SPECIAL
MATERHORN
COLLECTION

Observation and Model Bias Estimation in the Presence of Either or Both Sources of Error

RAQUEL LORENTE-PLAZAS

Research Application Laboratory, National Center for Atmospheric Research, Boulder, Colorado, and Department of Civil and Environmental Engineering and Earth Science, University of Notre Dame, Notre Dame, Indiana

JOSHUA P. HACKER

Research Application Laboratory, National Center for Atmospheric Research, Boulder, Colorado

ABSTRACT

In numerical weather prediction and in reanalysis, robust approaches for observation bias correction are necessary to approach optimal data assimilation. The success of bias correction can be limited by model errors. Here, simultaneous estimation of observation and model biases, and the model state for an analysis, is explored with ensemble data assimilation and a simple model. The approach is based on parameter estimation using an augmented state in an ensemble adjustment Kalman filter. The observation biases are modeled with a linear term added to the forward operator. A bias is introduced in the forcing term of the model, leading to a model with complex errors that can be used in imperfect-model assimilation experiments.

Under a range of model forcing biases and observation biases, accurate observation bias estimation and correction are possible when the model forcing bias is simultaneously estimated and corrected. In the presence of both model error and observation biases, estimating one and ignoring the other harms the assimilation more than not estimating any errors at all, because the biases are not correctly attributed. Neglecting a large model forcing bias while estimating observation biases results in filter divergence; the observation bias parameter absorbs the model forcing bias, and recursively and incorrectly increases the increments. Neglecting observation bias results in suboptimal assimilation, but the model forcing bias parameter estimate remains stable because the model dynamics ensure covariance between the parameter and the model state.

1. Introduction

In statistics, the term *bias* is broadly used when errors are systematic instead of random (i.e., when the mean of the error distribution is not zero). Data assimilation (DA) algorithms in wide use today rely on the basic assumptions of unbiased observations and models. In those systems, observations with assumed random errors are used to correct the random errors in a model-forecast background estimate. The underlying theories allow for known biases to be corrected prior to assimilation, thereby yielding an unbiased assimilation. But spatially and temporally varying contributions to bias are difficult to quantify in complex geophysical models and observing networks. The result is that forecasts

inevitably have biases that cannot be perfectly corrected. Abundant observations with nonzero and possibly unknown mean errors are part of the regular observing network. Some examples of biased observations can be found in aircraft data (Tenenbaum 1996), satellite radiances (Eyre 1992), radiosonde observations (Wang et al. 2002), and systematic representativeness errors in surface observations (Bédard et al. 2015). In prediction models, the initial conditions, the parameterization of subgrid-scale physical processes, and other deficiencies can cause systematic errors at various scales. The assimilation can be far from optimal when biases in either the observations, the models, or both are present (e.g., Dee 2005; Eyre 2016). Here a simple method, based on state augmentation in ensemble filter data assimilation, is explored for simultaneously estimating and correcting observation biases and a bias in model forcing. The emphasis is on understanding the effectiveness

Corresponding author: Raquel Lorente-Plazas, lorente.plazas@gmail.com

of observation bias estimation in the presence of varying levels of model error.

It is impossible to a priori determine whether biases in the state estimates result from biased observations or model deficiencies, because both can cause systematic departures of the predicted state from observations. An optimal data assimilation system is attainable if both the source, and the structure, of the errors are known and accounted for in the system. In most cases empirical evidence combined with intuition lead us to conclude whether biases result from a model or the observations. In some cases observation biases can be of the same order of magnitude as biases in the short-term forecasts that provide background fields for data assimilation.

Dee (2005) suggested using *bias-blind* assimilation if the source of bias is uncertain, because incorrect attribution can cause the assimilation adapt to an unknown bias. Dee (2004) speculated that background departures can increase using observation bias correction due to the presence of systematic model errors. Although all atmospheric models are biased, and biased observations are common, most past work has focused on addressing one source of bias (e.g., Baek et al. 2006; Auligné et al. 2007). Few studies have addressed both model and observation biases together (Pauwels et al. 2013; Eyre 2016), and investigated the ability to optimize the data assimilation while considering both sources of error.

A wide variety of algorithms can adaptively estimate bias as part of the assimilation. The general approach is to include parameters that represent the biases in the assimilation system, and augment the state vector (or control vector) with the parameters (Friedland 1969). The parameters can represent biases in a model and/or observations, and can be estimated using variational methods or filters. One common approach is based on a two-step design, where the state is first estimated and then biases are estimated in a following step (Dee and Da Silva 1998; Dee and Todling 2000; Fertig et al. 2009). An optimal estimate for the state variables and parameters can also be simultaneously estimated through Jacobian minimization (Dee 2005; Auligné et al. 2007), or Kalman filter equations (Aksoy et al. 2006).

Most effort in estimating observation bias has been focused on satellite radiances. Since Derber and Wu (1998) implemented a variational bias correction to satellite radiances at NCEP, a variety of work has emerged (Auligné et al. 2007; Dee and Uppala 2009; Fertig et al. 2009). The bias is modeled by a linear series with several flow-dependent and instrument-dependent predictors, and coefficients are estimated in the second step of the two-step process. Efforts to correct and estimate in situ

surface observation bias are scarce. Recently, Bédard et al. (2015) introduced a geostatistical observation operator that corrects for systematic and representativeness errors for near-surface winds.

The work presented here builds on past work by exploring the interplay between a model forcing bias and observation biases, and estimation of both, in ensemble data assimilation. Experiments with and without observation and model forcing bias, and with and without bias estimation, are performed with a range of bias magnitudes meant to elucidate the effectiveness of the assimilation under various conditions. In this work spatially correlated biases that may result from atmospheric state correlations are ignored, and the bias is assumed to be uncorrelated in space.

The organization of this paper is as follows. Section 2 describes the approach developed to estimate and correct model forcing bias and location-dependent observation biases. Section 3 describes the Lorenz (2005, hereafter L05) model and the experimental design. Results are presented in section 4. The main conclusions are summarized in section 5.

2. Bias aware data assimilation in the ensemble filter

The approach for estimating and correcting observation and model bias is to use an ensemble Kalman filter to estimate the state augmented with parameters describing the biases. The augmented state vector leads to simultaneous estimation of state variables and parameters.

The observation biases addressed here apply to spatially distributed observations with fixed locations. Following a philosophy similar to bias estimation for satellite radiances (e.g., Derber and Wu 1998; Auligné et al. 2007; Fertig et al. 2009), the bias of each observing location is assumed to be given by a linear series of predictors with coefficients. Here for simplicity only the leading term is retained. The model for biased observations ($\mathbf{y}$) is

$$\mathbf{y} = h(\mathbf{x}^t) + \boldsymbol{\beta}_O + \epsilon_O, \tag{1}$$

so that the observation error includes both a bias $\boldsymbol{\beta}_O$ and Gaussian random error ϵ_O with zero mean and standard deviation σ_O. The vectors in Eq. (1) have a dimension corresponding to the number of observations suspected to suffer from systematic error, which may be all observations. The forward operator $h(\mathbf{x}^t)$ is a linear interpolation from the nearest model grid points. The advantage of modeling the errors with Eq. (1) is its

ability to correct the observation bias with known $\boldsymbol{\beta}_O$ regardless of whether it results from a systematic instrument error, representativeness error, or both (Fertig et al. 2009).

A forcing bias in the model, β_M, is also added to the right-hand side of the model equations (described below in section 3a). As seen later, a nonzero β_M changes both the mean and variance of the model state, indicating a complex model error response.

Observation and model biases are estimated by including $\widehat{\boldsymbol{\beta}_O}$ and $\widehat{\boldsymbol{\beta}_M}$ parameters in the augmented state vector $\mathbf{z} = [\mathbf{x}, \widehat{\boldsymbol{\beta}_O}, \widehat{\boldsymbol{\beta}_M}]^{\mathrm{T}}$, where $\mathbf{x}$ is the usual model state vector and the hat indicates estimates of the true parameters. Then, the atmospheric state and the parameters are simultaneously updated by solving the analysis equations with the ensemble Kalman filter applied to the augmented state. The analysis equations are given by

$$\mathbf{z}^a = \mathbf{z}^f + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{z}^f), \tag{2}$$

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^{\mathrm{T}} (\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}, \tag{3}$$

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^f, \tag{4}$$

where the superscripts a and f represent the analysis and forecast, respectively. The model state is related to the observation through the (linear) forward operator $\mathbf{H}$, and $\mathbf{R}$ is the error covariance matrix for the observations, which is diagonal because the observation biases are assumed to be uncorrelated. Here $\mathbf{P}$ are the covariance matrices for the augmented state given by (for either forecast f or analysis a):

$$\mathbf{P}^{f,a} = \begin{bmatrix} \mathbf{P}^{f,a}_{xx} & \mathbf{P}^{f,a}_{x\beta} \\ \mathbf{P}^{f,a}_{x\beta} & \mathbf{P}^{f,a}_{\beta\beta} \end{bmatrix}. \tag{5}$$

Matrix $\mathbf{P}^{f,a} = \mathbf{Z}^{f,a}(\mathbf{Z}^{f,a})^{\mathrm{T}}$ is computed from matrix of ensemble perturbations for the forecast or analysis augmented state $\mathbf{Z}^{f,a}$. Blocks $\mathbf{P}^{f,a}_{x\beta}$ are the covariances linking the model state variables to the parameters.

3. Model and experiments

a. The L05 model

Experiments are carried out using Model III developed by L05, because it has two characteristics that are useful for our investigations. First, Model III has large-scale correlations between neighboring grid points. Second, the model combines small and large scales analogous to mesoscales and synoptic scales. The superposition of two scales, and realistic spatial correlations, provide a useful platform for experimentation with observations that sample multiple scales of motion.

In Model III, the small scale (short waves, X_n), and large scale (long waves, Y_n) are superposed in the variables Z_n. The model equation is

$$\frac{dZ_n}{dt} = [X,X]_{K,n} + b^2[Y,Y]_{1,n} + c[Y,X]_{1,n} - X_n - bY_n + F, \tag{6}$$

where c is a coupling coefficient, and b reduces the amplitude and increases the frequency of the small scales. The variable F represents the forcing term and does not modify the long wavelength, but increases the number of shorter waves imposed per long wave. The subscript n indexes the N model grid points. Here K is inversely proportional to the wavelength of the long waves. The advection terms are defined as a sum of pairs of products in order to introduce the spatial continuity (the reader is referred to L05 for further details about the model).

All model constants are selected based on previous work (e.g., L05; Lei and Hacker 2015). Values follow: $N = 960$ (corresponding to $0.375°$ grid spacing), $K = 32$, $b = 10$, and $c = 2.5$. Here $F = 15$ in the perfect model. Equation (6) is integrated in time using a fourth-order Runge–Kutta method with a nondimensional time step of 0.001, equivalent to 432 s based on error growth arguments that link one dimensionless time unit to 5 days L05.

Adding β_M to the right-hand side of Eq. (6) is equivalent to adding a bias to the forcing F. Both perfect-model and imperfect-model experiments help to demonstrate the effects of model error, and estimates of the forcing bias, on estimates of observation biases.

b. Data assimilation strategies

The bias estimation is implemented in the Data Assimilation Research Testbed (DART; Anderson et al. 2009), and the assimilation experiments use the serial implementation of the ensemble adjustment Kalman filter (EAKF; Anderson 2001, 2003), but the approach can apply to any ensemble filter algorithm. The EAKF is a deterministic filter, which does not rely on perturbed observations. Observations are sampled from a truth simulation at a discrete time (here every 50 time steps) by applying the forward operators. They are assimilated one at a time by using each in serial to update the joint state vector defined as $[\mathbf{x}^f, \mathbf{H}\mathbf{x}^f]$ (i.e., the state and all of the prior observation values; the forward operators applied to the model state). Parameters appended to that joint state vector are simultaneously updated as described in section 2. Multiplicative covariance inflation and covariance localization are applied as described below. As is typical in ensemble filter implementations,

no formal initialization procedure such as a digital filter or normal mode initialization is applied to the analysis.

To select the ensemble size, the forecast or background (prior) error was analyzed for different ensemble sizes and numbers of parameters (not shown). As expected, errors increase with the number of parameters estimated, and decrease with the ensemble size. In this work, unless stated otherwise, 100 ensemble members are used to provide stable estimates.

Multiplying elements in the prior covariance matrix by a factor greater than 1.0 (covariance inflation) helps maintain enough spread in the ensemble and ensure sufficient overlap between the prior and observation likelihood in the filter (e.g., Anderson and Anderson 1999). A Bayesian, adaptive, and spatially varying state space inflation for the prior distribution, described in Anderson (2009), is applied here. Distributions of the inflation value for each state variable are updated according to the error statistics each time observations are assimilated. A 1.1 mean inflation factor with a standard deviation of 0.6 is applied at initial time. Typically, inflation values increase when observations are dense, to counteract the tendency toward very small analysis uncertainty on the nearby grid points. In the absence of observations, the inflation factor is damped by a factor of 0.9 each assimilation cycle so that the variance can decay. The damping mitigates the potential for large inflation values, and therefore large background uncertainty that can lead to large analysis increments where observations are lacking. The adaptive inflation is applied equally to the model state variables and the parameters.

The assimilation will tend to reduce the variance in the parameter distributions. A lack of a prognostic equation with error growth for the parameters means that the variance of the parameter distribution cannot grow as the state advances in time. Without measures beyond the adaptive inflation, parameter variance will tend to zero. To avoid this, a minimum variance in the parameter distributions is enforced after each assimilation. The variance persists to be the prior variance at the next assimilation. A minimum variance of 0.2 and 0.5 for observation bias and model forcing bias parameters, respectively, is imposed. The selection of these values will be discussed in section 4e.

Covariance localization is applied to mitigate sampling error, and a reduced subset of observations within a specified region can impact a particular state variable. Localization is carried out using the fifth-order piecewise rational function developed by Gaspari and Cohn (1999), which is a function of only distance. The half-width of the localization function is manually chosen based on tuning experiments to be 0.3 rad in the domain—approximately 17.20° longitude or 46 grid points.

Each observation bias parameter is collocated with an observing station at a single location. They are assumed to be independent of each other. The effect of any spatial correlations that may exist between a bias parameter and observations located elsewhere will be examined in future research.

c. Evaluation metrics

Defining an error as the ensemble-mean forecast (background or prior) minus the truth at each assimilation time (i.e., $\mathbf{z}^f - \mathbf{z}^t$), standard metrics are used to evaluate the system. The root-mean-square error (RMSE), as well as its decomposed bias and error standard deviation (STD) components, assesses the quality of the assimilation experiments. The errors are quantified for the background/prior to evaluate how the assimilation is affected by the biases and bias estimation. RMSE of the parameter estimates (cf. the true observation bias or model forcing bias) are also useful to assess whether the biases are properly attributed to the model or the observations, and are denoted RMSE β_O and RMSE β_M. A spinup period of 100 assimilation cycles is not included in the statistics.

d. Experimental design

Five sets of experiments are carried out to assess the performance of the observation bias correction and its interaction with model bias. First, the performance of a suboptimal assimilation with unknown biases in the observations and/or in the model is assessed to provide a baseline. Second, the accuracy of the observation bias correction approach is assessed with and without a biased model. Third, the consequences of incorrectly attributing the source of bias to the model (β_M) or the observations ($\boldsymbol{\beta}_O$) is examined. Fourth, the sensitivity of observation bias correction to the magnitude of model forcing bias is evaluated. Finally, the optimal values of the minimum specified parameter variances for $\widehat{\boldsymbol{\beta}_O}$ and $\widehat{\beta_M}$ are analyzed to understand trade-offs between the two. Table 1 summarizes the experiments.

Experiments are constructed to assess effects of both magnitude and spatial variability of the observation bias on the assimilation. First, a homogeneous (spatially constant) bias of 0.3 or 1 is imposed. Later, more realistic heterogeneous, but temporally constant, biases are added according to $\mathcal{N}(0, 0.09)$ or $\mathcal{N}(0, 0.25)$. Because it simplifies the interpretation, most of the sensitivity experiments use a homogeneous 0.3 bias.

The bias estimation and correction are evaluated using both perfect- and imperfect-model scenarios. In the perfect model, the forcing term in Eq. (6) is $F = 15$ ($\beta_M = 0$). Imposing $\beta_M \neq 0$, equivalent to $F \neq 15$, results in imperfect-model experiments. In this study, $\beta_M = 2$ is chosen, because it produces a state RMSE

TABLE 1. Summary of the experiments performed with different sources of biases and parameters estimated. The sources of bias can be attributed to model forcing (β_M) and/or observations ($\boldsymbol{\beta}_O$). Biases are not estimated in bias-blind assimilation or estimated in bias-aware assimilation using the state augmentation.

Model	Assigned bias	Augmented state	Assimilation
Perfect ($\beta_M = 0$)	$\boldsymbol{\beta}_O = 0$	$\mathbf{z} = [\mathbf{x}]$	Control
	$\boldsymbol{\beta}_O \neq 0$	$\mathbf{z} = [\mathbf{x}]$	β_O blind
		$\mathbf{z} = [\mathbf{x}, \widehat{\boldsymbol{\beta}_O}]$	β_O aware
Imperfect ($\beta_M \neq 0$)	$\boldsymbol{\beta}_O = 0$	$\mathbf{z} = [\mathbf{x}]$	β_M blind
	$\boldsymbol{\beta}_O \neq 0$	$\mathbf{z} = [\mathbf{x}]$	$\beta_{O,M}$ blind
		$\mathbf{z} = [\mathbf{x}, \widehat{\boldsymbol{\beta}_O}]$	β_O aware/β_M blind
		$\mathbf{z} = [\mathbf{x}, \widehat{\boldsymbol{\beta}_M}]$	β_O blind/β_M aware
		$\mathbf{z} = [\mathbf{x}, \widehat{\boldsymbol{\beta}_O}, \widehat{\boldsymbol{\beta}_M}]$	$\beta_{O,M}$ aware

similar to the results from imposing a homogeneous $\beta_O = 0.3$ (shown below). Note that vector $\boldsymbol{\beta}_O$ has dimension equal to the number of biased observations (here 240), and β_M is a scalar.

To generate an ensemble of initial conditions, a single initial condition is created and then perturbed by adding a small Gaussian-random perturbation to every element of the state vector. The perturbed states are integrated for 1000 days to generate a climatological state distribution, from which the ensemble initial conditions are sampled.

The observation network consists of 240 fixed random locations in the domain, which observe the true state every 50 time steps (6 h). Synthetic observations are taken by applying Eq. (1) to the true state (the state of a model with $\beta_M = 0$), where the random error is sampled from the same Gaussian distribution used for assimilation, and the (true) observation bias $\boldsymbol{\beta}_O$ is specified. The observation error variance is $\sigma_O^2 = 0.5$.

4. Results

a. Blind-bias assimilation

Context for the parameter estimation experiments is best established by first illustrating the assimilation suboptimality in the presence of unknown observation or model biases. Table 2 compares RMSE, error STD, and bias for several experiments that assimilate biased observations in a perfect model, and unbiased observations in a biased model. The results indicate that either a model forcing bias or observation bias can lead to similar errors in prediction.

A perfect model assimilating unbiased observations results in negligible background bias, and an error STD of 0.292. A spatially constant observation bias of 1 leads to a background bias of 0.757. The positive bias indicates a forecast systematically greater than the truth, consistent with a systematically positive error in the observations. Constant observation bias also introduces an apparently random error component, with an error STD nearly double the value resulting from perfect observations. When the mean state of the model changes from assimilating biased observations, the nonlinear terms in the model can contribute to the error STD. A smaller homogeneous observation bias of 0.3 reduces both error components.

The spatially varying observation bias contributes directly to the background error STD, and shifts some of the background error from the bias term to the error STD term. The mean value of the observation bias distribution when it is spatially varying is zero, which results in a background bias much smaller than when the

TABLE 2. RMSE, error STD, and bias computed of the prior ensemble mean against the truth, and prior ensemble spread (standard deviation). Error measurements are computed for perfect model with biased observations and imperfect model with unbiased observations removing spinup.

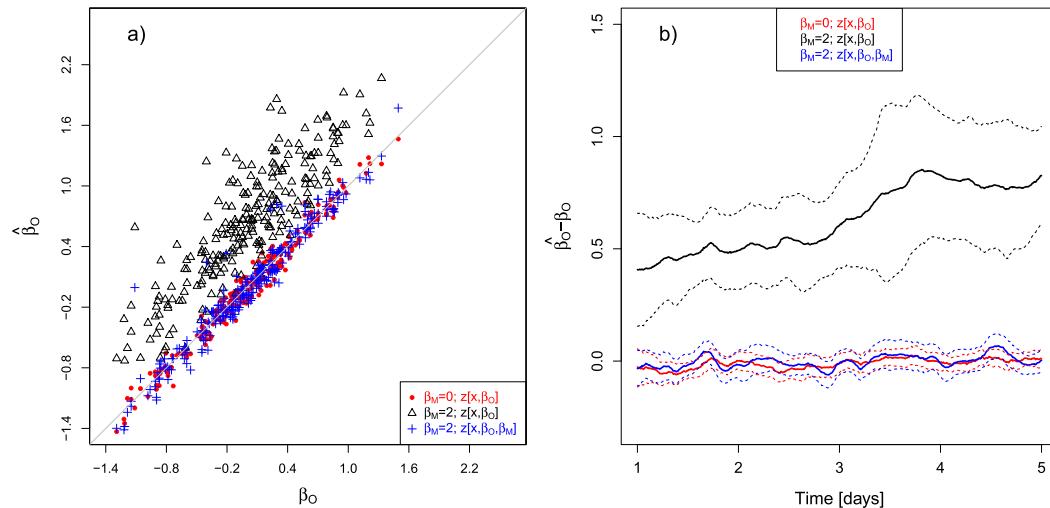|	Assigned bias	RMSE	STD	Bias
Perfect model $\beta_M = 0$	—	0.292	0.292	0.001
	$\beta_O = 1$	0.960	0.590	0.757
	$\beta_O = 0.3$	0.415	0.363	0.202
	$\beta_O \sim \mathcal{N}(0, 0.25)$	0.547	0.542	-0.077
	$\beta_O \sim \mathcal{N}(0, 0.09)$	0.360	0.359	-0.018
Imperfect model $\boldsymbol{\beta}_o = 0$	$\beta_M = 2$ ($F = 13$)	0.421	0.396	-0.141
	$\beta_M = 1$ ($F = 14$)	0.332	0.325	-0.070
	$\beta_M = -1$ ($F = 16$)	0.325	0.318	0.066
	$\beta_M = -2$ ($F = 17$)	0.385	0.364	0.126

FIG. 1. (a) Estimated observation bias as a function of specified bias and (b) time series of estimated observation bias minus assigned bias. In (b) the mean (solid lines) and standard deviation (dashed lines) of the 240 observing locations are shown. Colors distinguish different experiments: a perfect model with $\widehat{\boldsymbol{\beta}_O}$ estimated (red), an imperfect model with $\widehat{\boldsymbol{\beta}_O}$ estimated (black), and an imperfect model with both $\widehat{\boldsymbol{\beta}_M}$ and $\widehat{\boldsymbol{\beta}_O}$ estimated (blue).

observation bias is spatially invariant. Observation biases with a magnitude less than 0.3 (the observation bias standard deviation) occur 68% of the time, but the resulting error STD is approximately the same as when the observation bias is constant in space at 0.3 (0.359 and 0.363).

A biased model forcing also leads to more than just a biased model background forecast. It changes the mean state of the model, and contributes to random error growth through the nonlinear model equations. The asymmetric background error magnitudes around $\beta_M = 0$ also suggests a nonlinear response to the forcing bias.

Different dynamical response to observation and model forcing biases, as measured by error components, make it difficult to specify observation and forcing bias magnitudes that lead to the same error components. Spatially invariant observation bias of 0.3 and a model forcing bias with $F = 13$ have more similar errors for any other pair of experiments in Table 2. Although the background bias is greater for the biased observation experiment than for the biased forcing experiment, they are of the same order of magnitude. The error STDs are nearly identical.

b. Estimation of spatially varying observation biases

Experiments with spatially varying observation bias are presented first. They are the most realistic and demonstrate that the state augmentation can successfully estimate observation biases. The effectiveness of the observation bias estimates under three different scenarios is presented: a perfect model, an imperfect model, and an imperfect model with simultaneous

estimation of the forcing bias. Here the observations have spatially heterogeneous biases with a normal distribution as $\mathcal{N}(0, 0.25)$.

Figure 1a compares the estimated observation bias to the true (assigned) bias. Each symbol represents the bias at a single observing location, averaged over the experiment period (removing the spinup). In the perfect-model scenario (red circles), the estimates lie close to the 1:1 line, indicating a near-perfect estimation. When an imperfect model is used ($F = 13$, black triangles), the observation bias estimates are biased themselves; the model bias is assigned to the parameters meant to estimate observation bias. The estimated observation biases are close to the 1:1 line [blue plus symbols $(+)$] when the model forcing bias is estimated. The observation biases are better estimated, and apparently attributed correctly, if the assimilation is aware of the model forcing bias as opposed to blind to it.

Time series of the spatially varying observation bias estimates also show that the state augmentation can recover the assigned biases as long as a parameter to estimate the model forcing bias is included (Fig. 1b). Within sampling error for the 240 observing locations, the true minus estimated observation bias is zero, as long as model forcing bias does not exist or is estimated and corrected. The temporal mean $\widehat{\beta_M}$ is 2.142, with a 0.654 temporal standard deviation. The values of $\widehat{\beta_M}$ are positive, compensating for the negative forcing bias generated by subtracting 2 from the forcing term in the perfect model, so that $F = 13$.

Judgement about whether the relative magnitudes of observation and model bias in this experiment are the
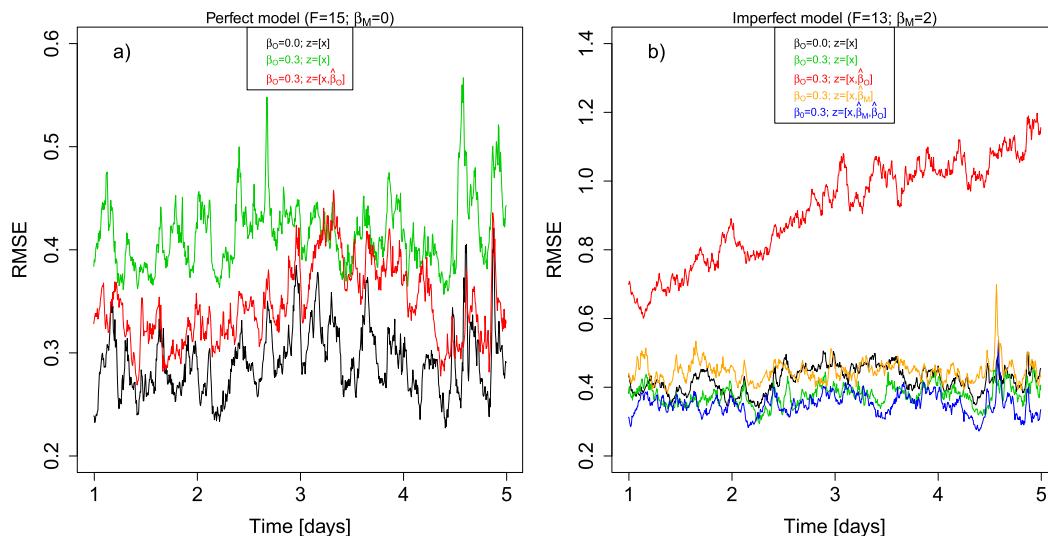
FIG. 2. Time series of prior RMSE for different experiments (colors) using (a) a perfect model with and without spatially constant observation bias of $\boldsymbol{\beta}_O = 0.3$, and (b) an imperfect model with $\beta_M = 2$. Legends show specified observation biases and the augmented vector estimated in the assimilation.

same as those for a real model, and the real atmosphere, is not possible at this point. The results here are in agreement with the analytical results from Eyre (2016), who pointed out that observation bias correction is relatively straightforward in the absence of model bias, but that observation bias correction is more complicated in the presence of model bias.

c. Attributing observation and model bias

In this section several experiments with various bias parameters (observation, model, and none) are constructed to explore whether the biases are correctly attributed to the observations or the model. The focus is on spatially invariant observation bias because it facilitates interpretation, and because of the similar background error magnitude that results when $F = 13$. The experiments are summarized in Table 1. The effect of the bias correction is evaluated by analyzing the time evolution of the RMSE (Fig. 2). Accuracy of the estimated biases is evaluated with RMSE β_O and RMSE β_M (Fig. 3).

Three experiments use a perfect model (Fig. 2a): (i) unbiased observations, (ii) biased observations without estimation (bias blind), and (iii) parameter estimation for $\boldsymbol{\beta}_O$ with a spatially invariant bias of 0.3 (i.e., bias aware). The state RMSE when the observations are biased (green curve) is approximately 0.4, compared to about 0.3 when the observations are unbiased (black curve). Observation bias parameter estimation (red curve) reduces the RMSE compared to the bias-blind assimilation, but the error is not quite reduced to the

error level resulting from unbiased observations. The observation bias estimation improves the assimilation and ensuing prediction, but it does not completely eliminate the effects of the biased observations.

Five experiments are performed with an imperfect model ($F = 13$, Fig. 2b). Experiments with and without observation biases, and no parameter estimation, are the first two (green and black curves). Three experiments with parameter estimation are also presented. One estimates $\boldsymbol{\beta}_O$ (red), the second estimates β_M (gold), and the third estimates both (blue). The specified observation bias is again 0.3 for all observations when they are biased. Note the change in scale from Fig. 2a. Assimilating unbiased observations into the imperfect model raises the RMSE to greater than 0.4, compared with the perfect-model RMSE of ~ 0.3, as expected (more precise RMSE values are in Table 2). Biased observations slightly reduce the overall RMSE to below 0.4, because the sign of the observation bias is opposite the effect of the model forcing bias (green curve in Fig. 2b).

Simultaneously estimating and applying the parameters representing both forcing and observation biases (blue curve in Fig. 2b) results in the lowest errors measured in the presence of biases from either source. Estimating the observation bias $\boldsymbol{\beta}_O$, while ignoring the model forcing bias, results in the state diverging from the truth (i.e., filter divergence, red curve). The observation bias parameter is adapting to both the observation bias and model forcing bias distributions. Consequently the observation bias estimates are too large, and the resulting analysis increment is too large. Recursively this
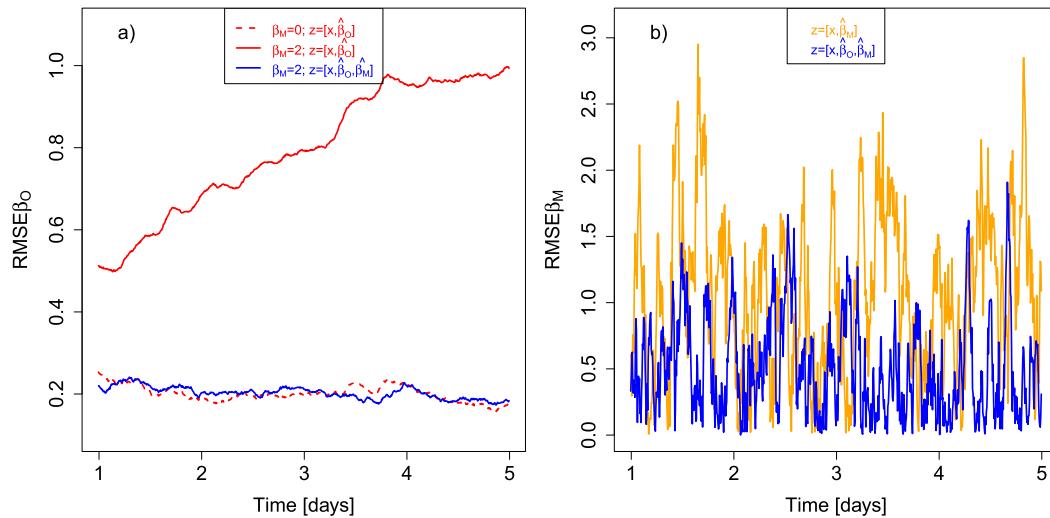
Fig. 3. Time series of (a) RMSE β_O and (b) RMSE β_M for different parameter estimation experiments. The legend in (a) shows the model forcing bias and the augmented vector estimated in the assimilation. In (b) the model forcing bias is $\beta_M = 2$.

produces continually degrading analyses and predictions for both observation bias parameter and model state.

Estimating β_M, while ignoring the observation bias (gold curve), also degrades the prior RMSE compared to not estimating the bias at all (green curve). Here the model bias parameter is incorrectly accounting for some of the observation bias, but the filter does not diverge. This is addressed further in later sections.

To determine whether the prior RMSE is improved for the right reason, bias attribution can be assessed by comparing the estimated parameters with the assigned (true) biases in the model or the observations (Fig. 3). The error in the observation bias parameter estimate is approximately the same whether the model forcing is unbiased (red dashed curve), or the forcing is biased but the bias is estimated and corrected (blue curve). Error in estimating the observation biases grows through the length of the experiment when the assimilation is blind to the model forcing bias (solid red curve in Fig. 3a), again reflecting filter divergence.

The RMSE β_M (blue curve in Fig. 3b) is about 20% of the overall error in the L05 forcing parameter ($\beta_M = 2$). The estimates are variable in time because, as discussed above, a biased F introduces error STD in addition to bias. It is clear that a reasonable estimate of β_M is important for good estimates of the observation biases, and results in appropriate bias attribution and skillful state estimates. The β_M estimate is also less skillful when the assimilation is blind to observation bias (gold curve), consistent with the less accurate state estimates for the same experiment.

Results presented so far are consistent with Dee (2005), who suggested that incorrectly attributing the source of the bias could harm the assimilation, and recommended bias-blind assimilation when the source and characterization of the biases are unknown. Consistent with that recommendation, the experiments estimating observation biases, but blind to model forcing bias, result in filter divergence because the biases are incorrectly attributed. Experiments here that estimate both observation and model forcing biases avoid filter divergence, and lead to smaller state errors. In this case, it is not necessary to know the full characterization of the model error a priori. It appears to be sufficient to know that a model error exists, and have a parameter that represents at least some part of that error. It is true that in complex geophysical models, the form of a parametric model to represent those errors is not always clear, but additive biases certainly exist.

The next sections further address why β_M-blind with β_O-aware assimilation fails, but β_M-aware and β_O-blind assimilation does not. Why $\widehat{\beta_M}$ variability is greater than $\widehat{\boldsymbol{\beta}_O}$ is also further explored.

d. Sensitivity to model forcing bias magnitude

Results presented above show that the success in estimating observation bias can depend on whether the forcing bias parameter is simultaneously estimated. It follows that the accuracy of the state estimates, when observation biases are present, may depend on the magnitude of the forcing bias. To quantify how the forcing bias affects the estimates, β_M is varied from -3 to 3. The ability to estimate and correct observation bias
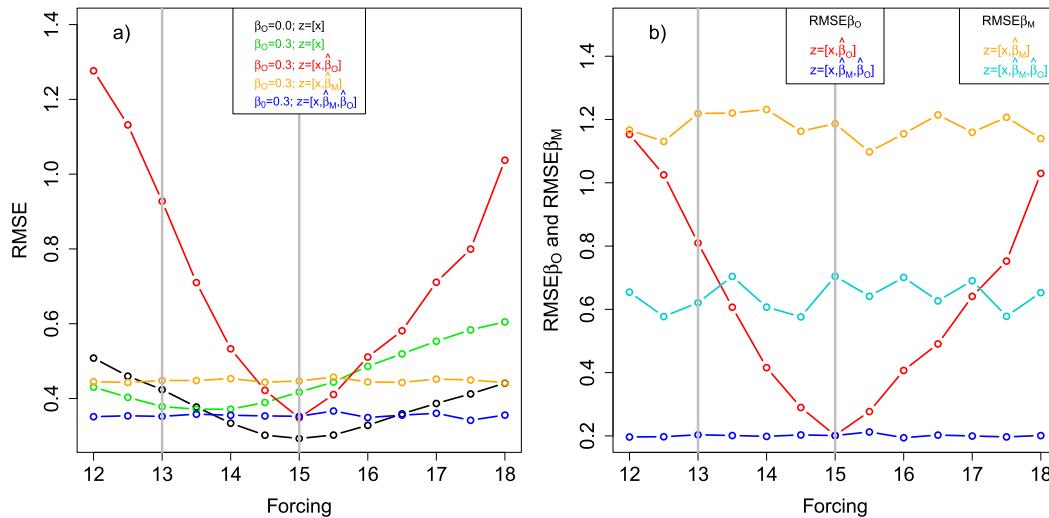
FIG. 4. (a) Prior RMSE, and (b) both RMSE β_O and RMSE β_M vs the forcing term F in the assimilating model. The legend in (a) shows the specified observation bias and the augmented vector estimated in the assimilation. In (b) the observation bias is $\boldsymbol{\beta}_O = 0.3$. The vertical gray lines show, for reference, the forcing values for the perfect and imperfect models used in the estimation experiments reported in prior figures.

is evaluated with experiments that are both aware of, and blind to, the model forcing bias.

For β_M-blind assimilation with no observation bias, the state RMSE increases monotonically away from the perfect-model RMSE (black curve in Fig. 4a). When the observations are biased (0.3 for all observations), the minimum RMSE shifts from $F = 15$ to $F = 14$ (green curve). In this case, it is because of the experiment design; the negative model bias compensates for the positive observation bias, and provides a better estimate of the truth.

When observation bias is estimated, and the assimilation is β_M blind, the RMSE increases steeply as β_M increases (red curve). The minimum error results when the model is perfect. The vertical gray line at $F = 13$ corresponds to the results shown in Figs. 2b and 3a, where the red curve shows a failure due to filter divergence. The steep RMSE increase with β_M in the red curve shows that the assimilation is especially sensitive when both model forcing and observation biases are present, but the forcing bias is not estimated.

Estimating β_M leads to state estimates that are insensitive to the error itself, as long as it is estimated, within the range examined here. Ignoring observation biases (gold curve) leads to greater RMSE compared to when observation biases are estimated simultaneously to the forcing bias (blue curve), as expected.

Estimating both β_M and β_O sets a minimum RMSE for experiments that have both biases, regardless of the magnitude of the forcing (blue curve). By allowing the assimilation to account for the model forcing bias,

the RMSE remains at that level rather than increasing with greater β_M magnitude. The state error cannot quite match the perfect-model and perfect-observation results ($F = 15$ on the black curve), consistent with Fig. 2a, but it is clear that simultaneous estimation is possible for a range of model forcing biases.

Similar behavior characterizes the estimated parameters (Fig. 4b). RMSE β_O increases with the forcing in β_M-blind assimilation (red curve), but it is independent of forcing bias in β_M-aware assimilation (blue curve). The estimate of β_M is insensitive to the β_M magnitude itself (gold curve) as long as it is estimated. RMSE β_M is lower for β_O-aware assimilation (cyan curve) than for β_O-blind assimilation.

To better understand what influences the state and parameter RMSE, Fig. 5a shows the sensitivity of the spatial and temporal mean prior state to the magnitude of the forcing bias. The prior mean state is not sensitive to forcing changes as long as β_M is estimated (gold and blue curves), which agrees with the lack of RMSE sensitivity to forcing changes (Fig. 4a). Because the mean state is proportional to the bias minus the true mean state, Fig. 5a also shows that the state error STD is also not sensitive to the forcing bias as long as it is estimated. When β_M is not estimated (black, green, and red curves), the mean state monotonically increases with value of the forcing bias, where greater values of forcing bias are equivalent to greater values of F.

A change in the mean state with the forcing leads to a shift in the background probability, reducing the overlap between the prior distribution and the observation
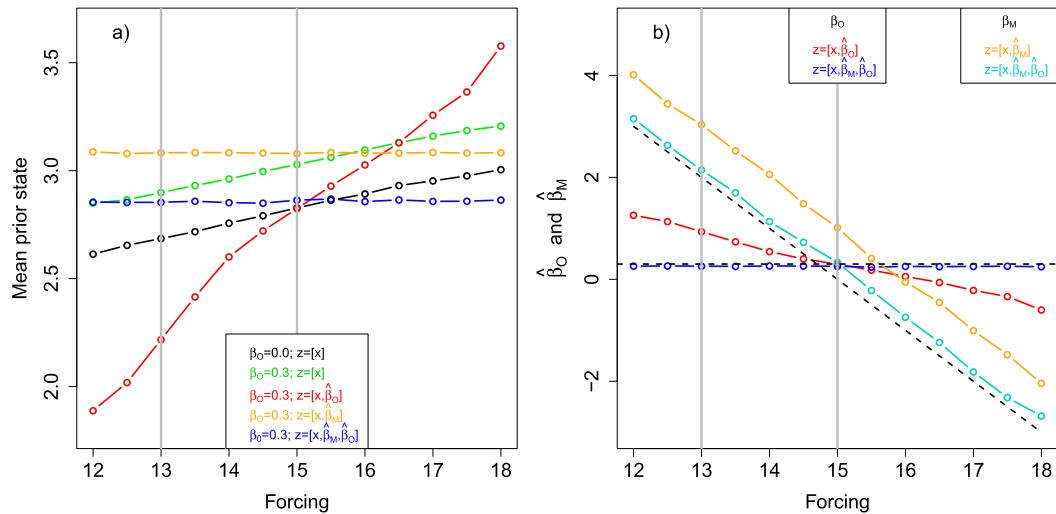
FIG. 5. (a) The prior mean state and (b) the estimated parameter values vs the forcing term F in the assimilating model. The legend in (a) shows the specified observation bias and the augmented vector estimated in the assimilation. In (b) the specified biases are shown in dashed black lines. The vertical gray lines show, for reference, the forcing values for the perfect and imperfect models used in the estimation experiments reported in prior figures.

likelihood. The prior ensemble must be inflated more to account for the model error (not shown).

Figure 5b shows the parameter estimates for different forcing biases in the assimilating model. The $\boldsymbol{\beta}_O$ is estimated reasonably well when $\boldsymbol{\beta}_M$ is also estimated (blue), but $\widehat{\boldsymbol{\beta}_O}$ worsens when $\boldsymbol{\beta}_M$ is not estimated. Likewise, the $\widehat{\boldsymbol{\beta}_M}$ is closer to the real value when $\boldsymbol{\beta}_O$ is estimated (cyan), but the $\widehat{\boldsymbol{\beta}_M}$ estimate is less accurate when the assimilation is blind to observation bias (gold). This demonstrates the fact that in the presence of two sources of bias (model and observations), one ignored bias will be incorrectly attributed to the other estimated parameter. The result is an overestimation of the bias parameter that is updated in the assimilation, harming the assimilation.

These results help explain the filter divergence observed in Fig. 2b. By construction the observation bias appears only in the forward observation operator, and is not dynamically correlated with the state. Ignoring the forcing bias in the estimation, when the bias exists in the model, leads to the observation bias parameter absorbing the state-dependent error component. The parameters then become correlated with the state. That unphysical correlation produces a feedback that eventually results in filter divergence.

e. Sensitivity to minimum parameter variance

The assimilation acts to reduce variance in the parameter distributions, and the lack of a prognostic equation for the parameter (besides persistence) means that the parameter variance has no way to grow during the period between assimilations. Although adaptive inflation is applied to the parameters, a minimum variance is enforced to ensure that the parameter retains sufficient spread to be updated by the observations. Minimum values for parameter error variances for the assimilation are most easily chosen through direct experimentation. This section quantifies the effects of the choice of minimum variance on the accuracy of the parameter estimates.

A useful minimum parameter variance can be selected by looking at how RMSE, RMSE β_O, and RMSE β_M vary with bias magnitude and the minimum imposed parameter variance (Fig. 6). From the perfect-model experiments (Fig. 6a), the RMSE is minimum at $\sigma^2_{\beta_O} \approx 0.2$, except for the experiment with a spatially invariant observation bias of 0.3. In that case, slightly lower RMSE results with $\sigma^2_{\beta_O} \approx 0.1$. RMSE β_O (Fig. 6b) shows analogous behavior. For an imperfect model, the prior RMSE is fairly constant with $\sigma^2_{\beta_M} > 0.2$ (Fig. 6c) and the RMSE β_M generally increases with the variance at $\sigma^2_{\beta_M} > 0.1$ (Fig. 6d).

The experiments reported in previous sections imposed a minimum $\sigma^2_{\beta_O} = 0.2$, corresponding to perfect-model minimums in state RMSE and parameter RMSE β_O. A more conservative $\sigma^2_{\beta_M} = 0.5$ was imposed because of the weaker state RMSE dependence on it.

Temporal evolution of state error and estimates of the model forcing bias parameter show how the variability of the estimates changes with the minimum variances (Fig. 7). In agreement with the RMSE results in Fig. 6, state errors are more sensitive to $\sigma^2_{\beta_O}$ than to $\sigma^2_{\beta_M}$. When
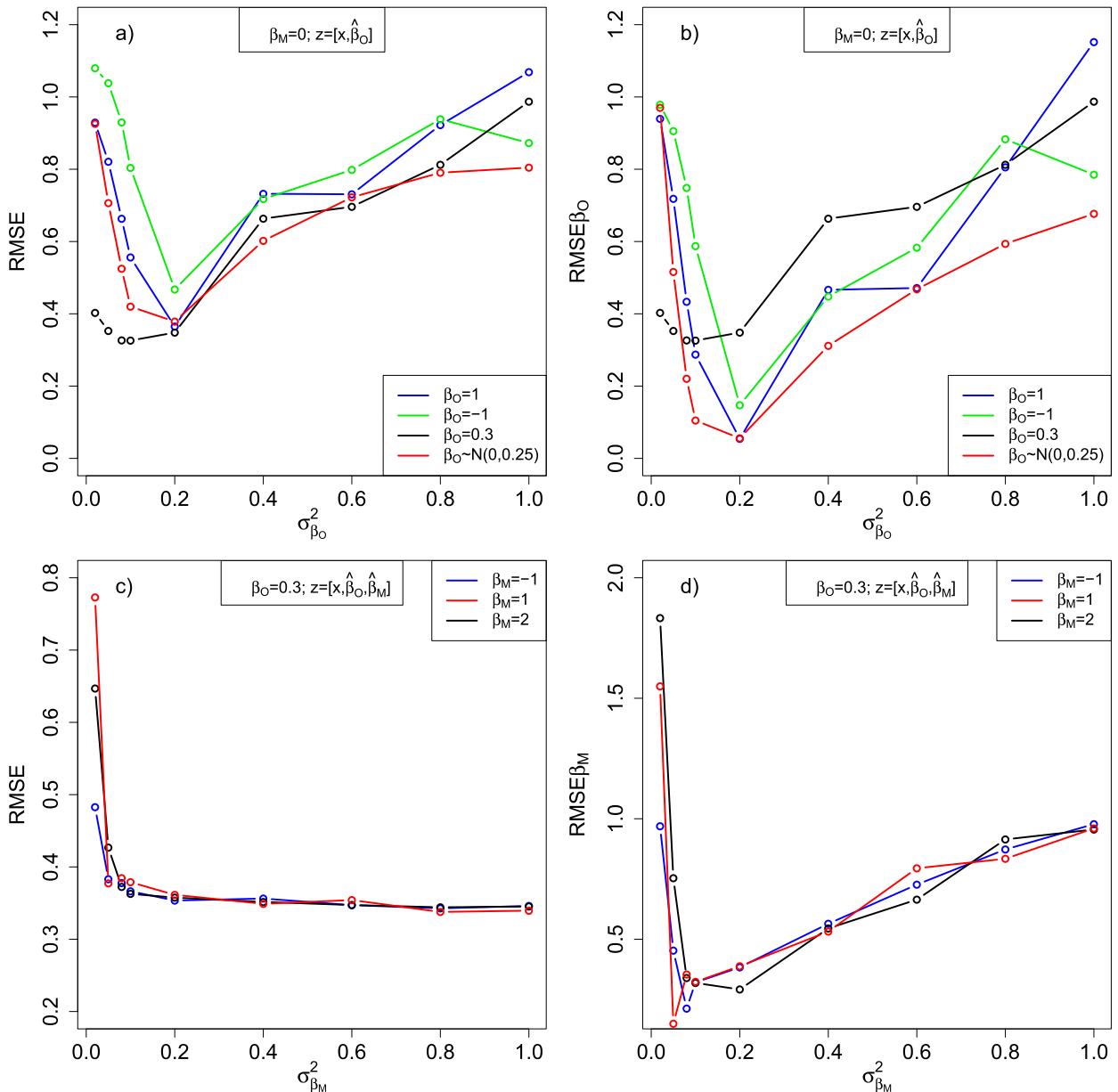
FIG. 6. Perfect-model results for (a) RMSE and (b) RMSE β_O as a function of $\sigma^2_{\beta_O}$. Colors indicate specified observation bias ($\boldsymbol{\beta}_O$), with $\boldsymbol{\beta}_O$ estimated in the assimilation. Imperfect-model results for (c) RMSE and (d) RMSE β_M as function of $\sigma^2_{\beta_M}$. Colors indicate specified forcing bias (β_M), with both $\boldsymbol{\beta}_O$ and β_M estimated in the assimilation.

$\sigma^2_{\beta_O}$ is less than 0.2, the distribution is not sufficiently updated and bias estimates are slow to converge (red curve in Fig. 7a). The prior RMSE is greater because the observation bias estimate is inaccurate.

The temporal state error variability is less sensitive to the choice of minimum $\sigma^2_{\beta_M}$ (Fig. 7b). The prior RMSE in Fig. 6c is insensitive to $\sigma^2_{\beta_M}$ because the state error STD is insensitive to it. The model forcing bias parameter changes the model dynamics, and the estimate of β_M responds. The true model error is a function of the

state, while the true observation bias is temporally constant and independent of the state.

These results help to further explain why a reasonable model forcing bias estimate can be obtained when the assimilation is blind to the observation bias, but accurate observation bias estimates are not possible under large model forcing bias that is ignored in the assimilation. The $\widehat{\beta_M}$ errors are sensitive to the parameter variance (Figs. 6d and 7c), but prior state errors are relatively insensitive to parameter variance as long as values are
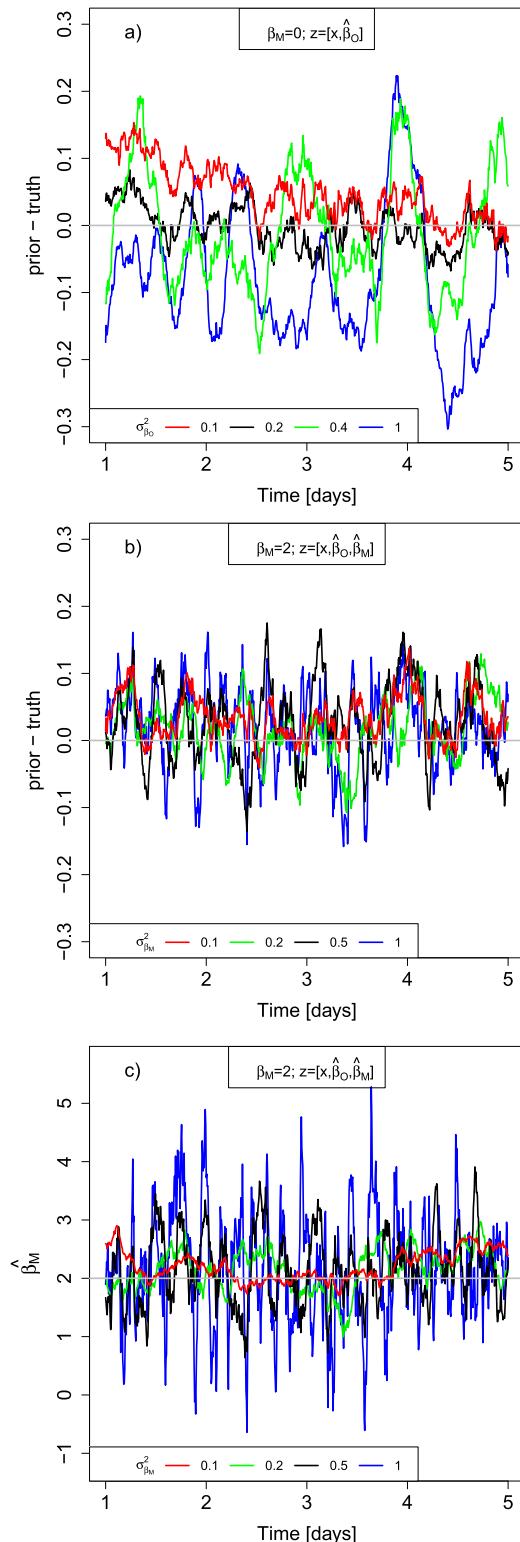
FIG. 7. (a) Perfect-model experiment error time series (prior minus truth) for different minimum enforced $\sigma^2_{\beta_O}$ (colors) when estimating a assigned observation bias $\boldsymbol{\beta}_O = 0.3$. (b) Imperfect-model experiment time series for state errors and (c) $\widehat{\beta_M}$ for different minimum enforced $\sigma^2_{\beta_M}$ (colors) and estimating both model and observation bias as shown in the upper legend.

greater than about 0.1 (Figs. 6c and 7b). In contrast, the state estimates degrade with poorer observation bias estimates (Fig. 7a). Presence of a model forcing bias parameter prevents the model error from being absorbed by the observation bias estimate, and allows an accurate estimate of the observation bias and consequently the model state.

5. Conclusions

This paper explores the interactions between model and observation biases, both when they are ignored and when they are estimated simultaneously to the state in data assimilation. Parameters representing observation biases are included as terms in the forward operators, and a parameter representing model forcing bias is added as an extra term in the model equations. Observation and model forcing biases are estimated and corrected in the assimilation by including them in an augmented state. The L05 Model III included in the DART software provides the basis for quantitative testing in a variety of perfect- and imperfect-model experiments.

The augmented state approach is able to estimate both spatially varying and constant observations biases using a perfect model. The assimilation suffers when an imperfect model is used, and the model forcing bias is ignored, while estimating biases in observations. State RMSE increases proportionally to the model forcing bias magnitude, and the filter diverges under sufficiently large forcing bias. This is a consequence of incorrectly attributing the bias source. When the model forcing bias is estimated and corrected with an additional parameter, the observation biases can be accurately estimated. Accurate parameter estimation improves the assimilation and subsequent predictions, as measured against the true state.

The quantitative effect from ignoring one of either model forcing or observation biases, when both are present, depends on which one is ignored. Experiments that estimate model forcing bias and ignore observation bias lead to lower errors than experiments that estimate observation bias and ignore model forcing bias. This is true even when the state errors resulting from the forcing or observation biases are of similar magnitude.

In this work, the model error appears in the model forcing term. Model forcing perturbations, which are the parameters, dynamically covary with the model state. But the observation bias appears only in the forward observation operator, and is not dynamically correlated with the state. When the forcing bias parameter is not estimated, but the model has an incorrect forcing value, the observation error parameters absorb the state-dependent

error component and become correlated with the state. The result is that the bias estimates and analysis increments feed back to each other, and the filter can eventually diverge if the model forcing bias is large enough.

A minimum value for parameter variance was enforced in these experiments, to ensure the parameter estimates retain some uncertainty. No prognostic equation is available to promote parameter spread growth, and the ensemble covariance inflation may not produce sufficient spread. Accuracy of the parameter estimates depend on the minimum variance specified, as shown with sensitivity experiments. Results also show that the state estimate is relatively insensitive to the accuracy of the model forcing bias estimate as long as a reasonable estimate is available. This allows for an accurate observation bias estimate, and an accurate state estimate when both model and observation errors are present.

Although the results presented here are specific to the model and experimental framework, care was taken to avoid unrealistic results. The potential benefits of the bias estimation algorithm explored here motivate its application to higher-dimensional models. Experimentation with the Weather Research and Forecasting (WRF) Model (Skamarock et al. 2008) is in progress. Different from the experiments here, correlations between the observations and observation bias may exist. One of the key challenges for application in a complex model such as the WRF is that the structure of the model error is unknown, and the relative magnitudes of the model and observation biases are also unknown.

REFERENCES

Aksoy, A., F. Zhang, and J. W. Nielsen-Gammon, 2006: Ensemble-based simultaneous state and parameter estimation in a two-dimensional sea-breeze model. *Mon. Wea. Rev.*, **134**, 2951–2970, doi:10.1175/MWR3224.1.

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903, doi:10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

——, 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634–642, doi:10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2.

——, 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus*, **61A**, 72–83, doi:10.1111/j.1600-0870.2008.00361.x.

——, and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758, doi:10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2.

——, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A community facility. *Bull. Amer. Meteor. Soc.*, **90**, 1283–1296, doi:10.1175/2009BAMS2618.1.

Auligné, T., A. McNally, and D. Dee, 2007: Adaptive bias correction for satellite data in a numerical weather prediction system. *Quart. J. Roy. Meteor. Soc.*, **133**, 631–642, doi:10.1002/qj.56.

Baek, S.-J., B. R. Hunt, E. Kalnay, E. Ott, and I. Szunyogh, 2006: Local ensemble Kalman filtering in the presence of model bias. *Tellus*, **58A**, 293–306, doi:10.1111/j.1600-0870.2006.00178.x.

Bédard, J., S. Laroche, and P. Gauthier, 2015: A geo-statistical observation operator for the assimilation of near-surface wind data. *Quart. J. Roy. Meteor. Soc.*, **141**, 2857–2868, doi:10.1002/qj.2569.

Dee, D. P., 2004: Variational bias correction of radiance data in the ECMWF system. *Proc. ECMWF Workshop on Assimilation of High Spectral Resolution Sounders in NWP*, Vol. 28, Reading, United Kingdom, ECMWF, 97–112. [Available online at https://www.ecmwf.int/sites/default/files/elibrary/2004/8930-variational-bias-correction-radiance-data-ecmwf-system.pdf.]

——, 2005: Bias and data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131**, 3323–3344, doi:10.1256/qj.05.137.

——, and A. M. Da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.*, **124**, 269–296, doi:10.1002/qj.49712454512.

——, and R. Todling, 2000: Data assimilation in the presence of forecast bias: The GEOS moisture analysis. *Mon. Wea. Rev.*, **128**, 3268–3282, doi:10.1175/1520-0493(2000)128<3268:DAITPO>2.0.CO;2.

——, and S. Uppala, 2009: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quart. J. Roy. Meteor. Soc.*, **135**, 1830–1841, doi:10.1002/qj.493.

Derber, J. C., and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287–2299, doi: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. Tech. Memo. 186, European Centre for Medium-Range Weather Forecasts, 34 pp. [Available online at https://www.ecmwf.int/sites/default/files/elibrary/1992/9330-bias-correction-scheme-simulated-tovs-brightness-temperatures.pdf.]

——, 2016: Observation bias correction schemes in data assimilation systems: A theoretical study of some of their properties. *Quart. J. Roy. Meteor. Soc.*, **142**, 2284–2291, doi:10.1002/qj.2819.

Fertig, E. J., and Coauthors, 2009: Observation bias correction with an ensemble Kalman filter. *Tellus*, **61A**, 210–226, doi:10.1111/j.1600-0870.2008.00378.x.

Friedland, B., 1969: Treatment of bias in recursive filtering. *IEEE Trans. Auto. Control*, **14**, 359–367.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, doi:10.1002/qj.49712555417.

Lei, L., and J. P. Hacker, 2015: Nudging, ensemble, and nudging ensembles for data assimilation in the presence of model error. *Mon. Wea. Rev.*, **143**, 2600–2610, doi:10.1175/MWR-D-14-00295.1.

Lorenz, E. N., 2005: Designing chaotic models. *J. Atmos. Sci.*, **62**, 1574–1587, doi:10.1175/JAS3430.1.

Pauwels, V., G. De Lannoy, H.-J. Hendricks Franssen, and H. Vereecken, 2013: Simultaneous estimation of model state variables and observation and forecast biases using a two-stage hybrid Kalman filter. *Hydrol. Earth Syst. Sci.*, **17**, 3499–3521, doi:10.5194/hess-17-3499-2013.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., doi:10.5065/D68S4MVH.

Tenenbaum, J., 1996: Jet stream winds: Comparisons of aircraft observations with analyses. *Wea. Forecasting*, **11**, 188–197, doi:10.1175/1520-0434(1996)011<0188:JSWCOA>2.0.CO;2.

Wang, J., H. L. Cole, D. J. Carlson, E. R. Miller, K. Beierle, A. Paukkunen, and T. K. Laine, 2002: Corrections of humidity measurement errors from the Vaisala RS80 radiosonde—Application to TOGA COARE data. *J. Atmos. Oceanic Technol.*, **19**, 981–1002, doi:10.1175/1520-0426(2002)019<0981:COHMEF>2.0.CO;2.