor observations increases ($R_{\downarrow} \to 0$), the contribution of the background bias and the model bias seen by both anchor and BC observations tends to zero. However, the magnitude of the contribution of the model bias seen by only one observation type increases as the precision of the anchor observations increases. Therefore it is only in Case I that infinitely precise anchor observations can provide an unbiased analysis of the bias-correction coefficients; this is consistent with the theory presented in Sections 3.1–3.3.

In Case II, the contribution of the model bias seen by only the anchor observation has a different sign from the other terms (see the last term in Equation 17) and so the total value of $\langle \boldsymbol{\epsilon}^{\mathrm{a}}_{\beta} \rangle$ for the range of $R_{\downarrow}$ plotted is smaller than the value of $\langle \boldsymbol{\epsilon}^{\mathrm{a}}_{\beta} \rangle$ in Case I. $\langle \boldsymbol{\epsilon}^{\mathrm{a}}_{\beta} \rangle$ can also be zero despite the different terms being non-zero, as the different sources of bias compensate one another.

In Case III, the contribution of the model bias seen by only the BC observation has the same sign as the other terms (see Equation 19). Therefore, the total value of $\langle \boldsymbol{\epsilon}^{\mathrm{a}}_{\beta} \rangle$ is never zero and always greater than the value of $\langle \boldsymbol{\epsilon}^{\mathrm{a}}_{\beta} \rangle$ in Case I.

Also plotted in Figure 2 is the total bias in the analysis of the state derived in Appendix A.2 (Equation A.10). In each case, we see that the magnitude of $\langle \boldsymbol{\epsilon}^{\mathrm{a}}_{x} \rangle$ increases as the anchor observations become more precise, with the largest analysis bias when precise anchor observations are at the end of the window (Cases I and II). This is because, to fit these observations, the assimilation system compensates for the positive model bias by introducing a negative bias at the initial time (the time of the state analysis: Howes *et al.*, 2017). This can be seen from the bias in the analysis at the end of the window (Equation A.11, dash–dotted line), which is zero when we have infinitely precise anchor observations at the end of the window. Note that the bias at the end of the window is only plotted for the given ranges. The bias in the analysis at the end of the window will be the background bias for the next window. It is interesting to note that, for Case II, when the value of $R_{\downarrow}$ is large we get a positive bias in the analysis of the state, due to the background bias becoming the dominant source.

In the left-hand panel of Figure 3, the error variance of the analysis bias correction coefficient (Equation 9) is plotted as a function of the precision of the anchor observations for the three scalar cases, assuming that the model error is not random. In Appendix B.2 it is shown that, as $\widehat{R}_{\downarrow} \to 0$, $A_{\beta} \to B_{\beta}\widehat{R}_{\mathrm{BC}}/(B_{\beta} + \widehat{R}_{\mathrm{BC}})$. For our parameter choices, this gives a lower limit for A_{β} of 1/3 for all cases. As $\widehat{R}_{\downarrow}$ increases, we see that, as shown in Equation B.2, the analysis of β is most precise when the BC observations

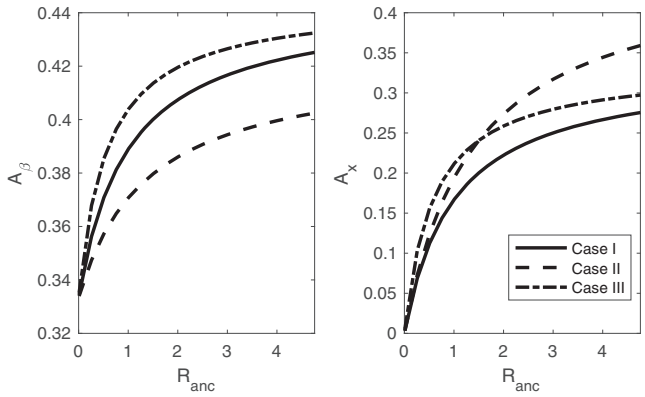

FIGURE 3 The analysis-error variance for (left) bias coefficients $\boldsymbol{\beta}$ and (right) the state as a function of the anchor observation-error variance for the three scalar observing configurations described in Section 3.4.

are earlier than the anchor observations (Case II). This is explained partly by the fact that the background evolved to the earlier BC observations will be able to provide a more precise reference from which the bias-correction coefficients can be estimated (because the model will increase the uncertainty the more the background is evolved). Comparing Cases I and III, when the BC observations are both at the end of the window, we see that we can have a more precise analysis of $\boldsymbol{\beta}$ when the anchor observations are at the same time as the BC observations (Case I). Again, this is because, in this case, the anchor observations provide a more precise reference, as there is no need for model propagation to project the information from the anchor observations forward to the time of the BC observations.

For comparison, in the right-hand panel of Figure 3 the error variance of the analysis state (Equation 10) is plotted as a function of the precision of the anchor observations for the three scalar cases, again assuming that the model error is not random. The analysis of the state is more precise if both observations are at the end of the window (Case I). When $R_{\mathbb{L}} < R_{\mathrm{BC}} = 1$, A_x is more sensitive to the precision of the anchor observations when they are in the middle of the window (Case III) and, when $R_{\mathbb{L}} > R_{\mathrm{BC}} = 1$, A_x is more sensitive to the precision of the anchor observations when they are at the end of the window (Case II). This is consistent with previous studies (e.g., Gauthier *et al.*, 2007; McNally, 2019). In each case as $R_{\mathbb{L}} \to 0$, $A_x \to 0$. This can be seen from Equation (10), as $R_{\mathbb{L}} \to 0 \Rightarrow K_{x,\mathrm{BC}} \to 0$, as the greater weight given to the anchor observations means that the analysis of the state is less sensitive to the BC observations. Additionally, $R_{\mathbb{L}} \to 0 \Rightarrow D \to 1$ from Equation (7).

Figures 2 and 3 highlight the differing observational needs of the analysis of the state and bias-correction coefficients:

- to reduce the contamination of the model bias in the analysis of the bias coefficients, we require precise anchor observations at the end of the window;
- precise observations at the end of the assimilation window increase the contamination of the model bias in the analysis of the state.

In practice, analyses of the state and the bias coefficients are not isolated from one another and the accuracy of each will impact the accuracy of the other. The interaction of the analysis of the state and the bias-correction coefficients is explored in the next section, in which we perform cycled DA experiments.

4 | NUMERICAL EXPERIMENTS USING THE LORENZ 96 MODEL

In the theory developed so far, we have studied only a single cycle, assuming that, in the previous cycle, VarBC was applied optimally for the chosen predictors. However, an inaccurate analysis of the bias-correction coefficient caused by the model bias will then result in a bias in the background estimate of $\boldsymbol{\beta}$ used in the next assimilation cycle. This in turn will mean that the bias in BC observations will not be completely removed and additional terms will appear in Equation (11).

We have also illustrated in the scalar example that the observing system providing the most accurate (smallest bias) and precise (smallest error variance) analysis of $\boldsymbol{\beta}$ will not necessarily provide the most accurate and precise analysis of $\mathbf{x}$. However, in a cycled system these two terms will interact and so both will need to perform well simultaneously.

Within this section, we perform cycled 4DVar assimilation experiments following a similar methodology to F23 using the Lorenz 96 model (Lorenz, 1995). The Lorenz 96 model is a system of coupled ordinary differential equations, which describe the evolution of a quantity via advection, dissipation, and external forcing. We use the standard setup, where a periodic circular domain is discretized into 40 grid points ($\mathbf{x} \in \mathbb{R}^{40}$). The evolution of the variable x^k at the kth grid point is described by

$$\frac{dx^k}{dt} = -x^{k-2}x^{k-1} + x^{k-1}x^{k+1} - x^k + F, \qquad (20)$$

where F, the forcing parameter, is independent of k (Lorenz, 1995). This system is solved numerically using the fourth-order Runge–Kutta scheme with timestep equal to 0.0125.

The true model trajectory is given by propagating the true initial conditions using a forcing of $F = 8$. In these

experiments, the true initial conditions are made to be consistent with the attractor of the Lorenz 96 model by spinning up from a sine wave for 1000 timesteps. Model bias is then generated by using a forcing value of $F = 12$ to assimilate the observations. The resulting model bias is very large, to emphasise the effect it has on data assimilation. The model error generated in this way is not random but systematic, so it is consistent with the theory. The accumulated model bias over the assimilation window can be estimated by initialising the biased model with the truth subsampled every 10 timesteps (the length of the assimilation window). Because the model is nonlinear, the accumulated model bias depends on the initialisation state. On average (over 50 assimilation windows and 40 state variables), the accumulated model bias is 0.337.

The observations at model timestep i observe all state variables directly at time i. Hence they are simulated from the truth at time i with added random error, $\mathbf{y}_i = m^{\dagger}_{0\to i}(\mathbf{x}^{\dagger}_0) + \epsilon_o$, where $\epsilon_o \sim N(\mathbf{b}, \mathbf{R})$. Herem $\mathbf{b}$ is the specified observation bias (zero for the anchor observations and 0.5 for the BC observations) and $\mathbf{R}$ is the specified observation-error covariance matrix. This is $\mathbf{I}_{40}$ for the BC observation. Two different experiments are performed with $\mathbf{R}_{\mathbb{L}} = 4\mathbf{I}_{40}$ and $\mathbf{R}_{\mathbb{L}} = 0.25\mathbf{I}_{40}$ to illustrate the role of the anchor observations. In practice, the BC and anchor observations only partially observe the state and the variables observed are not expected to coincide. The effect of more realistic observation operators can be considered by returning to Equation (11) and combining it with the theory developed in F23. The aim here is instead to focus on the timing of the anchor observations relative to the BC observations and the assimilation window.

In implementing VarBC, we correct the observation operator for each BC observation of the form $c(\beta) = \beta$. So in this case the true bias-correction coefficient to be retrieved is $\beta^{\dagger} = 0.5$.

The background-error covariance matrix, $\mathbf{B}_x$ is generated in the same way as in F23 (see appendix E therein) to ensure that it is consistent with the model dynamics. The derived $\mathbf{B}$ matrix has an error standard deviation of approximately 0.5 for each variable and a sharp drop in covariance, so that variables separated by more than two grid points are uncorrelated. At the initial time, the background is sampled from $N(\mathbf{0}, \mathbf{B})$, such that the background is initially unbiased but a bias in the background accumulates as the assimilation is cycled. The observations are assimilated using 4DVar with a window length of 10 timesteps for 10 assimilation cycles. This window length is short enough to be consistent with the assumption of linear propagation of errors that is at the heart of 4DVar and our developed theory.

For the first cycle, the background value $\boldsymbol{\beta}$ is given by the true value plus a random error drawn from $N(\mathbf{0}, B_\beta)$, where $B_\beta = 0.5$. Therefore, the results should be consistent with the theory in which we assumed $\langle \boldsymbol{\epsilon}^{\mathrm{b}}_{\beta} \rangle = 0$. As the system continues to cycle, this assumption will no longer hold.

4.1 | Comparison of different observing strategies

Following Section 3, we consider three cases in which the BC and anchor observations are made at different times within the assimilation window. In each case, the BC and

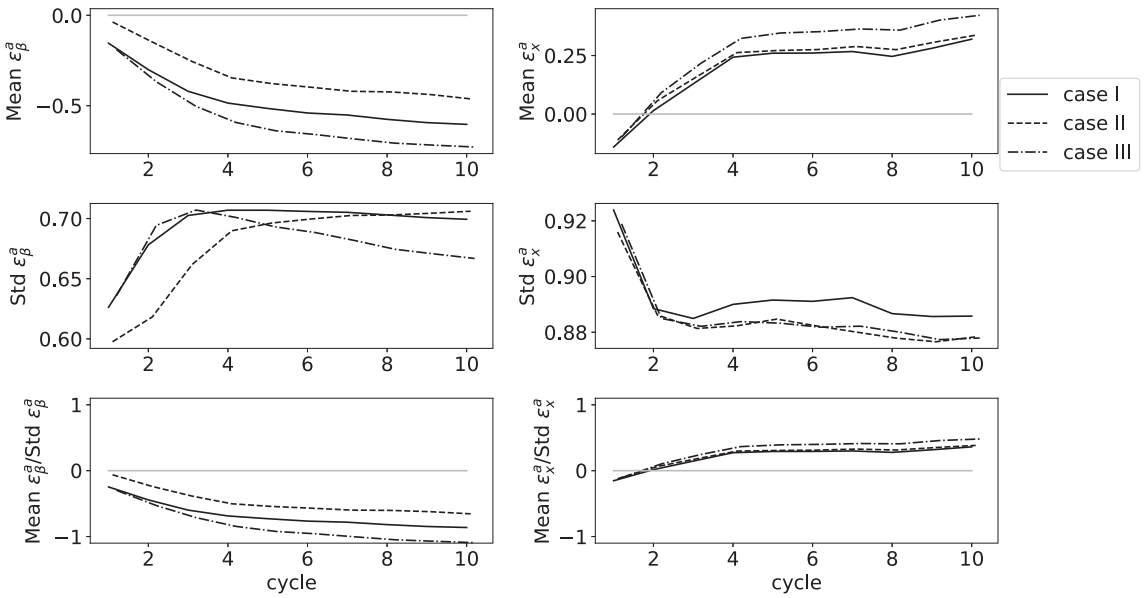

FIGURE 4 (Top row) Mean error, (middle row) error standard deviation, and (bottom row) the ratio of the two for analysis of (left) the bias coefficients and (right) the state averaged over 500 realisations for a cycled Lorenz 96 assimilation system. The error variance for the BC observations is 1 and that for the anchor observations is 4 ($\mathbf{R}_{\mathbb{L}} = 4\mathbf{I}_{40}$).

anchor observations are provided every 10 timesteps and observe every grid point.

- Case I: both the BC and anchor observations are at the end of assimilation window.
- Case II: the anchor observations are at the end of the window and the BC observations are halfway through the window.
- Case III: the anchor observations are halfway through the window and the BC observations are at the end of the window.

In Figures 4 and 5, the mean error and error standard deviation of the analysis of $\boldsymbol{\beta}$ (left) and $\mathbf{x}$ (right) are plotted for the three different observation configurations as a function of assimilation cycle. The statistics are averaged over 500 realisations of the random errors in the initial background and observations, and those for the analysis of the state are also averaged over the 40 variables.

In Figure 4, results are shown for the case in which the anchor observations are less precise than the BC observations. Consistent with the scalar illustrations and theory, we see that the smallest bias in the analysis of $\boldsymbol{\beta}$ occurs when the anchor observations are at the end of the window and the BC observations are halfway through the window (Case II). The largest bias in the analysis of $\boldsymbol{\beta}$ is when the BC observations are at the end of the window and the anchor observations are halfway through the window (Case III). As the assimilation is cycled, we see in each case that the magnitude of the bias in the analysis of $\boldsymbol{\beta}$ grows, and at the third cycle the bias drops below −0.5 for Case III. Recall that the true bias was 0.5, so an average error in $\boldsymbol{\beta}^{\mathrm{a}}$ equal to −0.5 will result in no correction on average and an average error in $\boldsymbol{\beta}^{\mathrm{a}}$ smaller than −0.5 will correct the observations in a direction away from the truth on average. However, in each case the standard deviation of the error for $\boldsymbol{\beta}^{\mathrm{a}}$ is larger than 0.5, so for each case there is a high probability that a single realisation of $\boldsymbol{\beta}^{\mathrm{a}}$ will correct the observations in a direction away from the truth.

For the early cycles, Case II provides the most precise analysis of $\boldsymbol{\beta}$ (see middle left panel). Again this is consistent with the scalar illustration. However, as the system is cycled and the biases in the analyses of $\boldsymbol{\beta}$ are propagated, we see that the consistency with the scalar illustration based on an initial cycle breaks down.

In the right-hand panels of Figure 4, we see that there is less sensitivity in the bias in the analysis of the state to the observing configuration. From the scalar example in Figure 2, we would have expected Case III to give the most accurate analysis of $\mathbf{x}$. However, as the DA system is cycled, Case III results in the largest bias in $\mathbf{x}^{\mathrm{a}}$. This can be explained by the inaccurate analysis of $\boldsymbol{\beta}$, which in this case is contaminated by model bias, and the accumulative effect of this inaccurate analysis of $\boldsymbol{\beta}$ as the DA system is cycled.

In Figure 5, results are shown for when the anchor observations are more precise than the BC observations. With more precise anchor observations, there is now a clearer separation between the three observing configurations for all of the statistics plotted. There is also less of an increase in the biases for $\boldsymbol{\beta}^{\mathrm{a}}$ and $\mathbf{x}^{\mathrm{a}}$ as the assimilation system is cycled.

As in Figure 4, we again see the smallest bias in the analysis of $\boldsymbol{\beta}$ in Case II and the largest bias in the analysis of

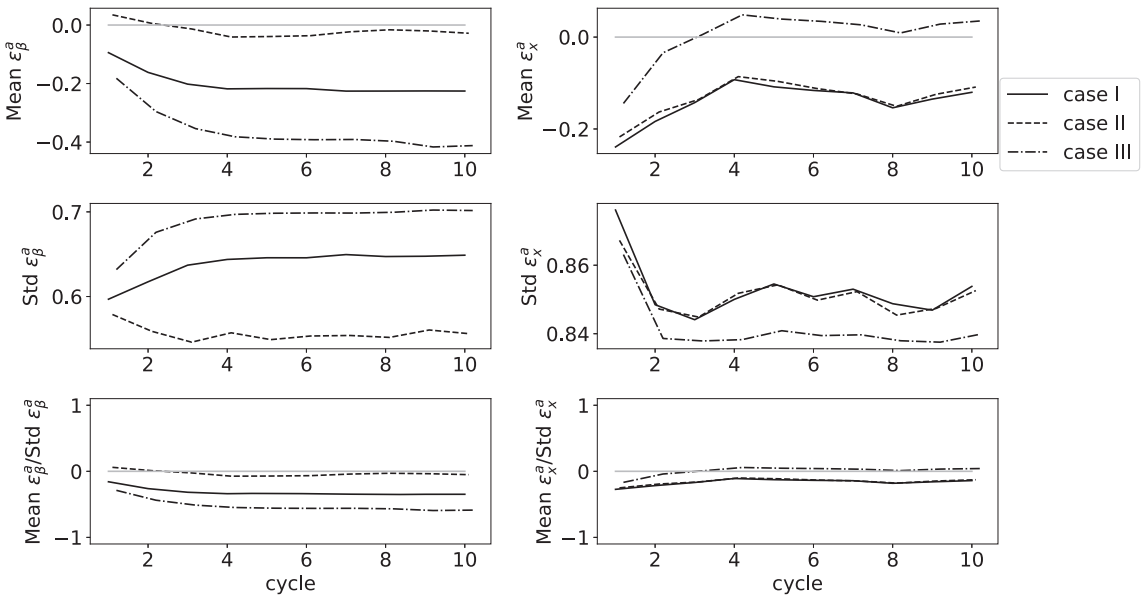

FIGURE 5 As in Figure 4 but with $\mathbf{R}_{\mathbb{L}} = 0.25\mathbf{I}_{40}$.

$\boldsymbol{\beta}$ in Case III. With more precise anchor observations, however, we now see that the bias in the analysis of $\boldsymbol{\beta}$ initially has a different sign in Case II compared with Cases I and III, which again is consistent with the scalar example, and can be explained by the contribution of the model bias only seen by the anchor observations having a different sign in this case (see Equation (17) and Figure 2).

With very precise anchor observations, we now see that Case III gives the smallest bias in the analysis state. Again, this is consistent with the scalar example and explained by the fact that the model bias can be expected to have a bigger impact on the analysis at the beginning of the window when trying to fit precise observations at the end of the window. The change in sign of the analysis bias for Cases I and II demonstrates that the analysis compensates for the positive model bias when fitting to the precise observations at the end of the window.

In each of the examples shown, we see that the ratio of the bias to the error standard deviation (plotted in the lower panels) in the analysis of $\boldsymbol{\beta}$ is generally larger than for the analysis of $\mathbf{x}$. This suggests that the potential biases in $\boldsymbol{\beta}^{\mathrm{a}}$ are more significant than those in $\mathbf{x}^{\mathrm{a}}$ and the benefits of controlling them with the choice of observing configuration are greater.

In addition to the experiments shown here, we also performed experiments more consistent with a real observing system, in which the state was partially observed by the anchor observations, indirectly observed by the BC observations, and either the anchor or BC observations are available at more than one time throughout the window. The results from these experiments (not shown) corroborated the conclusions from these experiments and the theory. That is, the most accurate estimate of $\boldsymbol{\beta}$ is achieved when anchor observations are available at the end of the window and the BC observations are earlier in the window. However, having the anchor observations observe different variables from the BC observations limits the ability of the anchor observations to constrain VarBC, as described in F23.

4.2 | Challenges in quantifying the value of anchor observations using the forecast sensitivity to observation impact

In this final results section, we demonstrate the difficulty in quantifying the value of the anchor observations in data assimilation with standard metrics. A popular metric that allows for the continual monitoring of the value of observations within a data assimilation system is the FSOI. The FSOI uses the adjoint of the forecast model to find the contribution of different subsets of observations to the change in forecast error due to the assimilation of all observations. A more thorough description of the method can be found in Langland and Baker (2004) and Lorenc and Marriott (2014).

Ideally, we would like the FSOI to be consistent with the impact of reducing a subset of observations as measured by data-denial experiments (DDEs: Eyre, 2021). DDEs allow for a direct assessment of the effect of changes to the observing system, such as reducing anchor observations, to quantify the importance of the removed observations. DDEs can be considered as a subset of observing-system experiments (OSEs), in which the impact of observations can be compared with a baseline in which observations are either added or removed. However, DDEs have a significant cost and hence only a limited number of experiments can be performed. This is why the FSOI has proven to be so popular in assessing the current observing network. Clearly, the FSOI and DDE measure different things and have been shown to have different sensitivities to model error and inconsistencies in the specification of the background and observation-error covariances (Daescu & Langland, 2022; Daescu & Todling, 2010; Eyre, 2024; Lupu *et al.*, 2015). Nevertheless, we would hope that we could still use the FSOI to assess the relative importance of different subsets of observations.

We show the results for two metrics of observation impact, computed in the Lorenz 96 model described at the beginning of this section for a forecast of length 20 timesteps. In generating the forecasts, we have used the biased model with $F = 12$. We define the forecast-error norm used to compute the two metrics as

$$e = \frac{1}{40}\sum_{i=1}^{40}(x_i^{\mathrm{f}} - x_i^{\mathrm{v}})^2,$$

where x_i^{f} is the forecast for the ith variable and x^{v} is a validation state.

The first metric is the FSOI computed change in e attributed to assimilation of either the anchor observations, $\delta e^{\mathrm{f}}(\mathbf{y}_{\perp})$, or BC observations, $\delta e^{\mathrm{f}}(\mathbf{y}_{\mathrm{BC}})$, as a percentage of the total change in forecast error due to the assimilation of all observations, $\delta e^{\mathrm{f}}(\mathrm{allobs})$:

$$\%FSOI(\mathbf{y}_{\perp}) \equiv 100\frac{\delta e^{\mathrm{f}}(\mathbf{y}_{\perp})}{\delta e^{\mathrm{f}}(\mathrm{allobs})} \quad \text{and}$$
$$\%FSOI(\mathbf{y}_{\mathrm{BC}}) \equiv 100\frac{\delta e^{\mathrm{f}}(\mathbf{y}_{\mathrm{BC}})}{\delta e^{\mathrm{f}}(\mathrm{allobs})}. \tag{21}$$

The second metric is the increase in e when either the anchor or BC observations are thinned so that only every other grid point is observed ($e^{\mathrm{f}}(\mathbf{y}_{\perp}\mathrm{thinned}) - e^{\mathrm{f}}(\mathrm{allobs})$ and $e^{\mathrm{f}}(\mathbf{y}_{\mathrm{BC}}\mathrm{thinned}) - e^{\mathrm{f}}(\mathrm{allobs})$, respectively) as a percentage

of the total forecast error, $e^{\mathrm{f}}(\mathrm{allobs})$, when both types of observations observe every grid point:

$$\%DDE(\mathbf{y}_{\perp}) \equiv 100\frac{e^{\mathrm{f}}(\mathbf{y}_{\perp}\mathrm{thinned}) - e^{\mathrm{f}}(\mathrm{allobs})}{e^{\mathrm{f}}(\mathrm{allobs})} \quad \text{and}$$
$$\%DDE(\mathbf{y}_{\mathrm{BC}}) \equiv 100\frac{e^{\mathrm{f}}(\mathbf{y}_{\mathrm{BC}}\mathrm{thinned}) - e^{\mathrm{f}}(\mathrm{allobs})}{e^{\mathrm{f}}(\mathrm{allobs})}. \quad (22)$$

We have defined the DDE in terms of halving the observations rather than completely removing one type, so that VarBC is still applicable when the observations are reduced and so that the assimilation systems are not dramatically changed in the DDE. Note that doubling $\%DDE(\mathbf{y}_{\perp})$ and $\%DDE(\mathbf{y}_{\mathrm{BC}})$ and then adding does not equal 100%, due to the nonlinearity of the DDE (Ishibashi, 2011), whereas the sum of $\%FSOI(\mathbf{y}_{\perp})$ and $\%FSOI(\mathbf{y}_{\mathrm{BC}})$ does equal 100% by definition.

For both metrics, the statistics are averaged over 10,000 assimilation cycles. The forecasts are validated against both the "truth" and the analysis valid at the time of the forecast.

In Figures 6 and 7, the DDE and FSOI metrics (Equations 21 and 22) are plotted against each other for the three different observing configurations described in Section 4.1. The solid markers indicate the case when the forecast is validated against the truth and the empty symbols that when the forecast is validated against the analysis. In Figure 6, the results are for the case when the anchor observations are less precise than the BC observations (consistent with Figure 4). In Figure 7, the results are for the case when the anchor observations are more precise than the BC observations (consistent with Figure 5).

We see a general lack of agreement between the two metrics regarding the relative importance of the anchor versus BC observations. The data-denial experiments always quantify a larger degradation in the forecast error when the anchor observations are reduced, compared with reducing the BC observations. In some situations, halving

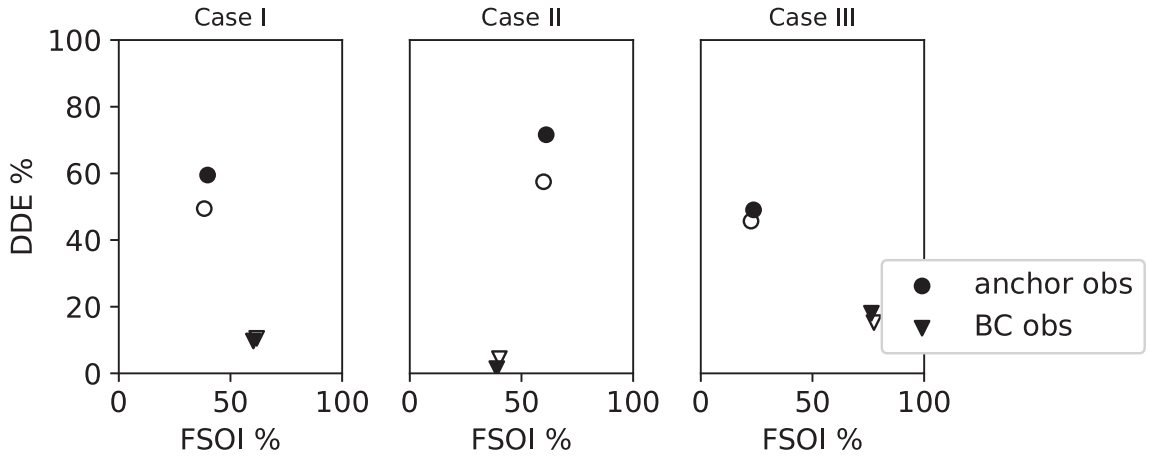

FIGURE 6 Comparison of the percentage of the FSOI (Equation 21) versus the percentage increase of forecast error when the observations are halved (Equation 22) for the anchor (circles) and BC (triangle) observations. The panels give results for the three different observation configurations. Full symbols validate the forecast against the truth. Empty symbols validate against the analysis. The error variance for the BC observations is 1 and that for the anchor observations is 4.

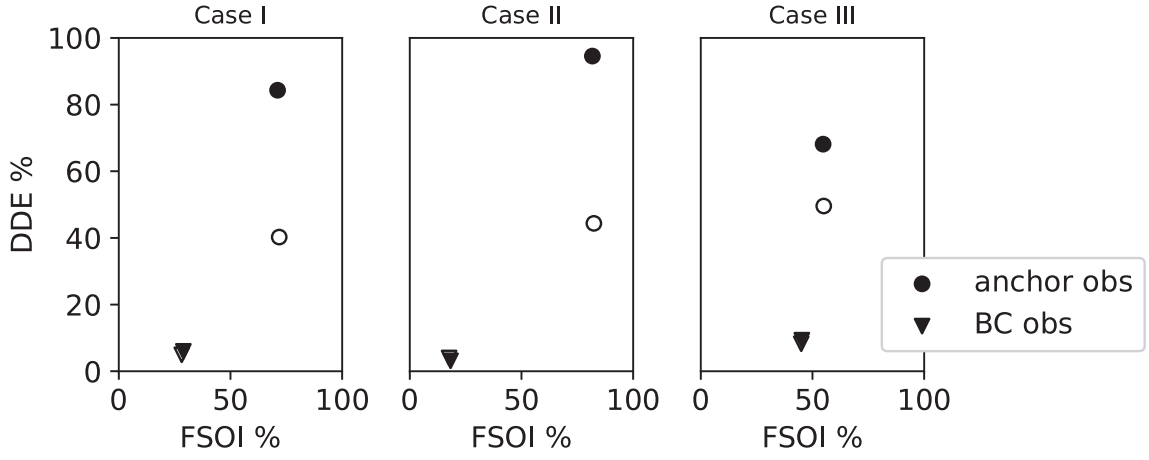

FIGURE 7 As in Figure 6, but with the error variance for the anchor observations set to 0.25.

the number of BC observations has a near-negligible effect on the forecast error measured by DDE (e.g., Case II in Figure 6). However, the FSOI sometimes assigns more impact to the BC observations than the anchor observations (e.g., Case III in Figure 6). The breakdown between relative importance of the two observation types in the metrics appears to be particularly large when model bias contamination is greatest and VarBC performance is at its worst, for example, Cases I and III when the anchor observations are less precise (cf. Figures 6 and 4). In contrast, the better performance of VarBC achieved with more precise anchor observations (cf. Figures 7 and 5) gives a better agreement between the relative importance of the anchor and BC observations as measured by the FSOI and DDE metrics. This is consistent with the conclusions of Eyre (2024), in which it was shown that FSOI gives erroneous results when the observation-error variances are misspecified.

The use of the truth or analysis for validation seems to have the biggest impact on the DDE results for the anchor observations. Validating against the analysis diagnoses a smaller increase in forecast error when halving the anchor observations than validating against the truth, especially when the anchor observations are more precise (see Figure 7).

The inability of the FSOI to capture the relative value of the anchor observations as measured by DDE could be anticipated. By its construction, the FSOI is unable to account for the cumulative effect of the anchoring observations, which is represented by DDE. However, in practice, this means that the FSOI is often dominated by radiance data (Candy *et al.*, 2021), leaving it open for the FSOI results to be misinterpreted and the anchor observations to be undervalued when different types of observations are ranked by their contribution to reducing forecast errors. Without the anchor observations, VarBC would not be able to correct the radiance data successfully and so the large FSOI scores for radiance data should in part be attributed to the anchor observations. For both the FSOI and DDE (as we defined it in Equation 22), it is difficult to separate the anchoring effect of the observations. One approach is to look directly at the contributions of the different observation types to the bias-correction coefficients, as proposed in Ishibashi (2011).

5 | SUMMARY OF RESULTS

To maximise the positive impact of satellite radiances in NWP assimilation, the correction of biases in these radiances should be minimally contaminated with model bias. A lack of theory has previously limited ability to assess the effect of model and background bias on the precision and accuracy of radiance bias corrections, $\boldsymbol{\beta}^{\mathrm{a}}$. We have shown that the accuracy of $\boldsymbol{\beta}^{\mathrm{a}}$ (its bias in estimating the true value, Equation 11) can be written as a sum of terms depending on the background bias, the model bias seen by both the anchor and BC observations, and the model bias seen by only one observation type (depending on which is later in the assimilation window).

We demonstrated that the anchor observations can play a clear role in reducing the effect of the first two bias sources (the background bias and the model bias seen by both the anchor and BC observations). This depends on the spreading of information from the anchor to the BC observed variables as controlled by the strength of the background-error correlations between the BC and anchor observed variables ($\widehat{\mathbf{H}}_{\mathrm{BC}}\mathbf{B}_x\widehat{\mathbf{H}}_{\perp}$). The strength of the correlations described by $\widehat{\mathbf{H}}_{\mathrm{BC}}\mathbf{B}_x\widehat{\mathbf{H}}_{\perp}$ will in part be determined by the relative spatial and multivariate sampling of the anchor and BC observations as investigated in F23. In this current work, we concentrate on the temporal distribution of the observations and as such $\widehat{\mathbf{H}}_{\mathrm{BC}}\mathbf{B}_x\widehat{\mathbf{H}}_{\perp}$ will include the evolved temporal error correlations given by $\mathbf{M}_{0\to t_{\mathrm{BC}}}\mathbf{B}_x\mathbf{M}^{\mathrm{T}}_{0\to t_{\perp}}$.

If the BC observations see an additional model error, then more precise anchor observations cannot reduce the effect of this bias source but can increase the effect via an increased sensitivity of the bias-correction coefficient to the BC observations, $\mathbf{K}_{\beta,\mathrm{BC}}$ (see Equation 19). If the anchor observations see an additional model error, then, because this term has a different sign from the other terms, more precise anchor observations can potentially reduce the overall bias in the analysis of the bias coefficients, $\boldsymbol{\beta}^{\mathrm{a}}$ (see Equation 17).

The precision of $\boldsymbol{\beta}^{\mathrm{a}}$ (the error variance) was shown to depend on the complementary role of the information about $\boldsymbol{\beta}$ in the BC observations and the references of the background state and the anchor observations. It was found that having the BC observations earlier in the window gave a more precise $\boldsymbol{\beta}^{\mathrm{a}}$, because the uncertainty in the background state propagated to the time of the BC observations is smaller. The precision of $\boldsymbol{\beta}^{\mathrm{a}}$ was also greater if the anchor observations are later in the window, because the information the anchor observations provide about $\boldsymbol{\beta}$ would be increased when back-propagated to the time of the BC observations. This was illustrated for a scalar case in Section 3.4. This was derived assuming that the model error is not random, but, if the model error were random in addition to biased, then this would only enhance this effect, because the growth in uncertainty with time would be greater.

We illustrated how these results compare with the impact of the anchor observations on the accuracy and precision of the analysis of the state, $\mathbf{x}^{\mathrm{a}}$, which is ultimately the metric of interest, in the presence of model

bias—a much more studied problem (Gauthier *et al.*, 2007; McNally, 2019). We demonstrated how precise anchor observations at the end of the assimilation window, although improving the accuracy of $\boldsymbol{\beta}^{a}$, can degrade the accuracy of $\mathbf{x}^{a}$.

These analytical results were then used to explain the behaviour of VarBC in cycled 4DVar experiments using the Lorenz 96 model when model bias is generated using an incorrect forcing parameter in the assimilation of the data. In a cycled system, the analyses of $\boldsymbol{\beta}$ and $\mathbf{x}$ will interact, so we cannot focus on the accuracy of one and not the other. The breakdown of some of the conclusions from the analytical results as the assimilation system cycles highlighted, in particular, the importance of VarBC for reducing the bias in the analysis of the state.

Lastly, using the same Lorenz 96 model experimental setup, we demonstrated how the use of FSOI and data-denial experiments, common metrics of observation impact, can give misleading results regarding the role of anchor observations. The breakdown in the relative importance of the anchor and BC observations as measured by the FSOI and data-denial experiments was seen to be greatest in the cases in which the contamination from model bias was greatest.

6 | DISCUSSION AND CONCLUSIONS

The correction of biases in radiance data, via VarBC, is essential for their successful assimilation and is fundamental to the skill of modern-day numerical weather prediction. VarBC is formulated assuming that the numerical model is unbiased. However, this assumption is often known to be far from valid. In formulating VarBC, the bias predictors used to model the characteristics of the radiance biases are chosen to avoid predictors in which model bias is expected, for example, those involving humidity, instead focusing on variables where we can have the most confidence that the model bias is small, for example, large-scale averages of the background temperature field (Eyre, 1992). Similarly, the performance of VarBC can also be improved if the radiances used to compute the bias corrections are limited to regions where the model bias is less significant, for example, limiting the radiances to those over the sea, where the surface skin temperature is more accurately known compared with over land (Eyre, 1992), or to those that are not affected by cloud, when all-sky radiances are assimilated. However, these compromises are not trivial to implement and VarBC still depends on the availability of another reference for the true state.

Due to the limited reliability of metrics for assessing the value of anchor observations in NWP, this article instead took the approach of developing and exemplifying new theory to understand how to optimise the impact of anchor observations to minimise the contamination of model bias in VarBC. This new theory explores this question by looking at the sensitivity of the contamination of VarBC by model bias to the weighting of the anchor observations. In the theoretical results, the weighting was controlled by changing the precision of the anchor observations, but, in practice, we could also consider the weighting as a function of the number of anchor observations (with similar precision) assimilated. In this way, we have shown the importance of how much of the model bias seen by the BC observations is also seen by the anchor observations, concluding that, in general, it is better to have anchor observations at the end of the assimilation window so that they can see the maximum amount of model bias and pass this information to VarBC. It is also beneficial to have the observations that are to be bias-corrected closer to the beginning of the window, so that the background initial state provides a more useful reference. These results are complementary to our previous work focused on the spatial distribution of the anchor observations (F23).

Our work leads to several practical conclusions for maximising the use of anchor observations. Firstly, in the current observing system we have a good temporal coverage of both the anchor and radiance observations for temperature throughout the assimilation window. However, in applying VarBC to reanalysis there is a sudden drop off in the number of anchor observations pre-2006 due to the lack of GNSS-RO observations. Therefore, the theory presented here is crucial to understanding how to maximise the benefit of a smaller number of anchor observations and to help guide where efforts should be made in the pre-processing of observations to increase the historical anchor observation network: for example, choosing the assimilation windows such that the available anchor observations are more likely to fall at the end of the window. Secondly, looking to the future, the number of instruments on board satellites is increasing. With this, we can expect the proportion of satellite radiances assimilated to increase, as well as a greater variety in the variables to be estimated. Therefore, even with the current GNSS-RO network, there should be a continued effort to understand how the anchor observing network should evolve.

Within the theory presented, we did not consider the possibility that the anchor observations (i.e., any observations not bias-corrected) could themselves be biased. The impact of this source of bias in VarBC can again be expected to depend on the spreading of information between the anchor and BC observations and the weighting of the anchor observations. As could be anticipated,

the more precise the anchor observations are (i.e., the more weight they are given), the greater the impact of their biases. Therefore, a clear route for the ongoing improvement of VarBC is the correction of biases in each observation type. For some observations, this may involve extending VarBC if adaptive bias correction is needed (e.g., for atmospheric motion vectors: Hoffman *et al.*, 2022, 2024). However, to maintain a significant proportion of anchor observations with good spatial coverage of the key variables, it is also necessary to increase the network of anchor observations through the better pre-processing of observations with biases that can be corrected offline: for example, aircraft observations (de Haan *et al.*, 2022).

Some operational centres are moving towards online bias correction of both the observations and the model using weak-constraint methods. This is a promising avenue for allowing a greater separability of model and observation bias by formulating a well-constructed prior for each term (Lorente-Plazas & Hacker, 2017). However, the need for a network of anchor observations will persist. Preliminary numerical results have shown that, even in the case in which both the model and observation biases are corrected, the conclusions from this article hold (Francis, 2023, ch. 8), that is, a more accurate estimate of both biases can be obtained when the anchor observations are later in the assimilation window and the observations to be bias-corrected are earlier.

Model biases are a widespread concern at operational centres. This work highlights how careful consideration of the timing and coverage of anchoring observations can help mitigate the negative impacts of model bias on forecast skill.

ACKNOWLEDGEMENTS

This study was funded as part of NERC's support of the National Centre for Earth Observation, contract number PR140015, and the SCENARIO NERC Doctoral Training Partnership grant NE/S007261/1, with support by a CASE studentship from the Met Office. We thank the two anonymous reviewers for their thoughtful feedback on an earlier version of this article.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Alison M. Fowler https://orcid.org/0000-0003-3650-3948
Amos S. Lawless https://orcid.org/0000-0002-3016-6568
J. R. Eyre https://orcid.org/0000-0002-9877-392X
Stefano Migliorini https://orcid.org/0000-0002-9489-9867

REFERENCES

Auligné, T., McNally, A.P. & Dee, D.P. (2007) Adaptive bias correction for satellite data in a numerical weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, 133, 631–642. Available from: https://doi.org/10.1002/qj.56

Bouttier, F. & Kelly, G. (2001) Observing-system experiments in the ECMWF 4D-Var data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 127, 1469–1488. Available from: https://doi.org/10.1002/qj.49712757419

Buontempo, C., Jupp, A. & Rennie, M. (2008) Operational NWP assimilation of GPS radio occultation data. *Atmospheric Science Letters*, 9, 129–133. Available from: https://doi.org/10.1002/asl.173

Cameron, J. & Bell, W. (2018) *The testing and implementation of variational bias correction (VarBC) in the Met Office global NWP system. Weather Science Technical Report 631*. Exeter: Met Office. Available from: https://library.metoffice.gov.uk/Portal/Default/en-GB/RecordView/Index/633663

Candy, B., Cotton, J. & Eyre, J. (2021) *Recent results of observation data denial experiments. Forecasting Research Technical Report 641*. Exeter: Met Office. Available from: https://library.metoffice.gov.uk/Portal/Default/en-GB/RecordView/Index/645667

Cardinali, C. (2009) Monitoring the observation impact on the short-range forecast. *Quarterly Journal of the Royal Meteorological Society*, 135, 239–250. Available from: https://doi.org/10.1002/qj.366

Cardinali, C. & Healy, S. (2014) Impact of GPS radio occultation measurements in the ECMWF system using adjoint-based diagnostics. *Quarterly Journal of the Royal Meteorological Society*, 140, 2315–2320. Available from: https://doi.org/10.1002/qj.2300

Daescu, D.N. & Langland, R.H. (2022) An FSO-Based Optimization Framework for Improved Observation Performance: Theoretical Formulation and Experiments with NAVDAS-AR/NAVGEM. *Monthly Weather Review*, 150, 1335–1353. Available from: https://doi.org/10.1175/MWR-D-21-0305.1

Daescu, D.N. & Todling, R. (2010) Adjoint sensitivity of the model forecast to data assimilation system error covariance parameters. *Quarterly Journal of the Royal Meteorological Society*, 136, 2000–2012. Available from: https://doi.org/10.1002/qj.693

de Haan, S., de Jong, P.M.A. & van der Meulen, J. (2022) Characterizing and correcting the warm bias observed in Aircraft Meteorological Data Relay (AMDAR) temperature observations. *Atmospheric Measurement Techniques*, 15, 811–818. Available from: https://doi.org/10.5194/amt-15-811-2022

Dee, D.P. (2005) Bias and data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 131, 3323–3343. Available from: https://doi.org/10.1256/qj.05.137

Derber, J.C. & Wu, W.-S. (1998) The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Monthly Weather Review*, 126, 2287–2299. Available from: https://doi.org/10.1175/1520-0493(1998)126<2287:TUOTCC>2.0.CO;2

Eyre, J.R. (1992) *A bias correction scheme for simulated TOVS brightness temperatures. Technical Memorandum 186*. Reading: ECMWF. Available from: https://doi.org/10.21957/tmhrqv5cp

Eyre, J.R. (2016) Observation bias correction schemes in data assimilation systems: a theoretical study of some of their properties. *Quarterly Journal of the Royal Meteorological Society*, 142, 2284–2291. Available from: https://doi.org/10.1002/qj.2819

Eyre, J.R. (2021) Observation impact metrics in NWP: A theoretical study. Part I: Optimal systems. *Quarterly Journal of the Royal Meteorological Society*, 147, 3180–3200. Available from: https://doi.org/10.1002/qj.4123

Eyre, J.R. (2024) Observation impact metrics in NWP: a theoretical study. Part II: systems with suboptimal observation errors. *Quarterly Journal of the Royal Meteorological Society*, 250, 632–640. Available from: https://doi.org/10.1002/qj.4614

Eyre, J.R., Bell, W., Cotton, J., English, S.J., Forsythe, M., Healy, S.B. et al. (2022) Assimilation of satellite data in numerical weather prediction. Part II: Recent years. *Quarterly Journal of the Royal Meteorological Society*, 148, 521–556. Available from: https://doi.org/10.1002/qj.4228

Francis, D. (2023) Bias correction of satellite data in data assimilation for numerical weather prediction. https://doi.org/10.48683/1926.00114305

Francis, D.J., Fowler, A.M., Lawless, A.S., Eyre, J. & Migliorini, S. (2023) The effective use of anchor observations in variational bias correction in the presence of model bias. *Quarterly Journal of the Royal Meteorological Society*, 149, 1789–1809. Available from: https://doi.org/10.1002/qj.4482

Gauthier, P., Tanguay, M., Laroche, S., Pellerin, S. & Morneau, J. (2007) Extension of 3DVAR to 4DVAR: Implementation of 4DVAR at the Meteorological Service of Canada. *Monthly Weather Review*, 135, 2339–2354. Available from: https://doi.org/10.1175/MWR3394.1

Haimberger, L., Tavolato, C. & Sperka, S. (2012) Homogenization of the global radiosonde temperature dataset through combined comparison with reanalysis background series and neighboring stations. *Journal of Climate*, 25, 8108–8131. Available from: https://doi.org/10.1175/JCLI-D-11-00668.1

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J. et al. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049. Available from: https://doi.org/10.1002/qj.3803

Hoffman, R.N., Liu, H., Lukens, K.E., Garrett, K. & Ide, K. (2024) Assimilating atmospheric motion vector winds using a feature track correction observation operator. *Quarterly Journal of the Royal Meteorological Society*, 150, 5074–5093. Available from: https://doi.org/10.1002/qj.4857

Hoffman, R.N., Lukens, K.E., Ide, K. & Garrett, K. (2022) A collocation study of atmospheric motion vectors (AMVs) compared to Aeolus wind profiles with a feature track correction (FTC) observation operator. *Quarterly Journal of the Royal Meteorological Society*, 148, 321–337. Available from: https://doi.org/10.1002/qj.4207

Howes, K.E., Fowler, A.M. & Lawless, A.S. (2017) Accounting for model error in strong-constraint 4D-Var data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 143, 1227–1240. Available from: https://doi.org/10.1002/qj.2996

Ingleby, B., Rodwell, M. & Isaksen, L. (2016) Global radiosonde network under pressure. ECMWF newsletter 149, ECMWF.

Ishibashi, T. (2011) Tangent linear approximation based observation data impact estimation in 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, 137, 1898–1912. Available from: https://doi.org/10.1002/qj.871

Kalnay, E. (2003) *Atmospheric modeling, data assimilation and predictability*. New York: Cambridge University Press.

Kalnay, E., Ota, Y., Miyoshi, T. & Liu, J. (2012) A simpler formulation of forecast sensitivity to observations: application to ensemble Kalman filters. *Tellus A: Dynamic Meteorology and Oceanography*, 64, 18462. Available from: https://doi.org/10.3402/tellusa.v64i0.18462

Kull, D., Riishojgaard, L.P., Eyre, J. & Varley, R.A. (2021) *The value of surface-based meteorological observation data. Technical report*. Washington: World Bank. Available from: https://hdl.handle.net/10986/35178

Laloyaux, P., Bonavita, M., Chrust, M. & Gurol, S. (2020a) Exploring the potential and limitations of weak-constraint 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, 146, 4067–4082. Available from: https://doi.org/10.1002/qj.3891

Laloyaux, P., Bonavita, M., Dahoui, M., Farnan, J., Healy, S., Hólm, E. et al. (2020b) Towards an unbiased stratospheric analysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 2392–2409. Available from: https://doi.org/10.1002/qj.3798

Langland, R.H. & Baker, N. (2004) Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, 56, 189–201. Available from: https://doi.org/10.3402/tellusa.v56i3.14413

Li, H., Kalnay, E., Miyoshi, T. & Danforth, C.M. (2009) Accounting for model errors in ensemble data assimilation. *Monthly Weather Review*, 137, 3407–3419. Available from: https://doi.org/10.1175/2009MWR2766.1

Lorenc, A.C. (1986) Analysis method for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 112, 1177–1194. Available from: https://doi.org/10.1002/qj.49711247414

Lorenc, A.C. & Marriott, R.T. (2014) Forecast sensitivity to observations in the Met Office Global numerical weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, 140, 209–224. Available from: https://doi.org/10.1002/qj.2122

Lorente-Plazas, R. & Hacker, J.P. (2017) Observation and Model Bias Estimation in the Presence of Either or Both Sources of Error. *Monthly Weather Review*, 145, 2683–2696. Available from: https://doi.org/10.1175/MWR-D-16-0273.1

Lorenz, E.N. (1995) Predictability-a problem partly solved. In: *ECMWF seminar proceedings I*, Vol. 1, pp. 1–18. Reading: ECMWF. Available from: https://www.ecmwf.int/en/elibrary/75462-predictability-problem-partly-solved

Lupu, C., Cardinali, C. & McNally, A.P. (2015) Adjoint-based forecast sensitivity applied to observation-error variance tuning. *Quarterly Journal of the Royal Meteorological Society*, 141, 3157–3165. Available from: https://doi.org/10.1002/qj.2599

McNally, A.P. (2019) On the sensitivity of a 4D-Var analysis system to satellite observations located at different times within the assimilation window. *Quarterly Journal of the Royal Meteorological Society*, 145, 2806–2816. Available from: https://doi.org/10.1002/qj.3596

Poli, P., Healy, S.B. & Dee, D.P. (2010) Assimilation of Global Positioning System radio occultation data in the ECMWF ERA-Interim reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 136, 1972–1990. Available from: https://doi.org/10.1002/qj.722

Rabier, F. (2005) Overview of global data assimilation developments in numerical weather-prediction centres. *Quarterly Journal of the Royal Meteorological Society*, 131, 3215–3233. Available from: https://doi.org/10.1256/qj.05.129

Radnóti, G., Bauer, P., McNally, A. & Horányi, A. (2012) ECMWF study to quantify the interaction between terrestrial and space-based observing systems on Numerical Weather Prediction skill. In: *Technical memorandum*, Vol. 679. Reading: ECMWF. Available from: https://doi.org/10.21957/yvf1ceq7i

Sun, B., Reale, A., Schroeder, S., Seidel, D.J. & Ballish, B. (2013) Toward improved corrections for radiation-induced biases in radiosonde temperature observations. *Journal of geophysical research: Atmospheres*, 118, 4231–4243. Available from: https://doi.org/10.1002/jgrd.50369

Trémolet, Y. (2006) Accounting for an imperfect model in 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, 132, 2483–2504. Available from: https://doi.org/10.1256/qj.05.224

WMO. (2024) *Proceedings of the 8th WMO Workshop on the Impact of Various Observing systems on Numerical Weather Prediction and Earth System Prediction. Technical report*. Geneva: WMO. Available from: https://community.wmo.int/en/meetings/8th-wmo-impact-workshop-home

World Meteorological Organization. (2020) *Vision for the wmo integrated global observing system in 2040. Technical report, WMO-No. 1243*. Geneva: WMO. Available from: https://library.wmo.int/records/item/57028-vision-for-the-wmo-integrated-global-observing-system-in-2040

How to cite this article: Fowler, A.M., Francis, D.J., Lawless, A.S., Eyre, J.R. & Migliorini, S. (2025) The importance of the timing of anchor observations in 4D variational bias correction: Theory and idealised experiments. *Quarterly Journal of the Royal Meteorological Society*, e5043. Available from: https://doi.org/10.1002/qj.5043

APPENDIX A. DERIVATION OF THE ERROR IN THE ANALYSIS

The expression for the analysis given by Equation (3) can be separated out according to the bias-corrected observations, $\hat{\mathbf{y}}_{\mathrm{BC}}$, and anchor observations $\hat{\mathbf{y}}_{\downarrow}$ (i.e., all obs that are not corrected by VarBC):

$$\mathbf{v}^{\mathrm{a}} = \mathbf{v}^{\mathrm{b}} + \mathbf{K}_{\mathrm{v}} \begin{pmatrix} \hat{\mathbf{y}}_{\mathrm{BC}} - \hat{h}_{\mathrm{BC}}(\mathbf{v}^{\mathrm{b}}) \\ \hat{\mathbf{y}}_{\downarrow} - \hat{h}_{\downarrow}(\mathbf{x}^{\mathrm{b}}) \end{pmatrix}, \tag{A.1}$$

where $\hat{h}_{\mathrm{BC}}$ is the 4D observation operator mapping the control vector to the BC observations (i.e., it includes a bias correction) and $\hat{h}_{\downarrow}$ is the 4D observation operator mapping the initial state to the anchor observations (i.e., it does not include a bias correction and so is a function of $\mathbf{x}$ only).

To derive the error in the analysis of the bias coefficients, we first define the following errors.

- Error in the background augmented state vector: $\boldsymbol{\epsilon}_{\mathrm{v}}^{\mathrm{b}} = \mathbf{v}^{\mathrm{b}} - \mathbf{v}^{\dagger}$, where $\mathbf{v}^{\dagger}$ is the "true" control vector. This can be separated in terms of the error in the background state and the background bias-correction coefficients, $\boldsymbol{\epsilon}_{\mathrm{v}}^{\mathrm{b}} = ((\boldsymbol{\epsilon}_{x}^{\mathrm{b}})^{\mathrm{T}}, (\boldsymbol{\epsilon}_{\beta}^{\mathrm{b}})^{\mathrm{T}})^{\mathrm{T}}$.
- Instrument error for the BC observations: $\hat{\boldsymbol{\epsilon}}_{\mathrm{BC}} = \hat{\mathbf{y}}_{\mathrm{BC}} - \hat{h}_{\mathrm{BC}}^{\dagger}(\mathbf{v}^{\dagger})$, where $\hat{h}_{\mathrm{BC}}^{\dagger}$ is the "true" observation operator for the BC obs (i.e., it uses the correct model to propagate the initial state and corrects for the true radiance bias).
- Instrument error for the anchor observations: $\hat{\boldsymbol{\epsilon}}_{\downarrow} = \hat{\mathbf{y}}_{\downarrow} - \hat{h}_{\downarrow}^{\dagger}(\mathbf{x}_0^{\dagger})$, where $\hat{h}_{\downarrow}^{\dagger}$ is the "true" observation operator for the anchor obs acting on the "true" state ($\mathbf{x}^{\dagger}$).
- Model error as "seen" by the BC observations: $\hat{\boldsymbol{\eta}}_{\mathrm{BC}} = \hat{h}_{\mathrm{BC}}(\mathbf{v}^{\dagger}) - \hat{h}_{\mathrm{BC}}^{\dagger}(\mathbf{v}^{\dagger})$ (see Section A.3 for further explanation).
- Model error as "seen" by the anchor observations: $\hat{\boldsymbol{\eta}}_{\downarrow} = \hat{h}_{\downarrow}(\mathbf{x}_0^{\dagger}) - \hat{h}_{\downarrow}^{\dagger}(\mathbf{x}_0^{\dagger})$ (see Section A.3 for further explanation).

The error in the analysis is then given by

$$\begin{aligned} \boldsymbol{\epsilon}_{\mathrm{v}}^{\mathrm{a}} &= \mathrm{v}^{\mathrm{a}} - \mathrm{v}^{\dagger} \\ &= \mathrm{v}^{\mathrm{b}} - \mathrm{v}^{\dagger} + \mathrm{K}_{\mathrm{v}} \begin{pmatrix} \hat{\mathrm{y}}_{\mathrm{BC}} - \hat{h}_{\mathrm{BC}}^{\dagger}(\mathrm{v}^{\mathrm{t}}) + \hat{h}_{\mathrm{BC}}^{\dagger}(\mathrm{v}^{\dagger}) - \hat{h}_{\mathrm{BC}}(\mathrm{v}^{\dagger}) + \hat{h}_{\mathrm{BC}}(\mathrm{v}^{\dagger}) - \hat{h}_{\mathrm{BC}}(\mathrm{v}^{\mathrm{b}}) \\ \hat{\mathrm{y}}_{\downarrow} - \hat{h}_{\downarrow}^{\dagger}(\mathrm{x}^{\dagger}) + \hat{h}_{\downarrow}^{\dagger}(\mathrm{x}^{\dagger}) - \hat{h}_{\downarrow}(\mathrm{x}^{\dagger}) + \hat{h}_{\downarrow}(\mathrm{x}^{\dagger}) - \hat{h}_{\downarrow}(\mathrm{x}^{\mathrm{b}}) \end{pmatrix} \\ &= \boldsymbol{\epsilon}_{\mathrm{v}}^{\mathrm{b}} + \mathrm{K}_{\mathrm{v}} \begin{pmatrix} \hat{\epsilon}_{\mathrm{BC}} - \hat{\boldsymbol{\eta}}_{\mathrm{BC}} - \hat{\mathrm{H}}_{\mathrm{BC}} \boldsymbol{\epsilon}_{x}^{\mathrm{b}} - \hat{\mathrm{C}}_{\beta} \boldsymbol{\epsilon}_{\beta}^{\mathrm{b}} \\ \hat{\epsilon}_{\downarrow} - \hat{\boldsymbol{\eta}}_{\downarrow} - \hat{\mathrm{H}}_{\downarrow} \boldsymbol{\epsilon}_{x}^{\mathrm{b}} \end{pmatrix}, \end{aligned} \tag{A.2}$$

where $\hat{\mathbf{H}}_{\mathrm{BC}}$ and $\hat{\mathbf{H}}_{\downarrow}$ are the Jacobians of the 4D observation operators with respect to the state for the BC and anchor observations, respectively, and $\hat{\mathbf{C}}_{\beta}$ is the Jacobian of the bias-correction term with respect to the bias-correction coefficients.

Expressions for the error in the analysis of the bias-correction coefficient and state can then be isolated by separating out the Kalman gain matrix as

$$\mathbf{K}_{\mathrm{v}} = \begin{pmatrix} \mathbf{K}_{x,\mathrm{BC}} & \mathbf{K}_{x,\mathrm{anc}} \\ \mathbf{K}_{\beta,\mathrm{BC}} & \mathbf{K}_{\beta,\mathrm{anc}} \end{pmatrix}. \tag{A.3}$$

A.1 The error in the analysis of the bias coefficient

Substituting Equation (A.3) into Equation (A.2), we first obtain an expression for the error in the analysis of the bias-correction coefficient:

$$\begin{aligned} \boldsymbol{\epsilon}_{\beta}^{\mathrm{a}} = \boldsymbol{\epsilon}_{\beta}^{\mathrm{b}} &+ \mathbf{K}_{\beta,\mathrm{BC}} \left(\hat{\boldsymbol{\epsilon}}_{\mathrm{BC}} - \hat{\boldsymbol{\eta}}_{\mathrm{BC}} - \hat{\mathbf{H}}_{\mathrm{BC}} \boldsymbol{\epsilon}_{x}^{\mathrm{b}} - \hat{\mathbf{C}}_{\beta} \boldsymbol{\epsilon}_{\beta}^{\mathrm{b}} \right) \\ &+ \mathbf{K}_{\beta,\mathrm{anc}} \left(\hat{\boldsymbol{\epsilon}}_{\downarrow} - \hat{\boldsymbol{\eta}}_{\downarrow} - \hat{\mathbf{H}}_{\downarrow} \boldsymbol{\epsilon}_{x}^{\mathrm{b}} \right). \end{aligned} \tag{A.4}$$

In F23 it was shown that

$$\mathbf{K}_{\beta,\mathrm{anc}} = -\mathbf{K}_{\beta,\mathrm{BC}} \hat{\mathbf{H}}_{\mathrm{BC}} \mathbf{D}, \tag{A.5}$$

where $\mathbf{D} = \mathbf{B}_x\hat{\mathbf{H}}_{\downarrow}^{\mathrm{T}}(\hat{\mathbf{H}}_{\downarrow}\mathbf{B}_x\hat{\mathbf{H}}_{\downarrow}^{\mathrm{T}} + \hat{\mathbf{R}}_{\downarrow})^{-1}$. This allows us to rewrite and simplify (A.4) as

$$\boldsymbol{\epsilon}_{\beta}^{\mathrm{a}} = \boldsymbol{\epsilon}_{\beta}^{\mathrm{b}} + \mathbf{K}_{\beta,\mathrm{BC}}\Big(\hat{\boldsymbol{\epsilon}}_{\mathrm{BC}} - \hat{\boldsymbol{\eta}}_{\mathrm{BC}} - \hat{\mathbf{C}}_{\beta}\boldsymbol{\epsilon}_{\beta}^{\mathrm{b}} - \hat{\mathbf{H}}_{\mathrm{BC}}\mathbf{D}\hat{\boldsymbol{\epsilon}}_{\downarrow} + \hat{\mathbf{H}}_{\mathrm{BC}}\mathbf{D}\hat{\boldsymbol{\eta}}_{\downarrow} - \hat{\mathbf{H}}_{\mathrm{BC}}(\mathbf{I}\text{-}\mathbf{D}\hat{\mathbf{H}}_{\downarrow})\boldsymbol{\epsilon}_{x}^{\mathrm{b}}\Big). \tag{A.6}$$

A.2 The error in the analysis of the state

Following the same methodology as in Section A.1, we can obtain an expression for the error in the analysis of the state:

$$\boldsymbol{\epsilon}_{x}^{\mathrm{a}} = \boldsymbol{\epsilon}_{x}^{\mathrm{b}} + \mathbf{K}_{x,\mathrm{BC}}\Big(\hat{\boldsymbol{\epsilon}}_{\mathrm{BC}} - \hat{\boldsymbol{\eta}}_{\mathrm{BC}} - \hat{\mathbf{H}}_{\mathrm{BC}}\boldsymbol{\epsilon}_{x}^{\mathrm{b}} - \hat{\mathbf{C}}_{\beta}\boldsymbol{\epsilon}_{\beta}^{\mathrm{b}}\Big) + \mathbf{K}_{x,\mathrm{anc}}\Big(\hat{\boldsymbol{\epsilon}}_{\downarrow} - \hat{\boldsymbol{\eta}}_{\downarrow} - \hat{\mathbf{H}}_{\downarrow}\boldsymbol{\epsilon}_{x}^{\mathrm{b}}\Big). \tag{A.7}$$

From F23 we can also derive the relationship

$$\mathbf{K}_{x,\mathrm{anc}} = (\mathbf{I}_n - \mathbf{K}_{x,\mathrm{BC}}\hat{\mathbf{H}}_{\mathrm{BC}})\mathbf{D}, \tag{A.8}$$

so that

$$\begin{aligned}
\boldsymbol{\epsilon}_{x}^{\mathrm{a}} &= \boldsymbol{\epsilon}_{x}^{\mathrm{b}} + \mathbf{K}_{x,\mathrm{BC}}\Big(\hat{\boldsymbol{\epsilon}}_{\mathrm{BC}} - \hat{\boldsymbol{\eta}}_{\mathrm{BC}} - \hat{\mathbf{H}}_{\mathrm{BC}}\boldsymbol{\epsilon}_{x}^{\mathrm{b}} - \hat{\mathbf{C}}_{\beta}\boldsymbol{\epsilon}_{\beta}^{\mathrm{b}}\Big) + (\mathbf{I}_n - \mathbf{K}_{x,\mathrm{BC}}\hat{\mathbf{H}}_{\mathrm{BC}})\mathbf{D}\Big(\hat{\boldsymbol{\epsilon}}_{\downarrow} - \hat{\boldsymbol{\eta}}_{\downarrow} - \hat{\mathbf{H}}_{\downarrow}\boldsymbol{\epsilon}_{x}^{\mathrm{b}}\Big) \\
&= \mathbf{K}_{x,\mathrm{BC}}\Big(\hat{\boldsymbol{\epsilon}}_{\mathrm{BC}} - \hat{\boldsymbol{\eta}}_{\mathrm{BC}} - \hat{\mathbf{C}}_{\beta}\boldsymbol{\epsilon}_{\beta}^{\mathrm{b}}\Big) + (\mathbf{I}_n - \mathbf{K}_{x,\mathrm{BC}}\hat{\mathbf{H}}_{\mathrm{BC}})\mathbf{D}\Big(\hat{\boldsymbol{\epsilon}}_{\downarrow} - \hat{\boldsymbol{\eta}}_{\downarrow}\Big) + \Big(\mathbf{I}_n - \mathbf{K}_{x,\mathrm{BC}}\hat{\mathbf{H}}_{\mathrm{BC}} - (\mathbf{I}_n - \mathbf{K}_{x,\mathrm{BC}}\hat{\mathbf{H}}_{\mathrm{BC}})\mathbf{D}\hat{\mathbf{H}}_{\downarrow}\Big)\boldsymbol{\epsilon}_{x}^{\mathrm{b}} \\
&= \mathbf{K}_{x,\mathrm{BC}}\Big(\hat{\boldsymbol{\epsilon}}_{\mathrm{BC}} - \hat{\boldsymbol{\eta}}_{\mathrm{BC}} - \hat{\mathbf{C}}_{\beta}\boldsymbol{\epsilon}_{\beta}^{\mathrm{b}}\Big) + (\mathbf{I}_n - \mathbf{K}_{x,\mathrm{BC}}\hat{\mathbf{H}}_{\mathrm{BC}})\mathbf{D}\Big(\hat{\boldsymbol{\epsilon}}_{\downarrow} - \hat{\boldsymbol{\eta}}_{\downarrow}\Big) + (\mathbf{I} - \mathbf{K}_{x,\mathrm{BC}}\hat{\mathbf{H}}_{\mathrm{BC}})(\mathbf{I}_n - \mathbf{D}\hat{\mathbf{H}}_{\downarrow})\boldsymbol{\epsilon}_{x}^{\mathrm{b}} \\
&= \mathbf{K}_{x,\mathrm{BC}}\Big(\hat{\boldsymbol{\epsilon}}_{\mathrm{BC}} - \hat{\boldsymbol{\eta}}_{\mathrm{BC}} - \hat{\mathbf{C}}_{\beta}\boldsymbol{\epsilon}_{\beta}^{\mathrm{b}}\Big) + (\mathbf{I}_n - \mathbf{K}_{x,\mathrm{BC}}\hat{\mathbf{H}}_{\mathrm{BC}})\Big(\mathbf{D}\Big(\hat{\boldsymbol{\epsilon}}_{\downarrow} - \hat{\boldsymbol{\eta}}_{\downarrow}\Big) + (\mathbf{I}_n - \mathbf{D}\hat{\mathbf{H}}_{\downarrow})\boldsymbol{\epsilon}_{x}^{\mathrm{b}}\Big).
\end{aligned} \tag{A.9}$$

The expected error in the analysis of the state can be derived assuming that VarBC is applied correctly in the previous cycle. As explained in Section 2, this is analogous to assuming $\langle\boldsymbol{\epsilon}_{\beta}^{\mathrm{b}}\rangle = 0$ and $\langle\hat{\boldsymbol{\epsilon}}_{\mathrm{BC}}\rangle = 0$. We also have, by definition, $\langle\hat{\boldsymbol{\epsilon}}_{\downarrow}\rangle = 0$. Taking the expectation of Equation (A.9) and substituting in these assumptions then gives

$$\langle\boldsymbol{\epsilon}_{x}^{\mathrm{a}}\rangle = (\mathbf{I}_n - \mathbf{K}_{x,\mathrm{BC}}\hat{\mathbf{H}}_{\mathrm{BC}})\Big((\mathbf{I}_n - \mathbf{D}\hat{\mathbf{H}}_{\downarrow})\langle\boldsymbol{\epsilon}_{x}^{\mathrm{b}}\rangle - \mathbf{D}\langle\hat{\boldsymbol{\eta}}_{\downarrow}\rangle\Big) - \mathbf{K}_{x,\mathrm{BC}}\langle\hat{\boldsymbol{\eta}}_{\mathrm{BC}}\rangle. \tag{A.10}$$

Equation (A.10) gives the bias in the state at the beginning of the assimilation window, the time at which the analysis is defined. This can be propogated forward to derive the bias in the state at the end of the window, T_{end}:

$$\langle\boldsymbol{\epsilon}_{x}^{\mathrm{a}}\rangle(T_{\mathrm{end}}) = m_{0\to T_{\mathrm{end}}}(\langle\boldsymbol{\epsilon}_{x}^{\mathrm{a}}\rangle(T_0)) + \sum_{k=1}^{T_{\mathrm{end}}} m_{0\to k}(\boldsymbol{\eta}_k). \tag{A.11}$$

A.3 Derivation of model error "seen" by the observations

The model error "seen" by the observations is defined as the difference between the 4D observation operator used in the assimilation and the true observation operator acting on the true initial control vector. Assuming that the observation operator mapping the state at the same time as the observations to observation space is perfect, then the error in the 4D observation operator is due to model error only. In Equation (5) we define the model error as additive at each timestep, therefore if we have an observation at time i in the assimilation window, the model error "seen" by that observation will be

$$\begin{aligned}
\hat{\boldsymbol{\eta}}_i &= \hat{h}_i(\mathbf{x}_0^{\dagger}) - \hat{h}_i^{\dagger}(\mathbf{x}_0^{\dagger}) \\
&= h_i(m_{0\to i}(\mathbf{x}_0^{\dagger})) - h_i(m_{0\to i}^{\dagger}(\mathbf{x}_0^{\dagger})) \\
&= h_i\Big(m_{i\to i-1}^{\dagger}\Big[m_{i-1\to i-2}^{\dagger}\Big(\ \cdots\ \Big[m_{2\to 1}^{\dagger}\Big(m_{1\to 0}^{\dagger}(\mathbf{x}_0^{\dagger}) + \boldsymbol{\eta}_1\Big) + \boldsymbol{\eta}_2\Big]\ \cdots\ \Big) + \boldsymbol{\eta}_{i-1}\Big] + \boldsymbol{\eta}_i\Big) \\
&\quad - h_i\Big(m_{i\to i-1}^{\dagger}\Big[m_{i-1\to i-2}^{\dagger}\Big(\ \cdots\ \Big[m_{2\to 1}^{\dagger} \times \Big(m_{1\to 0}^{\dagger}(\mathbf{x}_0^{\dagger})\Big)\Big]\ldots\Big)\Big]\Big).
\end{aligned}$$

Using the first-order Taylor expansion of h_i and m_i about $\mathbf{x}_i^{\dagger}$ and $\mathbf{x}_{i-1}^{\dagger}$, respectively, we can simplify this to become

$$\begin{aligned}
\hat{\boldsymbol{\eta}}_i &= \mathbf{H}_i\Big(\mathbf{M}_{i-1}\Big[\mathbf{M}_{i-2}\Big(\ldots\Big[\mathbf{M}_1\Big(\boldsymbol{\eta}_1\Big) + \boldsymbol{\eta}_2\Big]\ldots\Big) + \boldsymbol{\eta}_{i-1}\Big] + \boldsymbol{\eta}_i\Big) \\
&= \mathbf{H}_i\sum_{k=1}^{i}\mathbf{M}_{k\to i}\boldsymbol{\eta}_k,
\end{aligned} \tag{A.12}$$

where $\mathbf{M}_{k\to i} = \mathbf{M}_{i-1}\mathbf{M}_{i-2}\ldots\mathbf{M}_k$ is the tangent linear (TL) model propagating a perturbation to the state at time k to time i, and $\mathbf{M}_{i\to i}$ is the identity. Note that, because the model error is assumed to be additive, the TLs of the true and assimilation models are the same.

If we have multiple observations throughout the assimilation window, then we can define $P = \sum_{t=1}^{T_y p_t}$ as the total number of observations such that $\hat{\boldsymbol{\eta}} \in \mathbb{R}^P$ and $\hat{\boldsymbol{\eta}}_i \in \mathbb{R}^{p_i}$:

$$\hat{\boldsymbol{\eta}} = \begin{pmatrix} \hat{\boldsymbol{\eta}}_{t_1} \\ \vdots \\ \hat{\boldsymbol{\eta}}_{t_i} \\ \vdots \\ \hat{\boldsymbol{\eta}}_{t_{T_y}} \end{pmatrix}. \tag{A.13}$$

As we have assumed that the bias correction is perfect if the true bias-correction coefficients are known, then the model error "seen" by both the BC and anchor observations will have the same form as given in Equation (A.12).

APPENDIX B. THE SENSITIVITY OF THE ANALYSIS OF THE BIAS-CORRECTION COEFFICIENTS TO THE BIAS-CORRECTED OBSERVATIONS, $\mathbf{K}_{\beta,\mathrm{BC}}$

The sensitivity of the analysis of the bias-correction coefficients to the bias-corrected observations, $\mathbf{K}_{\beta,\mathrm{BC}}$, was derived in Equation (8). Within this Appendix, we show how $\mathbf{K}_{\beta,\mathrm{BC}}$ relates to the analysis-error covariance matrix in Section B.1. We also explore its limits for a scalar example in Section B.2.

B.1 Relationship between the sensitivity of the analysis to the observations and the analysis-error covariance

The Kalman gain matrix is related to the analysis-error covariance for an optimal system via the following equation: $\mathbf{A}_\mathrm{v} = (\mathbf{I}_{n+r} - \mathbf{K}_\mathrm{v}\widehat{\mathbf{H}}_\mathrm{v})\mathbf{B}_\mathrm{v}$ (Kalnay, 2003).

$$\begin{aligned}\mathbf{A}_\mathrm{v} &= (\mathbf{I}_{n+r} - \mathbf{K}_\mathrm{v}\widehat{\mathbf{H}}_\mathrm{v})\mathbf{B}_\mathrm{v}\\ &= \left(\begin{pmatrix}\mathbf{I}_n & \mathbf{O}\\ \mathbf{O} & \mathbf{I}_r\end{pmatrix} - \begin{pmatrix}\mathbf{K}_{x,\mathrm{BC}} & \mathbf{K}_{x,\mathrm{anc}}\\ \mathbf{K}_{\beta,\mathrm{BC}} & \mathbf{K}_{\beta,\mathrm{anc}}\end{pmatrix}\begin{pmatrix}\widehat{\mathbf{H}}_\mathrm{BC} & \widehat{\mathbf{C}}_\beta\\ \widehat{\mathbf{H}}_\mathrm{anc} & \mathbf{O}\end{pmatrix}\right)\begin{pmatrix}\mathbf{B}_\mathrm{x} & \mathbf{O}\\ \mathbf{O} & \mathbf{B}_\beta\end{pmatrix}\\ &= \begin{pmatrix}\left(\mathbf{I} - \mathbf{K}_{x,\mathrm{BC}}\widehat{\mathbf{H}}_\mathrm{BC} - \mathbf{K}_{x,\mathrm{anc}}\widehat{\mathbf{H}}_\mathrm{anc}\right)\mathbf{B}_\mathrm{x} & -\mathbf{K}_{x,\mathrm{BC}}\widehat{\mathbf{C}}_\beta\mathbf{B}_\beta\\ \left(-\mathbf{K}_{\beta,\mathrm{BC}}\widehat{\mathbf{H}}_\mathrm{BC} - \mathbf{K}_{\beta,\mathrm{anc}}\widehat{\mathbf{H}}_\mathrm{anc}\right)\mathbf{B}_\mathrm{x} & \left(\mathbf{I}_r - \mathbf{K}_{\beta,\mathrm{BC}}\widehat{\mathbf{C}}_\beta\right)\mathbf{B}_\beta\end{pmatrix}.\end{aligned}$$

Therefore, the analysis-error covariance for the state is

$$\mathbf{A}_\mathrm{x} = \left(\mathbf{I} - \mathbf{K}_{x,\mathrm{BC}}\widehat{\mathbf{H}}_\mathrm{BC} - \mathbf{K}_{x,\mathrm{anc}}\widehat{\mathbf{H}}_\mathrm{anc}\right)\mathbf{B}_\mathrm{x}.$$

Using Equation (A.8), we can rewrite this as

$$\mathbf{A}_\mathrm{x} = \left(\mathbf{I} - \mathbf{K}_{x,\mathrm{BC}}\widehat{\mathbf{H}}_\mathrm{BC}\right)\left(\mathbf{I} - \mathbf{D}\widehat{\mathbf{H}}_\mathrm{anc}\right)\mathbf{B}_\mathrm{x}.$$

Similarly, the analysis-error covariance for the bias-correction coefficient is

$$\mathbf{A}_\beta = \left(\mathbf{I}_r - \mathbf{K}_{\beta,\mathrm{BC}}\widehat{\mathbf{C}}_\beta\right)\mathbf{B}_\beta. \tag{B.1}$$

Here we have assumed that the model error is deterministic. Alternatively, if the model error as seen by the observations, defined by Equation (A.13), is random with $\langle(\widehat{\boldsymbol{\eta}} - \langle\widehat{\boldsymbol{\eta}}\rangle)(\widehat{\boldsymbol{\eta}} - \langle\widehat{\boldsymbol{\eta}}\rangle)^\mathrm{T}\rangle = \widehat{\mathbf{Q}} \in \mathbb{R}^{P\times P}$, then this will inflate the analysis-error covariance matrix. It can be shown that a correction term of the form $\mathbf{K}_\mathrm{v}\widehat{\mathbf{Q}}\mathbf{K}_\mathrm{v}^\mathrm{T}$ should be added to $\mathbf{A}_\mathrm{v}$:

$$\begin{aligned}&\mathbf{K}_\mathrm{v}\widehat{\mathbf{Q}}\mathbf{K}_\mathrm{v}^\mathrm{T}\\ &= \begin{pmatrix}\mathbf{K}_{x,\mathrm{BC}} & \mathbf{K}_{x,\mathrm{anc}}\\ \mathbf{K}_{\beta,\mathrm{BC}} & \mathbf{K}_{\beta,\mathrm{anc}}\end{pmatrix}\begin{pmatrix}\widehat{\mathbf{Q}}_\mathrm{BC} & \widehat{\mathbf{Q}}_\mathrm{BC,anc}\\ \widehat{\mathbf{Q}}_\mathrm{BC,anc}^\mathrm{T} & \widehat{\mathbf{Q}}_\mathrm{anc}\end{pmatrix}\begin{pmatrix}\mathbf{K}_{x,\mathrm{BC}} & \mathbf{K}_{x,\mathrm{anc}}\\ \mathbf{K}_{\beta,\mathrm{BC}} & \mathbf{K}_{\beta,\mathrm{anc}}\end{pmatrix}^\mathrm{T}.\end{aligned} \tag{B.2}$$

B.2 Scalar case with anchor and bias-corrected observations at different times

Let us consider the scalar case when the anchor and bias-corrected observations are at different times. Let $\widehat{\mathbf{H}}_\mathrm{BC}\mathbf{B}_x\widehat{\mathbf{H}}_\mathrm{BC}^\mathrm{T} = B_x^\mathrm{BC}$, $\widehat{\mathbf{H}}_\downarrow\mathbf{B}_x\widehat{\mathbf{H}}_\downarrow^\mathrm{T} = B_x^\mathrm{anc}$, $\widehat{\mathbf{H}}_\mathrm{BC}\mathbf{B}_x\widehat{\mathbf{H}}_\downarrow^\mathrm{T} = \widehat{\mathbf{H}}_\downarrow\mathbf{B}_x\widehat{\mathbf{H}}_\mathrm{BC}^\mathrm{T} = B_x^\mathrm{BC,anc}$, and $\widehat{\mathbf{C}}_\beta = 1$. From Equation (8) we have

$$\begin{aligned}K_{\beta,\mathrm{BC}} &= B_\beta\left[\begin{pmatrix}B_x^\mathrm{BC} + B_\beta + \widehat{R}_\mathrm{BC} & B_x^\mathrm{BC,anc}\\ B_x^\mathrm{BC,anc} & B_x^\mathrm{anc} + \widehat{R}_\downarrow\end{pmatrix}^{-1}\right]_\mathrm{BC,BC}\\ &= \frac{B_\beta(B_x^\mathrm{anc} + \widehat{R}_\downarrow)}{(B_x^\mathrm{BC} + B_\beta + \widehat{R}_\mathrm{BC})(B_x^\mathrm{anc} + \widehat{R}_\downarrow) - (B_x^\mathrm{BC,anc})^2}.\end{aligned}$$

As $B_x^\mathrm{BC,anc}$ increases (i.e., the background-error correlations between the variables observed by the anchor and BC observations increases), $K_{\beta,\mathrm{BC}}$ increases.

If the model is growing, then the background-error covariances evolved to the time of the observations will be greater for whichever observations are later. For example, let $\widehat{\mathbf{H}}_\mathrm{BC} = \alpha\widehat{\mathbf{H}}_\downarrow$, implying $B_x^\mathrm{BC} = \alpha^2 B_x^\mathrm{anc}$ and $B_x^\mathrm{BC,anc} = \alpha B_x^\mathrm{anc}$. Substitute this into the above expression:

$$\begin{aligned}K_{\beta,\mathrm{BC}} &= \frac{B_\beta(B_x^\mathrm{anc} + \widehat{R}_\downarrow)}{(\alpha^2 B_x^\mathrm{anc} + B_\beta + \widehat{R}_\mathrm{BC})(B_x^\mathrm{anc} + \widehat{R}_\downarrow) - (\alpha B_x^\mathrm{anc})^2}\\ &= \frac{B_\beta(B_x^\mathrm{anc} + \widehat{R}_\downarrow)}{\alpha^2 B_x^\mathrm{anc}\widehat{R}_\downarrow + (B_\beta + \widehat{R}_\mathrm{BC})(B_x^\mathrm{anc} + \widehat{R}_\downarrow)}.\end{aligned}$$

If $\alpha > 1$, that is, the BC are later than the anchor observations, then $K_{\beta,\mathrm{BC}}$ will be smaller than if $\alpha < 1$, that is, the BC observations are earlier than the anchor observations. From Equation (B.1), we can therefore conclude that the analysis of the bias coefficients will be more precise when BC observations are earlier in the window and anchor observations are later in the window.

In the limit as $\widehat{R}_\downarrow \to 0$,

$$K_{\beta,\mathrm{BC}} \to \frac{B_\beta}{B_\beta + \widehat{R}_\mathrm{BC}}.$$

This is an upper limit. From Equation (B.1), this implies a lower limit for the analysis-error variance of A_β of

$$(1 - K_{\beta,\mathrm{BC}})B_\beta = \frac{B_\beta\widehat{R}_\mathrm{BC}}{B_\beta + \widehat{R}_\mathrm{BC}}.$$