Tellus (2006), 58A, 293–306
Printed in Singapore. All rights reserved

TELLUS

Local ensemble Kalman filtering in the presence of model bias

By SEUNG-JONG BAEK[1]*, BRIAN R. HUNT[2], EUGENIA KALNAY[3], EDWARD OTT[1,4] and ISTVAN SZUNYOGH[3], [1]*Institute for Research in Electronics and Applied Physics, and Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA;* [2]*Institute for Physical Science and Technology, and Department of Mathematics, University of Maryland, College Park, MD, USA;* [3]*Institute for Physical Science and Technology, and Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, USA;* [4]*Department of Physics, University of Maryland, College Park, MD, USA*

(Manuscript received 6 April 2005; in final form 5 December 2005)

ABSTRACT

We modify the local ensemble Kalman filter (LEKF) to incorporate the effect of forecast model bias. The method is based on augmentation of the atmospheric state by estimates of the model bias, and we consider different ways of modeling (i.e. parameterizing) the model bias. We evaluate the effectiveness of the proposed augmented state ensemble Kalman filter through numerical experiments incorporating various model biases into the model of Lorenz and Emanuel. Our results highlight the critical role played by the selection of a good parameterization model for representing the form of the possible bias in the forecast model. In particular, we find that forecasts can be greatly improved *provided* that a good model parameterizing the model bias is used to augment the state in the Kalman filter.

1. Introduction

In many situations the dynamics of a real process may differ from those of the best available model of that process. We refer to this difference as *model error*. Model error is thought to be a key issue in weather forecasting in that the presence of model error can lead to cause large discrepancies between the forecasts and the true atmospheric states. In this connection, we note (i) that Kalman filters have been considered for estimating atmospheric states to be used as initial conditions in forecast models (Ghil et al., 1981), and, (ii) that the general Kalman filter methodology has long been adapted to account for model error (Friedland, 1969). Recently, the ensemble technique has been proposed as a computationally feasible means of applying Kalman filtering to the very high dimensional states inherent in global atmospheric models (Evensen, 1994; Houtekamer and Mitchell, 1998, 2001; Anderson, 2001; Bishop et al., 2001; Hamill et al., 2001; Whitaker and Hamill, 2002; Ott et al., 2004a). One of our goals in this paper is to investigate the incorporation of model error correction in an ensemble Kalman filter. Here, we consider only the case where the evolution of the model error is governed by a deterministic equation, i.e. the model error has no random component. We refer to this type of error as model bias, since when the state of the model is described by a probabilistic variable, as is the case in data assimilation, such errors become equal to the expected error in the model forecast.

We will restrict our considerations to the example of one particular ensemble Kalman filter, the Local Ensemble Kalman Filter (LEKF) proposed by Ott et al. (2004a). [This scheme has been successfully tested (Szunyogh et al., 2005) on a reduced resolution version of the operational Global Forecast System (GFS) of the National Centers for Environmental Prediction (NCEP) for the perfect forecast model scenario.] We believe that our results in the present paper, using the LEKF example, may also be more generally applicable, providing an indication of what to expect if model bias correction is attempted using other related ensemble Kalman filter methods. In addition, some of our ideas may also be useful in designing weakly constrained 4DVAR schemes (e.g. Derber, 1989; Zupanski, 1997), which also allow for model errors.

The technique we propose belongs to the family of schemes usually called *state space augmentation* methods (Cohn, 1997). In these techniques the state vector is augmented with the uncertain model parameters, and the augmented state is estimated using the forecast model in conjunction with observations. In principle, the uncertain parameters can occur in otherwise completely known forecast model equations. In such a case, the augmented state space approach may provide an accurate

*Corresponding author.
e-mail: sjbaek@glue.umd.edu
DOI: 10.1111/j.1600-0870.2006.00178.x

estimate of the parameters even for a highly chaotic system, as recently demonstrated on a simple model by Annan and Hargreaves (2004). In reality, the equations governing the motion of the atmospheric flow are not known exactly, thus uncertainties also arise due to our limited knowledge of the dynamics. Also, estimating all parameters of the forecast model equations would not be computationally feasible due to the large number of parameters. A practical approach, first suggested by Derber (1989), is to assume that the uncertainties in the forecast model can be approximately represented in the form of a limited number of bulk error terms. Then the task is to estimate the parameters of the bulk error terms. We recall that since the error terms are modeled as random vectors, the parameters to be estimated are the mean errors (model biases).

The information encapsulated in the bias can be used either to modify the forecast model equations or to modify the analysis scheme. Here we follow the second approach. That is, we treat the forecast model as a 'black box', that does not yield the true time evolution of the atmosphere, and we attempt to use this black box in conjunction with observations to account for model bias in the state estimation.

The two key components of the aforementioned strategy are the selection of a bias model that efficiently represents the bias and the design of a computational strategy that can efficiently estimate the parameters of the bias model. The most common assumption is that the bias is constant or has a simple evolution in time. It is also frequently assumed that the uncertainties in the forecast model state and the bias are uncorrelated. These assumptions were used to derive the bias estimation schemes of Dee and Da Silva (1998), Dee and Todling (2000), Carton et al. (2000), Martin et al. (2002) and Bell et al. (2004).

The scheme we propose allows for correlations between the uncertainties of the forecast model state and the bias. This additional flexibility is necessitated by the structure of our technique (see Section 2.4), and is affordable due to the high computational efficiency of the LEKF approach. In Section 2, we introduce three different bias models. Bias Model I is a simple additive correction to the model forecast. Bias Model II is motivated by envisioning a situation in which the forecast model evolution takes place on an attractor that is shifted from the attractor for the true system evolution. Bias Model III is essentially a combination of Bias Models I and II. Section 3 presents the results of numerical experiments with the Lorenz-96 model (e.g. Lorenz and Emanuel, 1998) for several cases of the difference between the forecast model evolution and the evolution of the true state. Conclusions and discussion follow in Section 4.

A main result is the importance of selecting a bias model that effectively parameterizes the form of possible bias in the forecast model. In particular, if the bias model can parameterize the possible bias of the forecast model, then our results suggest that substantial improvement in forecasts may result. On the other hand, if the parameterization of the model bias through the bias model does not sufficiently capture the form of the true biases in the forecast model, then substantial forecast improvements were not obtained in our numerical experiments.

2. Bias modeling and data assimilation

The discrepancy between the forecast model evolution and the evolution of the real atmosphere has two sources: (i) due to numerical solution on a grid of a finite number of points, the forecast model state is a finite dimensional representation of the infinite dimensional atmospheric fields, and (ii) the equations that govern the true evolution of the atmosphere are not known exactly. These two sources of forecast model errors are not independent, since the errors in the forecast model formulation are mainly associated with the inherent problems of considering only a limited number of interactions between the finite number of components of the state vector and the imperfect representation of the effects of the subgrid processes on the motions at the resolved scales.

Denoting the true atmospheric state at t_n by $\mathbf{x}_n^t$, the true atmospheric evolution is denoted

$$\mathbf{x}_{n+1}^t = \mathbf{F}^t\left(\mathbf{x}_n^t\right), \tag{1}$$

where $\mathbf{F}^t$ is the (unknown) true atmospheric evolution operator and $\mathbf{x}_{n+1}^t$ is the true atmospheric state at time $t_{n+1} = t_n + \Delta t$. Denoting the forecast model state at time t_n by $\mathbf{x}_n^m$, the black box produces a forecast model state at time t_{n+1},

$$\mathbf{x}_{n+1}^m = \mathbf{F}^m\left(\mathbf{x}_n^m\right), \tag{2}$$

where $\mathbf{F}^m$ is the forecast model evolution operator that mimics $\mathbf{F}^t$. Note that the dimensions of $\mathbf{x}^t$ and $\mathbf{x}^m$ are, in general, different; for example, in the case of real weather forecasting the true state is infinite dimensional and the forecast model state is finite dimensional. In what follows we will treat a scenario in which $\mathbf{x}^m$ and $\mathbf{x}^t$ have the same (finite) dimensionality, and we henceforth assume this circumstance. With respect to the situation of atmospheric weather forecasting, this assumption restricts the character of the errors and their means (biases) that can be addressed by our bias models. In particular, we regard our treatment to follow as addressing only those types of forecast model biases that can be represented as dynamics in the state space of the forecast model variables. Thus, we ignore dynamics that occur at smaller scales than the forecast model resolves. On the other hand, if we think of the dynamics at the unresolved scales as random perturbations to the forecast model dynamics, our methods may be able to correct for the mean bias due to such perturbations. Meanwhile, the uncertainty in the small-scale fluctuations is modeled as representativeness error in the observation error statistics.

In this section, we define three ways of modeling the bias that can arise due to forecast model error. We refer to these as Bias Model I, Bias Model II and Bias Model III.

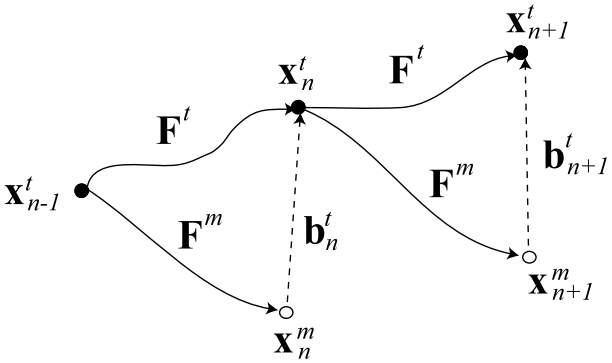

Fig. 1. Illustration of Bias Model I: $\mathbf{x}^t_n$ is the true state evolving according to $\mathbf{F}^t$ from the previous true state, and $\mathbf{x}^m_n$ is the forecast model state evolving according to $\mathbf{F}^m$ from the previous true state.

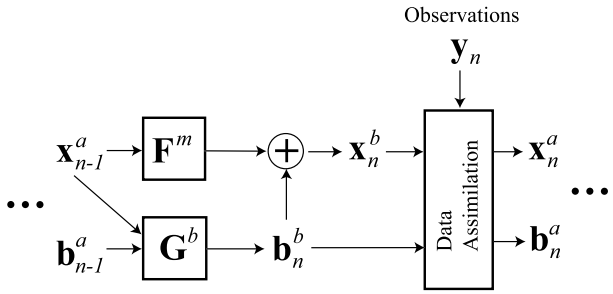

Fig. 2. Illustration of data assimilation with Bias Model I: Data assimilation produces an unbiased analysis for the true state, $\mathbf{x}^a_n$, and an analysis for the bias correction term, $\mathbf{b}^a_n$.

2.1. Bias Model I

In general, it is desired to have the forecast model state as close to the true state as possible so that, assuming that the forecast model and the true evolution operators are the same, the forecast model state stays near the true state after its evolution. In practice, however, the forecast model evolution operator differs from the true evolution operator. As a result, even if we evolve the forecast model state from an initial condition corresponding to the true state at the initial time (e.g. $\mathbf{x}^m_{n-1} = \mathbf{x}^t_{n-1}$), it is likely that the forecast model state departs from the true state as it evolves. In Bias Model I, we attempt to estimate

$$\mathbf{b}^t_n \;=\; \mathbf{F}^t\left(\mathbf{x}^t_{n-1}\right) - \mathbf{F}^m\left(\mathbf{x}^t_{n-1}\right), \tag{3}$$

i.e. the departure of the forecast model state from the true state as illustrated in Fig. 1.

In order to estimate $\mathbf{b}^t_n$, we must use some model of how it is related to past values of the bias $\mathbf{b}^t_i$ $(i < n)$. For example, we can assume that the bias is constant in time. If so, the true system evolution can be written in terms of the model evolution as follows:

$$\mathbf{x}^t_n = \mathbf{F}^m\left(\mathbf{x}^t_{n-1}\right) + \mathbf{b}^t_n, \tag{4}$$

$$\mathbf{b}^t_n = \mathbf{b}^t_{n-1}. \tag{5}$$

Though we could write this system more concisely as $\mathbf{x}^t_n = \mathbf{F}^m(\mathbf{x}^t_{n-1}) + \mathbf{b}$, where $\mathbf{b}$ is an unknown parameter vector, we write the system in terms of the augmented state vector $(\mathbf{x}^t_n, \mathbf{b}^t_n)$ in order to facilitate the iterative estimation of both $\mathbf{x}^t_n$ and $\mathbf{b}^t_n$ by our data assimilation procedure. More generally, we can replace eq. (5) with another model for the bias of the form,

$$\mathbf{b}^t_n \;=\; \mathbf{G}^b\left(\mathbf{b}^t_{n-1}, \mathbf{x}^t_{n-1}\right), \tag{6}$$

where $\mathbf{G}^b$ is the evolution operator for the bias correction term. Another alternative is to assume that the model error evolution is a Markov process. In that case, $\mathbf{b}^t_n$ is represented as $\mathbf{b}^t_{n-1}$ multiplied by a matrix that describes the temporal covariance between the model errors at different spatial location, and the right-hand side of eq. (6) also includes an additive random term (e.g. Jazwinski, 1970; Daley, 1992; Zupanski, 1997).

Then given an estimate $(\mathbf{x}^a_{n-1}, \mathbf{b}^a_{n-1})$ of the augmented state vector at time t_{n-1} (the 'analysis' from the previous data assimilation), we take the forecast (or 'background') of this vector $(\mathbf{x}^b_n, \mathbf{b}^b_n)$ at time t_n to be

$$\mathbf{x}^b_n = \mathbf{F}^m\left(\mathbf{x}^a_{n-1}\right) + \mathbf{b}^b_n, \tag{7}$$

$$\mathbf{b}^b_n = \mathbf{G}^b\left(\mathbf{b}^a_{n-1}, \mathbf{x}^a_{n-1}\right), \tag{8}$$

where we have assumed bias evolution by eq. (6). We then perform data assimilation using $(\mathbf{x}^b_n, \mathbf{b}^b_n)$, and the observations at time t_n to obtain the analysis $(\mathbf{x}^a_n, \mathbf{b}^a_n)$.

This way of taking forecast model error into account is illustrated in Fig. 2, and is the general scheme used in several previous methods appearing in the literature (e.g. Dee and Da Silva, 1998; Dee and Todling, 2000; Carton et al., 2000; Martin et al., 2002; Bell et al., 2004). The vector $\mathbf{y}_n$ in Fig. 2 is the observation of the true state at time t_n, which we assume to obey a model equation of the form,

$$\mathbf{y}_n \;=\; \mathbf{H}\left(\mathbf{x}^t_n\right) + \boldsymbol{\epsilon}_n, \tag{9}$$

where $\mathbf{H}$ is the observation operator mapping the true states to the observations. and $\boldsymbol{\epsilon}_n$ is the observational noise.

Basically, in Bias Model I it is supposed that the best forecast is produced when the input to the forecast model evolution is as close to the truth as possible. One can imagine problems with this. For example, say that atmospheric balance for the forecast model is not the same as that for the true atmosphere. Then, if a very good estimate of the true state at time t_n is inserted into the forecast model, the forecast model state at time t_n could often be out of balance, and spurious gravity wave activity might be excited. In the practice of numerical weather prediction, such spurious gravity wave activity is prevented by a filtering process, called initialization, applied to the fields provided by the data assimilation process (e.g. Machenhauer, 1977; Baer and Tribbia, 1977; Lynch and Huang, 1992; Lynch, 1997). The general wisdom is that a well-designed Kalman filter might eliminate the

need for initialization process. This consideration motivates Bias Model II.

2.2. Bias Model II

A consequence of the imperfect model is that the forecast model system has a different attractor from the true system. In some cases, it might be desirable to let the forecast model state follow its own attractor, since plugging a very good approximation of the true state into the forecast model system can result in completely different dynamics (like gravity wave excitation). In addition, one can envision a situation in which forecast model dynamics and true dynamics become more similar through an (a priori unknown) coordinate transformation. For instance, such transformations were rigorously derived to correct for truncation errors in numerical solution of the two-dimensional Navier–Stokes equations (Margolin et al., 2003). Having found a similar transformation for the weather prediction model, we may obtain a better estimate of the true trajectory by applying this transformation to an appropriate forecast model trajectory after it has been computed than by forcing the forecast model state to be close to the truth and then computing its trajectory. For simplicity, we assume the transformation is just a shift of the forecast model state to the true state, and we define the bias $\mathbf{c}_n^t$ at time t_n by

$$\begin{aligned}\mathbf{c}_n^t &= \mathbf{F}^t\left(\mathbf{x}_{n-1}^t\right) - \mathbf{F}^m\left(\mathbf{x}_{n-1}^m\right) \\ &= \mathbf{F}^t\left(\mathbf{x}_{n-1}^t\right) - \mathbf{F}^m\left(\mathbf{x}_{n-1}^t - \mathbf{c}_{n-1}^t\right).\end{aligned} \tag{10}$$

A schematic illustration of this bias model is shown in Fig. 3. The forecast model state is not pushed to the true state. Instead, it mimics the true dynamics in a shifted location of the state space.

Unlike the bias in Bias Model I, the bias in Bias Model II at time t_n depends not only on $\mathbf{F}^t$, $\mathbf{F}^m$, and $\mathbf{x}_{n-1}^t$, but also on the previous bias $\mathbf{c}_{n-1}^t$. Nonetheless, we may assume that for some choice of $\mathbf{c}_{n-1}^t$, the correction term $\mathbf{c}_n^t$ approximately obeys a simplified evolution model such as $\mathbf{c}_n^t = \mathbf{c}_{n-1}^t$, or more generally $\mathbf{c}_n^t = \tilde{\mathbf{G}}^c(\mathbf{c}_{n-1}^t, \mathbf{x}_{n-1}^t)$. In terms of this model, we approximate the true system evolution by the augmented model system,

$$\mathbf{x}_n^t = \mathbf{F}^m\left(\mathbf{x}_{n-1}^t - \mathbf{c}_{n-1}^t\right) + \mathbf{c}_n^t, \tag{11}$$

$$\mathbf{c}_n^t = \tilde{\mathbf{G}}^c\left(\mathbf{c}_{n-1}^t, \mathbf{x}_{n-1}^t\right). \tag{12}$$

For this bias model (and for Bias Model III to follow), our goal is not that the analysis state vector $\mathbf{x}_n^a$ closely approximates the true state $\mathbf{x}_n^t$, but rather that it approximates the best forecast model state $\mathbf{x}_n^m = \mathbf{x}_n^t - \mathbf{c}_n^t$ from which to make future forecasts. Thus, we rewrite eqs. (11) and (12) as

$$\mathbf{x}_n^m = \mathbf{F}^m\left(\mathbf{x}_{n-1}^m\right), \tag{13}$$

$$\mathbf{c}_n^t = \mathbf{G}^c\left(\mathbf{c}_{n-1}^t, \mathbf{x}_{n-1}^m\right), \tag{14}$$

where $\mathbf{G}(\mathbf{c}, \mathbf{x}) = \tilde{\mathbf{G}}^c(\mathbf{c}, \mathbf{x} + \mathbf{c})$. We can then write the background augmented state vector $(\mathbf{x}_n^b, \mathbf{c}_n^b)$ in terms of the previous analysis $(\mathbf{x}_{n-1}^a, \mathbf{c}_{n-1}^a)$ as follows:

$$\mathbf{x}_n^b = \mathbf{F}^m\left(\mathbf{x}_{n-1}^a\right), \tag{15}$$

$$\mathbf{c}_n^b = \mathbf{G}^c\left(\mathbf{c}_{n-1}^a, \mathbf{x}_{n-1}^a\right), \tag{16}$$

In taking this approach, one must keep in mind that the bias should be added to the forecast model state vector whenever making comparisons to observations. Thus instead of eq. (9), we use the observation model,

$$\mathbf{y}_n = \mathbf{H}\left(\mathbf{x}_n^m + \mathbf{c}_n^t\right) + \boldsymbol{\epsilon}_n \tag{17}$$

when performing data assimilation. The analysis $(\mathbf{x}_n^a, \mathbf{c}_n^a)$ represents an approximation to the augmented state vector $(\mathbf{x}_n^m, \mathbf{c}_n^t)$, and thus forecasts made using $\mathbf{x}_n^a$ as the initial condition should also be corrected by the approximated bias in order to better predict the true system. Data assimilation with Bias Model II is illustrated in Fig. 4.

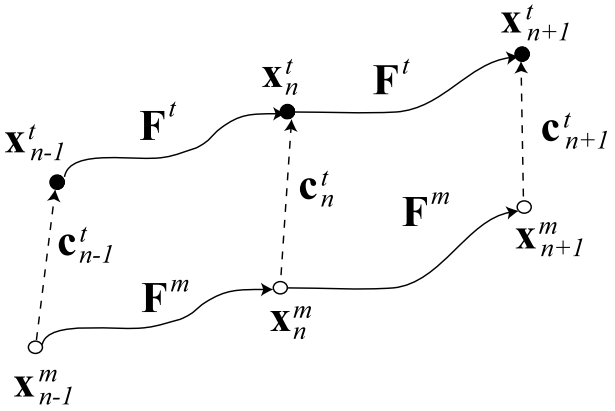

Fig. 3. Illustration of Bias Model II: $\mathbf{x}_n^t$ and $\mathbf{x}_n^m$ evolve according to their own dynamics but the behavior of the forecast model is similar to the behavior of the truth.

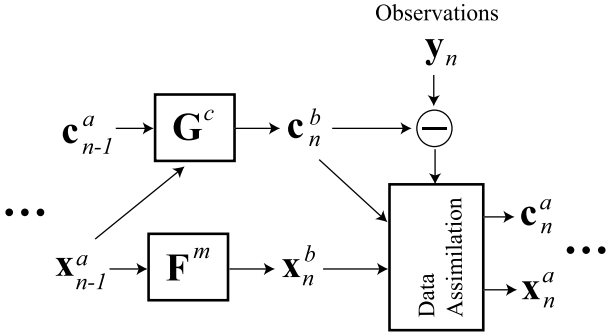

Fig. 4. Illustration of data assimilation with Bias Model II in the case $\mathbf{H}(\mathbf{x}) = \mathbf{x}$: Data assimilation produces an analysis of the best forecast model state, $\mathbf{x}_n^m = \mathbf{x}_n^t - \mathbf{c}_n^t$, and an analysis for the correction term, $\mathbf{c}_n^t$.

To the best of our knowledge, Bias Model II is a novel approach to the effects of model errors on the accuracy of the state estimates. Hansen (2002) also argued for the model attractor, but he suggested the use of a multi-model approach as opposed to the state augmentation we propose.

2.3. Bias Model III

In Bias Model III, we combine Bias Model I and Bias Model II. Formally, we combine the equations describing the previous two bias models in the following manner. We evolve the analysis augmented state vector $(\mathbf{x}^a_{n-1}, \mathbf{b}^a_{n-1}, \mathbf{c}^a_{n-1})$ to the background at the next step using the model

$$\mathbf{x}^b_n = \mathbf{F}^m\left(\mathbf{x}^a_{n-1}\right) + \mathbf{b}^b_n \tag{18}$$

$$\mathbf{b}^b_n = \mathbf{G}^b\left(\mathbf{x}^a_{n-1}, \mathbf{b}^a_{n-1}, \mathbf{c}^a_{n-1}\right) \tag{19}$$

$$\mathbf{c}^b_n = \mathbf{G}^c\left(\mathbf{x}^a_{n-1}, \mathbf{b}^a_{n-1}, \mathbf{c}^a_{n-1}\right), \tag{20}$$

and we compare the background state with observations according to eq. (17). Since $\mathbf{x}^b_n$ and $\mathbf{c}^b_n$ represent the best available approximations to $\mathbf{x}^m_n$ and $\mathbf{c}^t_n$ prior to the data assimilation at time t_n, the observation increment we use is

$$\mathbf{y}_n - \mathbf{H}\left(\mathbf{x}^b_n + \mathbf{c}^b_n\right). \tag{21}$$

Notice that if $\mathbf{G}^b(\mathbf{x}, \mathbf{b}, \mathbf{c}) = \mathbf{0}$, then this model reduces to Bias Model II, while if $\mathbf{G}^c(\mathbf{x}, \mathbf{b}, \mathbf{c}) = \mathbf{0}$, this model reduces to Bias Model I. In its simplest form, our model uses $\mathbf{G}^b(\mathbf{x}, \mathbf{b}, \mathbf{c}) = \mathbf{b}$ and $\mathbf{G}^c(\mathbf{x}, \mathbf{b}, \mathbf{c}) = \mathbf{c}$. However, we find a slightly different bias evolution function to be advantageous in some situations (see Section 3.6).

2.4. Augmented local ensemble Kalman filter

For the purposes of all subsequent discussion we henceforth take the system state at time t_n to be a scalar variable $x_{n,i}$ defined on a discrete one-dimensional spatial domain, $i = 1, 2, \ldots, N$. Thus we represent the system state as a vector $\mathbf{x}_n = [x_{n,1}, x_{n,2}, \ldots, x_{n,N}]^\mathsf{T}$, where the superscript T denotes the transpose.

Once a suitable model for the bias is chosen, it can be incorporated into the formulation of the Kalman filter. For example, in the case of Bias Model III, the new equations can be obtained by replacing the state $\mathbf{x}_n$, by the augmented state, $\mathbf{v}_n = [\mathbf{x}_n, \mathbf{b}_n, \mathbf{c}_n]^\mathsf{T}$, in the Kalman filter equations. Here, the correction terms, $\mathbf{b}_n = [b_{n,1}, b_{n,2}, \ldots, b_{n,N}]^\mathsf{T}$, and $\mathbf{c}_n = [c_{n,1}, c_{n,2}, \ldots, c_{n,N}]^\mathsf{T}$, have the same dimension, N, which is typically equal to the number of grid-point variables in a numerical weather prediction model. By inserting the augmented state into the Kalman filter equations, we assume that $\psi\ (\mathbf{v}^b_n)$, the background probability distribution of the augmented state, is Gaussian; that is,

$$\psi\left(\mathbf{v}^b_n\right) \sim \exp\left[-\frac{1}{2}\left(\mathbf{v}^b_n - \bar{\mathbf{v}}^b_n\right)^\mathsf{T}\left(\mathbf{P}^b_\mathbf{v}\right)^{-1}\left(\mathbf{v}^b_n - \bar{\mathbf{v}}^b_n\right)\right], \tag{22}$$

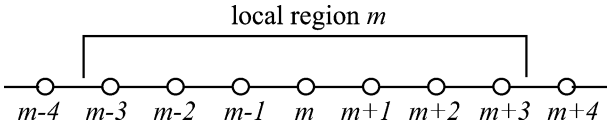

Fig. 5. Illustration of a local state centered about location m.

where $\bar{\mathbf{v}}^b_n$ is the background mean of the augmented state, and $\mathbf{P}^b_\mathbf{v}$ is the background error covariance matrix for the augmented state.

The main computational challenge in designing an augmented Kalman filter is to find a computationally efficient approach to estimate $\mathbf{P}^b_\mathbf{v}$, whose dimension increases by N when a new parameter is added to the state. One frequently applied approach to reduce the computational burden associated with the estimation of $\mathbf{P}^b_\mathbf{v}$ is to assume that many entries of the matrix are zero, e.g. by assuming that the (non-augmented) state and the bias parameters are uncorrelated. We propose a different approach, which involves estimating the background mean and the background error covariance matrix by an ensemble, and solving the ensemble Kalman filter equations locally in grid space applying the Local Ensemble Kalman Filter (Ott et al., 2004a) to the augmented state. The LEKF scheme estimates 'local states' as illustrated in Fig. 5. In particular, considering the LEKF procedure without model bias correction (i.e. as in Ott et al., 2004a), for each point m on the spatial grid, we consider a neighborhood consisting of the $2l + 1$ points centered at m; these points have locations $m - l, m - l + 1, \ldots, m, \ldots, m + l - 1, m + l$ (e.g. $l = 3$ in Fig. 5). At time $t = t_n$, the LEKF does data assimilation on local regions centered at each grid point using the local state,

$$\mathbf{x}_n(m) = \begin{bmatrix} x_{n,m-l} \\ x_{n,m-l+1} \\ \vdots \\ x_{n,m} \\ \vdots \\ x_{n,m+l-1} \\ x_{n,m+l} \end{bmatrix}. \tag{23}$$

The global analysis state (i.e. the analysis state at $t = t_n$ at each grid point over the entire grid) is then taken to be the state at the center of each local region (see Ott et al., 2004a, for further discussion).

In order to adapt the LEKF to correct for model bias, we augment each local state to include the bias estimate of the bias model employed. For example, for Bias Model III, we form an augmented local state, $\mathbf{v}_n(m) = [\mathbf{x}_n(m), \mathbf{b}_n(m), \mathbf{c}_n(m)]^\mathsf{T}$, for the data assimilation at location m. Similarly, for Bias Model I, $\mathbf{v}_n(m) = [\mathbf{x}_n(m), \mathbf{b}_n(m)]^\mathsf{T}$, and, for Bias Model II, $\mathbf{v}_n(m) = [\mathbf{x}_n(m), \mathbf{c}_n(m)]^\mathsf{T}$. Since the augmented local state is derived from the global state $\mathbf{v}_n$, it can be also assumed to have a Gaussian

distribution,

$$\psi_m\left(\mathbf{v}_n^b(m)\right) \sim \exp\left\{-\frac{1}{2}\left[\mathbf{v}_n^b(m)-\bar{\mathbf{v}}_n^b(m)\right]^{\mathsf{T}} \times\left[\mathbf{P}_n^b(m)\right]^{-1}\left[\mathbf{v}_n^b(m)-\bar{\mathbf{v}}_n^b(m)\right]\right\}, \quad (24)$$

where $\bar{\mathbf{v}}_n^b(m)$ is the background mean of the augmented local state, and $\mathbf{P}_n^b(m)$ is the background error covariance matrix for the augmented local state. In this way, the dimension of the space for data assimilation is reduced to $2(2l+1)$ for Bias Model I or II and to $3(2l+1)$ for Bias Model III. An important property of this scheme is that it allows for (and also requires) the estimation of cross-correlations between uncertainties in the state estimates and uncertainties in the estimation of the model bias terms.

3. Numerical experiments

3.1. Experimental setup

For testing our assimilation scheme, we consider the Lorenz-96 model (Lorenz and Emanuel, 1998) which is defined by the system of differential equations,

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = (x_{i+1}-x_{i-2})x_{i-1}-x_i+\Theta, \quad i=1,\ldots,N, \quad (25)$$

where $x_{-1}=x_{N-1}, x_0=x_N, x_{N+1}=x_1$ and Θ is a constant. The variables form a cyclic chain and may be thought of as roughly analogous to the values of some unspecified scalar meteorological quantity at N equally spaced sites along a latitude circle. For compactness of notation, we will also represent eq. (25) as

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{L}(\mathbf{x}), \quad (26)$$

where $\mathbf{x}=[x_1, x_2, \ldots, x_N]^{\mathsf{T}}$. We solve eq. (25) with a fourth-order Runge–Kutta method using a time step of 0.05 dimensionless units for which the system is computationally stable. Lorenz and Emanuel (1998) consider this time step as roughly corresponding to 6 h of real atmospheric evolution. With $\Theta=8.0$ and $N=40$, Lorenz and Emanuel demonstrate that the system (25) results in a westward (i.e. in the direction of low index of locations) progression of individual maxima and minima and an eastward progression of the center of activity with a dominant wavenumber of 8. In addition, they also find that the system is chaotic with 13 positive Lyapunov exponents and a Lyapunov dimension of 27.1. Throughout our numerical experiments we use $\Theta=8.0$ and $N=40$.

In what follows we will assume that our forecast model dynamics is given by eq. (25) but that the true dynamics of the system whose state we are concerned with obeys dynamics that may differ from those of our forecast model. We will consider situations in which the true dynamics differ from the forecast model in three ways, which we refer to as Type A truth bias, Type B truth bias, and Type C truth bias. The dynamical behaviors of the true system in these three cases are as follows:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{L}(\mathbf{x})+\boldsymbol{\beta} \quad \text{(Type A)}, \quad (27)$$

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{L}(\mathbf{x}+\boldsymbol{\zeta}) \quad \text{(Type B)}, \quad (28)$$

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{L}(\mathbf{x}+\boldsymbol{\zeta})+\boldsymbol{\beta} \quad \text{(Type C)}, \quad (29)$$

where $\boldsymbol{\beta}=[\beta_1,\beta_2,\ldots,\beta_N]^{\mathsf{T}}$ and $\boldsymbol{\zeta}=[\zeta_1,\zeta_2,\ldots,\zeta_N]^{\mathsf{T}}$. When the true dynamics is described by the same equation [eq. (25)] as the forecast model, we say that the forecast model is 'perfect'. Note that Bias Model I would be a natural choice if it were known that the deviation of the true dynamics from the model dynamics (26) was such that the true dynamics belonged to a family of systems of the form given by (27) (Type A truth bias). Similar statements apply with regard to the relation between Bias Model II and Type B truth bias and between Bias Model III and Type C truth bias.

With small values of $\boldsymbol{\beta}$ and $\boldsymbol{\zeta}$, we conjecture that the systems (27)–(29) exhibit behaviors similar to those of system (25). In our numerical experiments, the elements of $\boldsymbol{\beta}$ and $\boldsymbol{\zeta}$ vary in space (i) and have the forms

$$\beta_i = A\sin\left(2\pi\frac{i-1}{N}\right), \quad (30)$$

$$\zeta_i = B\sin\left(2\pi\frac{i-1}{N}\right), \quad i=1,\ldots,N, \quad (31)$$

where A and B are scalar constants.

The true states are generated by integrating one of the three eqs. (27)–(29), while the forecast model states are generated by integrating eq. (26). The evolution operators, $\mathbf{F}^t$ and $\mathbf{F}^m$, are the integrations of the above eqs. (26)–(29) from some time t to $t+\Delta t$ where $\Delta t=0.05$ and the states are available at every discrete time $t_n=t_0+n\Delta t$, where t_0 is the time at which an experiment begins and n is a positive integer.

We assume that the observations are available at every time t_n for $n\geq 0$ and the state variables themselves are directly observed. Thus the observation operator in eqs. (9) and (17) is the identity operator [i.e. $\mathbf{H}(\mathbf{x})=\mathbf{x}$]. We also assume that the observational noise ϵ_n has zero-expected value and is uncorrelated, white and Gaussian with variance σ^2. Thus the local observation error covariance matrix is a diagonal matrix whose components are σ^2. Correspondingly, we generate our simulated 'observations' (9) by adding uncorrelated Gaussian random numbers with variance σ^2 to the true state variables x_i^t and form a local observation $\mathbf{y}_n(m)=[y_{n,m-l},\ldots,y_{n,m+l}]^{\mathsf{T}}$. Throughout our numerical experiments, we take $\sigma^2=0.09$.

Data assimilations are done at every integration time t_n. The analysis error is defined as

$$\mathbf{e}_n^a = \bar{\mathbf{x}}_n^a-\mathbf{x}_n^t, \quad (32)$$

for Bias Model I, and

$$\mathbf{e}_n^a = \bar{\mathbf{x}}_n^a+\bar{\boldsymbol{\zeta}}_n^a-\mathbf{x}_n^t, \quad (33)$$

for Bias Model II and Bias Model III, where $\bar{\mathbf{x}}_n^a$ is the ensemble mean of the analysis and where $\bar{\mathbf{c}}_n^a$ is the ensemble mean of the estimate of the Type II bias. We use the root-mean-square (rms) of the analysis error,

$$\mathrm{RMS}\{\mathbf{e}_n^a\} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(e_{n,i}^a\right)^2}, \tag{34}$$

to assess the quality of the analysis at a given time, and the time mean of the rms error over a long time interval T,

$$\langle\langle \mathbf{e}^a \rangle\rangle = \frac{1}{T}\sum_{n=n_0+1}^{n_0+T} \mathrm{RMS}\{\mathbf{e}_n^a\}, \quad T \gg 1, \tag{35}$$

to measure the overall performance of the assimilation scheme. Here, n_0 is the time we allow for the analysis to converge to the true state.

To improve the analyses in our numerical experiments, we employ variance inflation,

$$\hat{\mathbf{P}}_n^a(m) \rightarrow \hat{\mathbf{P}}_n^a(m) + \frac{\mu\Lambda}{k}\mathbf{I}_k, \tag{36}$$

where $\hat{\mathbf{P}}_n^a(m)$ is the local analysis error covariance matrix defined in the 'internal' coordinate system (Ott et al., 2004a) whose basis is the set of eigenvectors of the local background error covariance matrix $\mathbf{P}_n^b(m)$, μ is an inflation coefficient, and $\Lambda = \mathrm{Trace}\{\hat{\mathbf{P}}_n^a(m)\}$. This particular form of variance inflation was proposed in Ott et al. (2004a) where it is referred to as 'enhanced variance inflation'. Enhanced variance inflation has the effect of enhancing the estimated probability of error in directions that formally show only very small error probability. [This modification of $\hat{\mathbf{P}}_n^a(m)$ also modifies the ensemble perturbations through the square root filter; see Ott et al. (2004a).] The general purpose of employing a variance inflation is to correct for the loss of variance in the ensemble due to non-linearities and sampling errors. Most importantly, variance inflation can also stabilize the Kalman filter in the presence of model errors, as it was shown in Ott et al. (2004b) for the Loranz-96 model and in Whitaker et al. (2004) for the NCEP GFS model preparing a historical reanalysis data set. Since variance inflation schemes are computationally less expensive than the state augmentation method, we hope to see that the technique we propose here lead to larger improvements in the accuracy of the state estimates than what can be achieved by simply tuning the variance inflation coefficient μ.

For the dimension of local states used in the LEKF, we select 13 (i.e. $l = 6$) which is known to be a good choice for the Lorenz model (25) with $\Theta = 8$ and $N = 40$ (Ott et al., 2004a). Hence, the augmented local states have 26 dimensions for the states used in Bias Model I or Bias Model II, and 39 dimensions for the states used in Bias Model III. In our numerical experiments, we choose the number of ensemble members to be the same as the dimension of the augmented local states so that the local background error covariance matrix has full rank. This choice means that the ensemble size is 13 when bias is not estimated in the assimilation, 26 when Bias Model I or II is used, and 39 when Bias Model III is used. Thus we take into account the added dimensionality of the augmented local states, anticipating that this increased dimensionality necessitates correspondingly increased ensemble size in order to properly represent it. This increased ensemble size is part of the added computational cost that is paid in order to correct for model bias. In practice, for given computer resources, the need for a large ensemble may thus necessitate consideration of benefit trade-offs among ensemble size, local domain size, model resolution, etc.

Finally, for bias evolution [eqs. (8), (16), (19) and (20)] we use $\mathbf{G}^b(\mathbf{x}, \mathbf{b}, \mathbf{c}) = \mathbf{b}$ and $\mathbf{G}^c(\mathbf{x}, \mathbf{b}, \mathbf{c}) = \mathbf{c}$, until Section 3.6 where we consider different evolution.

3.2. Perfect forecast model

We first test our bias models for the case of a perfect forecast model, i.e. for the case when the true values of $\boldsymbol{\beta}$ and $\boldsymbol{\zeta}$ in eqs. (27)–(29) are $\mathbf{0}$. In this case, the evolution operators for the true state and the forecast model state are identical, $\mathbf{F}^t = \mathbf{F}^m$.

In order to generate the true states for this run, we first integrate eq. (25) for 10^4 time steps from a random initial condition, allowing the system to approach its attractor. After this, we perform data assimilation at every time step. The initial ensemble members for the first data assimilation are generated by adding independent, zero mean, normally distributed random numbers of variance 1.3 to the true state at every spatial point i. Before obtaining the rms time mean of the analysis error (35), we run 2×10^4 data assimilations to allow convergence. Past this time (denoted n_0), it is found that the rms analysis error reaches a statistically steady state in which it fluctuates about a temporally constant mean value which we denote $\langle\langle \mathbf{e}^a \rangle\rangle$.

The data plotted in Fig. 6 show the time-averaged rms analysis error (35) as a function of the inflation coefficient μ (36). Here, the rms analysis error is averaged over $T = 3 \times 10^4$ time steps. For the case in which no state augmentation is employed in the assimilation (data plotted as $*$ symbols) the best performance is obtained near $\mu = 0.005$ for which $\langle\langle \mathbf{e}^a \rangle\rangle = 0.057$. The other three curves in Fig. 6 show the analysis errors for the cases when the same observations are assimilated with using the three different bias models of Section 2.4 in the state estimation. In these cases, we try to estimate a bias that is zero in reality. The estimated bias terms tend to fluctuate about zero, resulting in a slight ($\sim$8%) increase of the error. The above results from examining this case provide a standard against which we can compare results that we will subsequently obtain for situations with error in the forecast model.

In Fig. 6, we see that the minima of $\langle\langle \mathbf{e}^a \rangle\rangle$ appear at a lower inflation coefficient when the states are augmented by an estimate of the bias, $\mathbf{b}^b$ (i.e. for Bias Models I and III). In order to see why this occurs we consider the local perturbations for the augmented

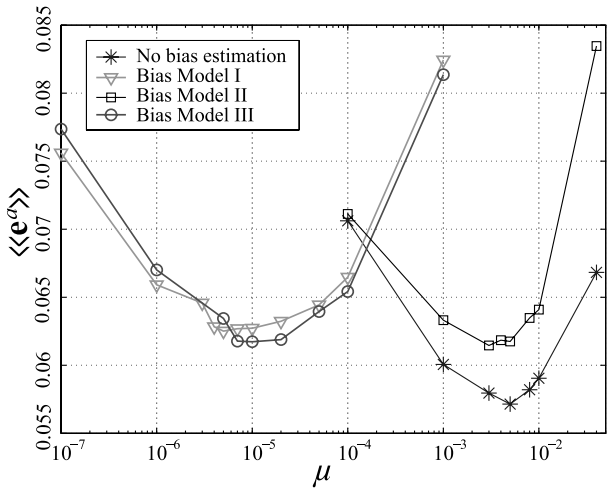

Fig. 6. Time-averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, versus variance inflation coefficient, μ, for the perfect forecast model experiment: With a perfect forecast model, any attempt to estimate and correct for a bias results in slightly higher analysis error.

local background,

$$\delta \mathbf{v}_n^{b(j)}(m) = \mathbf{v}_n^{b(j)}(m) - \bar{\mathbf{v}}_n^b(m), \tag{37}$$

where $\{\mathbf{v}_n^{b(j)}(m)\}$ are the ensemble members of the augmented local background. These perturbations are used to estimate the local background error covariance matrix (Ott et al., 2004a)

$$\mathbf{P}_n^b(m) \;=\; \mathbf{V}_n^b(m)\left[\mathbf{V}_n^b(m)\right]^{\mathsf{T}}, \tag{38}$$

where

$$\mathbf{V}_n^b(m) \;=\; k^{-\frac{1}{2}}\left[\delta \mathbf{v}_n^{b(1)}(m)\middle|\delta \mathbf{v}_n^{b(2)}(m)\middle|\cdots\middle|\delta \mathbf{v}_n^{b(k+1)}(m)\right] \tag{39}$$

and $k+1$ is the number of the local ensemble members. We rewrite eq. (37) using eq. (7) as

$$\delta \mathbf{v}_n^{b(j)}(m) = \begin{bmatrix} \delta \mathbf{x}_n^{b(j)}(m) \\ \delta \mathbf{b}_n^{b(j)}(m) \end{bmatrix} = \begin{bmatrix} \delta \tilde{\mathbf{x}}_n^{b(j)}(m) + \delta \mathbf{b}_n^{b(j)}(m) \\ \delta \mathbf{b}_n^{b(j)}(m) \end{bmatrix}, \tag{40}$$

where $\{\delta \mathbf{x}_n^{b(j)}(m)\}$ are perturbations for the local background, $\tilde{\mathbf{x}}_n^b = \mathbf{F}^m(\mathbf{x}_{n-1}^a)$, and $\{\delta \mathbf{b}_n^{b(j)}(m)\}$ are perturbations for the local prediction for Bias Model I. In our experiments we observe that $\{\delta \mathbf{b}_n^{b(j)}(m)\}$ are only weakly correlated with each other and almost uncorrelated with $\{\delta \tilde{\mathbf{x}}_n^{b(j)}(m)\}$. Effectively, therefore, uncorrelated random vectors are added to the state perturbations $\{\delta \tilde{\mathbf{x}}_n^{b(j)}(m)\}$ in eq. (40). In consequence, $\mathbf{P}_n^b(m)$ (38) is effectively inflated mostly on the diagonal components by the amount of the variance of $\{\delta \mathbf{b}_n^{b(j)}(m)\}$. (This effective inflation created by using eq. (7) to obtain the background will also be observed when the forecast model is not perfect as shown in the following subsections.)

3.3. Data assimilation with Type A truth bias

In this experiment, we perform data assimilation using the three augmented local states as described in Section 2.4 and an unaugmented state when the true state is evolved using eq. (27) with $A = 0.2\Theta = 1.6$ in eq. (30) corresponding to Type A truth bias. Again the forecast model state is evolved using eq. (25). We can approximate the bias $\mathbf{b}_n^t$ given by eq. (3) as follows. Recall that $\mathbf{F}^t$ and $\mathbf{F}^m$ are the time Δt maps of the true dynamics [eq. (27)] and the forecast model [eq. (26)] and that eq. (3) is based on the assumption that $\mathbf{x}^m(t_{n-1}) = \mathbf{x}^t(t_{n-1})$. Taking the difference between the true equation (27) and the model equation (26),

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{x}^t - \mathbf{x}^m) \;=\; \mathbf{L}(\mathbf{x}^t) + \boldsymbol{\beta} - \mathbf{L}(\mathbf{x}^m) \;\approx\; \boldsymbol{\beta}, \tag{41}$$

then, integrating eq. (41) for the time interval $t_{n-1} \le t \le t_n$ with the initial condition $\mathbf{x}^t(t_{n-1}) = \mathbf{x}^m(t_{n-1}) = \mathbf{x}_{n-1}^t$, to obtain $\mathbf{b}_n^t = \mathbf{x}^t(t_n) - \mathbf{x}^m(t_n)$, yields

$$\mathbf{x}^t - \mathbf{x}^m \;\approx\; \int_{t_{n-1}}^{t_n} \boldsymbol{\beta}\, dt \;=\; \boldsymbol{\beta}\Delta t. \tag{42}$$

Using eqs. (3) and (42) we obtain $\mathbf{b}_n^t \approx \beta\Delta t$. For the situation in this section, $\Delta t = 0.05$, and we have taken $\boldsymbol{\beta}$ to be constant in time,

$$\beta_i \;=\; A\sin\left(2\pi\frac{i-1}{N}\right) \;=\; 1.6\sin\left(2\pi\frac{i-1}{N}\right). \tag{43}$$

Time-averaged rms analysis errors for each case are shown in Fig. 7. In the case where the bias is not estimated, the error is around 0.167 at $\mu \approx 0.7$, which is still lower than the rms error of the noisy observations. If we, however, augment the state using Bias Model I the error is reduced dramatically, and slightly more if we augment the state using Bias Model III, yielding $\langle\langle \mathbf{e}^a \rangle\rangle = 0.068$ and 0.061, respectively, at $\mu \approx 10^{-5}$ [Fig. 7(b)]. If, however, we augment the state using Bias Model II, then the

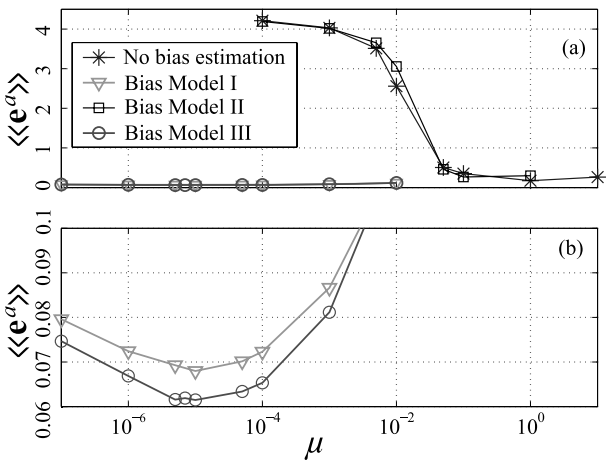

Fig. 7. Time-averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, versus μ for the case of Type A truth bias: Note that (b) shows the same results as (a) for Bias Model I and Bias Model III but for a different vertical scale.

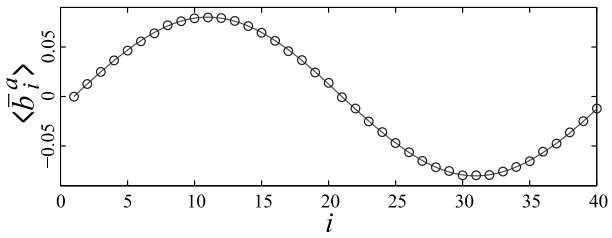

Fig. 8. The average bias estimate of location i is shown as ○. The approximate true bias $\beta_i \Delta t$ is shown as the solid curve.

rms error is around 0.225 at $\mu \approx 0.3$, worse than what is obtained when no bias estimation is employed. Here, we see again that using the unbiased background (7) for data assimilation effectively inflates the local background error covariance matrix, and a smaller variance inflation yields the lowest analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$. The good results obtained when the state is augmented using either Bias Model I or Bias Model III might reasonably be ascribed to the fact that estimation of $\mathbf{b}_n^t$ can be regarded as correcting for precisely the form of truth bias that is present when the truth evolves by eq. (27).

In Fig. 8, we plot $\langle \bar{\mathbf{b}}^a \rangle$, the time average of the ensemble mean of the bias estimate,

$$\langle \bar{\mathbf{b}}^a \rangle = \frac{1}{T} \sum_{n=n_0+1}^{n_0+T} \bar{\mathbf{b}}_n^a, \tag{44}$$

$$\bar{\mathbf{b}}_n^a = \frac{1}{k+1} \sum_{i=1}^{k+1} \mathbf{b}_n^{a(j)}, \tag{45}$$

where $T = 2000$ and $n_0 = 15\,000$, for the experiment with the state augmented using Bias Model I at $\mu = 10^{-5}$ (where the rms error is minimum in Fig. 7). We see that $\langle \bar{\mathbf{b}}^a \rangle$ agrees well with the approximation to $\mathbf{b}_n^t$ given by (42) and (43) (shown as the solid curve). Also, although the shape is not shown here, for the case that the state is augmented using Bias Model III, $\langle \bar{\mathbf{b}}^a \rangle$ again agrees well with (42) and (43).

We now examine the analysis errors using Bias Model I and III. In Fig. 9, we plot the analysis error (32) averaged over 5000 time steps,

$$\langle \mathbf{e}^a \rangle = \frac{1}{T} \sum_{n=n_0+1}^{n_0+T} \mathbf{e}_n^a, \tag{46}$$

(here, $T = 5000$ and $n_0 = 15\,000$) for the case that no bias estimation is performed in data assimilation (∗), the case that the estimation is performed using Bias Model I in the assimilation (▽), and the case that the bias estimation is performed using Bias Model III (○). The variance inflation coefficients are $\mu = 1.0$, 10^{-5}, and 10^{-5}, respectively, at which the errors, $\langle\langle \mathbf{e}^a \rangle\rangle$, are minimum for each case.

In order to understand why Bias Model III does better than Bias Model I for this case, recall that both models determine the background state $\mathbf{x}_n^b$ at time t_n by $\mathbf{x}_n^b = \mathbf{F}^m(\mathbf{x}_{n-1}^a) + \mathbf{b}_n^b$ where $\mathbf{b}_n^b$

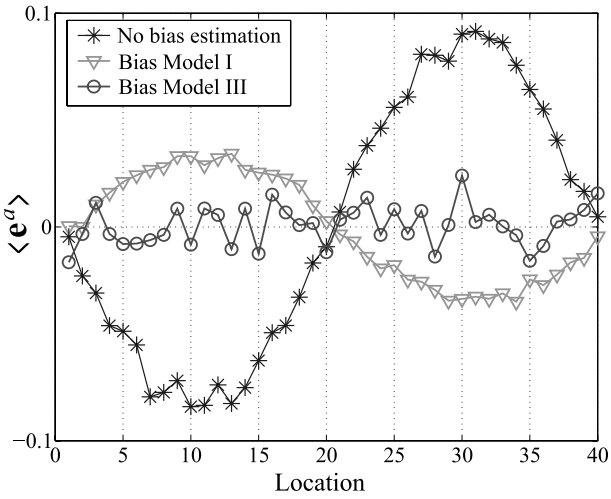

Fig. 9. Time-average of the analysis error as a function of location: The Type A truth bias is corrected best when we perform the assimilation using Bias Model III.

is approximately constant in time [eqs. (7) and (18)]. However, Bias Model I tries to make $\mathbf{x}_n^b$ close to $\mathbf{x}_n^t$, whereas Bias Model III tries to make $\mathbf{x}_n^b + \mathbf{c}_n^b$ close to $\mathbf{x}_n^t$, where $\mathbf{c}_n^b$ is also approximately constant in time. Suppose that $\mathbf{b}_n^b$ has converged to the time-independent approximation $\boldsymbol{\beta}\Delta t$ of $\mathbf{b}_n^t$ [see eq. (42)], that $\mathbf{c}_n^b$ has converged to a constant vector $\mathbf{c}$, and that the analysis at time t_{n-1} is perfect: $\mathbf{x}_{n-1}^a + \mathbf{c} = \mathbf{x}_n^t$. Consider the model trajectory $\mathbf{x}^m(t)$ of eq. (26) with $\mathbf{x}^m(t_{n-1}) = \mathbf{x}_{n-1}^a$ and the true trajectory $\mathbf{x}^t(t)$ of eq. (27) with $\mathbf{x}^t(t_{n-1}) = \mathbf{x}_{n-1}^t$. Then $\mathbf{x}_n^t = \mathbf{x}^t(t_n)$ and $\mathbf{x}_n^b = \mathbf{x}^m(t_n) + \boldsymbol{\beta}\Delta t$, and the background is most accurate if $\mathbf{x}_n^b + \mathbf{c} = \mathbf{x}_n^t$. As in eq. (41),

$$\begin{aligned}
\mathbf{x}_n^t - \left(\mathbf{x}_n^b + \mathbf{c}\right) &= \mathbf{x}^t(t_n) - \left(\mathbf{x}^m(t_n) + \boldsymbol{\beta}\Delta t + \mathbf{c}\right) \\
&= \mathbf{x}^t(t_{n-1}) - \left(\mathbf{x}^m(t_{n-1}) + \boldsymbol{\beta}\Delta t + \mathbf{c}\right) \\
&\quad + \int_{t_{n-1}}^{t_n} [\mathbf{L}(\mathbf{x}^t(t)) + \boldsymbol{\beta} - \mathbf{L}(\mathbf{x}^m(t))]\,\mathrm{d}t \\
&= \mathbf{x}_{n-1}^t - \left(\mathbf{x}_{n-1}^a + \boldsymbol{\beta}\Delta t + \mathbf{c}\right) \\
&\quad + \int_{t_{n-1}}^{t_n} [\mathbf{L}(\mathbf{x}^t(t)) - \mathbf{L}(\mathbf{x}^m(t))]\,\mathrm{d}t + \boldsymbol{\beta}\Delta t \\
&= \int_{t_{n-1}}^{t_n} [\mathbf{L}(\mathbf{x}^t(t)) - \mathbf{L}(\mathbf{x}^m(t))]\,\mathrm{d}t.
\end{aligned} \tag{47}$$

Thus, we desire that the average value of $\mathbf{L}(\mathbf{x}^t(t)) - \mathbf{L}(\mathbf{x}^m(t))$ to be as small as possible over the interval $t_{n-1} \le t \le t_n$. Using $\mathbf{c} = -\boldsymbol{\beta}\Delta t/2$ makes this average zero to first order, and is thus superior to using $\mathbf{c} = \mathbf{0}$, which corresponds to Bias Model I. (see Fig. 10.) This is confirmed in Fig. 11, from which one observes $\langle \bar{\mathbf{c}}^a \rangle \approx -\boldsymbol{\beta}\Delta t/2$ in the experiment with Bias Model III.

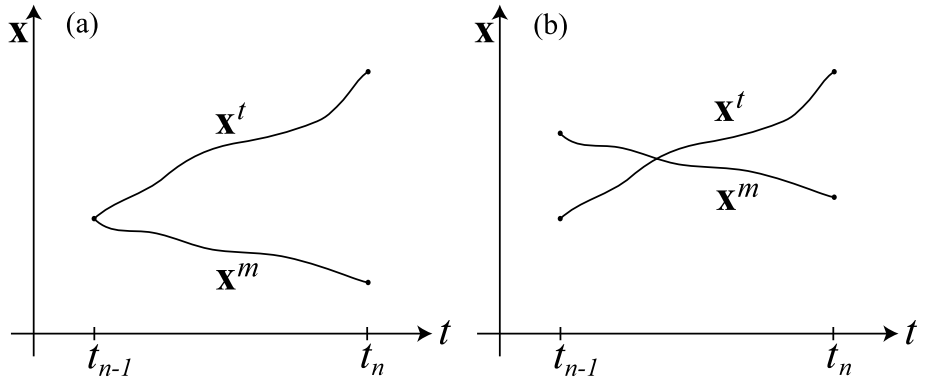

Fig. 10. Figure (a) depicts the case $\mathbf{c} = \mathbf{0}$, corresponding Bias Model I, while figure (b) depicts the case $\mathbf{c} \approx -\beta/2$ in Bias Model III.

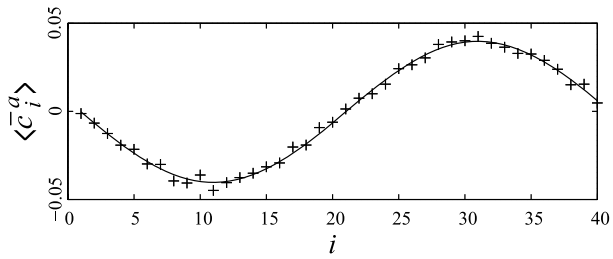

Fig. 11. The average bias estimate of location i is shown as $+$. The the value of $-\beta_i \Delta t/2$ is shown as the solid curve.

3.4. Data assimilation with Type B truth bias

In this experiment, we simulate a bias by evolving the true state with eq. (28) with temporally constant ζ and estimate it with three different augmentation methods as done in Section 3.3. To obtain the true value of $\mathbf{c}^t$, defined by (10), we use

$$\begin{aligned} \frac{d\mathbf{c}^t}{dt} &= \frac{\mathrm{d}\mathbf{x}^t}{\mathrm{d}t} - \frac{\mathrm{d}\mathbf{x}^m}{\mathrm{d}t} \\ &= \mathbf{L}(\mathbf{x}^t + \zeta) - \mathbf{L}(\mathbf{x}^m) \\ &= \mathbf{L}(\mathbf{x}^t + \zeta) - \mathbf{L}(\mathbf{x}^t - \mathbf{c}^t). \end{aligned} \tag{48}$$

A trivial solution to eq. (48) is $\mathbf{c}^t = -\zeta$ whose ith element is given by

$$c_{n,i}^t \;=\; -B \sin\left(2\pi \frac{i-1}{N}\right) \;=\; -1.6 \sin\left(2\pi \frac{i-1}{N}\right), \tag{49}$$

where we take $B = 0.2\Theta = 1.6$ in this experiment.

The time-averaged rms analysis errors for each case are shown in Fig. 12. We see that, when we augment the state using Bias Model II in the assimilation, we can correct for the bias. The minimum rms error for this assimilation is about 0.061 and occurs near $\mu = 0.003$. The minimum rms error for the assimilation with the augmented state using the Bias Model III is about 0.062 and occurs at a lower μ value (as expected) of about $\mu = 2 \times 10^{-5}$. Without bias correction, the minimum rms error is 0.263 and occurs near $\mu = 4.0$, similar to what is obtained using Bias Model I ($\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.234$ near $\mu = 0.5$). Also, due to the variance inflation effect of Bias Model I, the best assimilation result

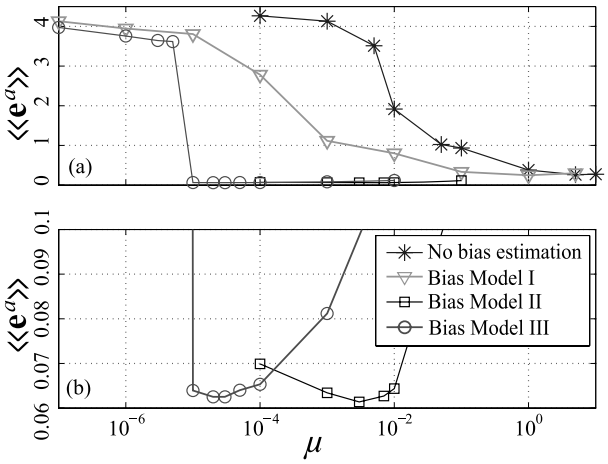

Fig. 12. Time-averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, versus μ for the case of Type B truth bias. (b) has a different vertical scale from (a) for transparency.

with the augmented state using the Bias Model III occurs at a lower value of variance inflation than the assimilation with the augmented state using Bias Model II. For the same reason, the rms error, $\langle\langle \mathbf{e}^a \rangle\rangle$, for the case of assimilation with the augmented state using Bias Model I has lower values in the region of variance inflation, $10^{-5} \leq \mu \leq 1.0$, than for the case of assimilation with no state augmentation.

We show $\langle \bar{\mathbf{c}}^a \rangle$ in Fig. 13 for the case in which the state is augmented using Bias Model II at $\mu = 0.003$. It is seen that the

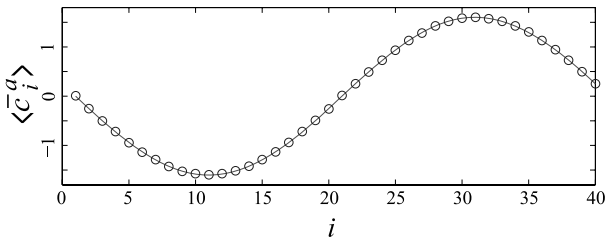

Fig. 13. Bias estimate for the Bias Model II is shown as $\circ$. The true bias (49) is shown as the solid curve.

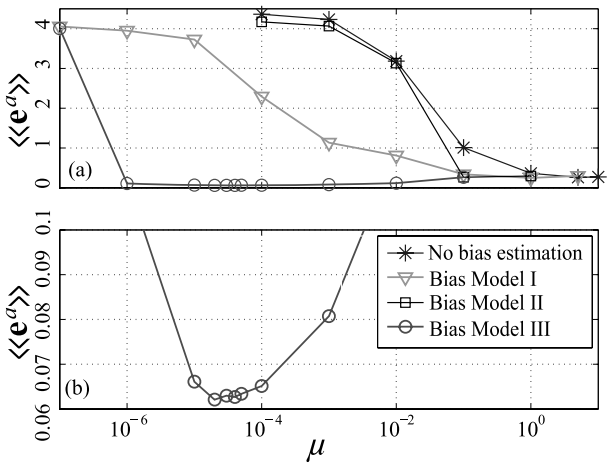

Fig. 14. Time-averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, versus μ for the case in which the truth has Type C truth bias: In order to correct for the biases, the augmented state used in the assimilations must contain both the **b** and **c** bias estimates, Here, we again use a different vertical scale in (b) from (a) for transparency.

result (plotted as ○) agrees very well with eq. (49) (plotted as the solid line). We obtain the same result for $\langle \bar{\mathbf{c}}^a \rangle$ when the state is augmented using Bias Model III at $\mu = 2 \times 10^{-5}$.

3.5. Data assimilation with Type C truth bias

Now, we combine the two biases in the truth [see eq. (29)], and estimate them with three augmented states as done in previous sections. In this case, we can regard the Type A truth bias as added to a system that already has Type B truth bias. Hence, a differential equation for $\mathbf{b}^t$ can be written as

$$\frac{\mathrm{d}\mathbf{b}^t}{\mathrm{d}t} = \mathbf{L}(\mathbf{x}^t + \boldsymbol{\zeta}) + \boldsymbol{\beta} - \mathbf{L}(\mathbf{x}^t - \mathbf{c}^t), \tag{50}$$

where $\boldsymbol{\zeta}$ and $\boldsymbol{\beta}$ are constant in time in the present experiment. A solution to eq. (50) is

$$\frac{\mathrm{d}\mathbf{b}^t}{\mathrm{d}t} = \boldsymbol{\beta} \quad \text{if} \quad \mathbf{c}^t = -\boldsymbol{\zeta}; \tag{51}$$

that is, the individual true bias is the same as if the system has only one bias. The quantities $\boldsymbol{\beta}$, $\boldsymbol{\zeta}$, are thus again given by eqs. (43) and (49).

Figure 14 shows the resulting time-averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, for each estimation method. The best result is obtained using Bias Model III. This might be anticipated since augmentation by one bias estimate alone cannot satisfy the solution (51). The minimum rms error is around 0.062 and occurs near 2×10^{-5}; this is the same as in the previous experiments. The other assimilation methods yields $\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.261$ at $\mu \approx 4.0$ without state augmentation, $\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.236$ at $\mu \approx 0.5$ using Bias Model I, and $\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.224$ at $\mu \approx 0.3$ using Bias Model II.

3.6. Settling time

Even though the state augmented LEKF can correct for various biases in the true state, for the truth biases considered here it requires longer settling time for the forecast model state to converge toward the true state as compared to the LEKF without state augmentation. We regard the time it takes for the rms analysis error (34) to settle near its time-averaged rms analysis error (35) as the settling time. With a perfect forecast model as in Section 3.2, the settling time is around 50 time steps with the $k + 1 = 13$ ensemble members we are using for the regular (i.e. without state augmentation) LEKF assimilation scheme. The settling time, however, becomes between 100 and 200 time steps when either Bias Model I or Bias Model II are used in the assimilation to respectively correct for Type A or B truth bias. The longest settling time, which is near 15 000 time steps, appears when assimilations are done using Bias Model III to correct for Type B and Type C truth biases (when using Bias Model III to correct for Type A truth bias, we found settling times that were generally below 500).

However, it turns out that we can easily correct this problem by using a priori information on the bias which could be obtained by looking at the innovation, the difference between the forecast and the observation,

$$\bar{\mathbf{d}}_n = \mathbf{H}(\bar{\mathbf{x}}_n^b) - \mathbf{y}_n, \tag{52}$$

for the case when no bias estimation is performed. We plot the time-average of the innovation in Fig. 15 for the case that no bias estimations are performed even though Type A, B or C truth bias is present in the truth. Since we take the observations to be unbiased, we can think of the time averaged innovation as the forecast bias. We see that these averages are large and vary slowly in space.

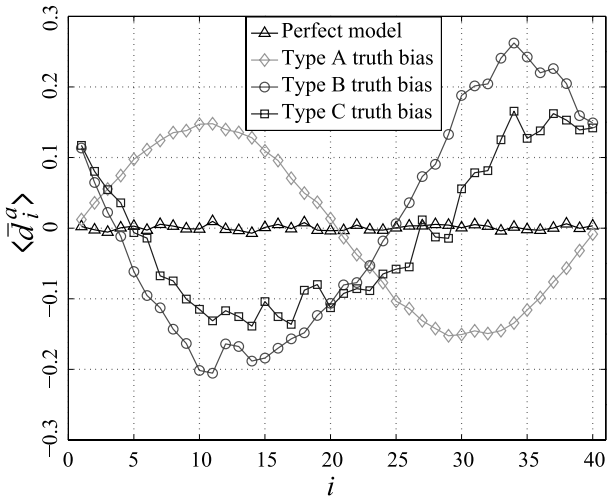

Fig. 15. Time-averaged innovation for the case that no bias estimation is performed with various biases in the truth.

In the previous experiments, the initial ensemble variance for the bias estimate is 0.1. By increasing the initial variance to 1.0 (somewhat larger than the spread in the time-averaged innovations), we can dramatically decrease the longest settling time from 15 000 time steps to 800 time steps, while for the case that the settling time is already small (between 100 and 200 time steps) no significant change in the settling time is observed. We can decrease the settling time further if we exploit the fact that the biases vary slowly in space. To incorporate this added knowledge into our data assimilation scheme we now use a diffusion process for the time evolution, eqs. (8) and (16), of the biases,

$$b^b_{n+1,i} = (1-2\alpha_b)b^a_{n,i} + \alpha_b b^a_{n,i-1} + \alpha_b b^a_{n,i+1}, \tag{53}$$

$$c^b_{n+1,i} = (1-2\alpha_c)c^a_{n,i} + \alpha_c c^a_{n,i-1} + \alpha_c c^a_{n,i+1}, \tag{54}$$

where α_b and α_c are diffusion coefficients. By introducing diffusion in this way, rapid spatial variation of the bias estimates is damped, leading to smooth spatial variation consistent with the actual case, eqs. (30) and (31), and the evidence of Fig. 15.

Our experiments show that there is a modest improvement in the settling time in the case that there is one type of bias in the truth (truth bias A or B) and that it is corrected using the corresponding bias model (Bias Model I or II, respectively); the settling time is decreased from between 100 and 200 time steps to between 80 and 130 time steps when α_b (or α_c) is increased from zero to $\alpha_b = 0.01$ (or $\alpha_c = 0.01$). When we consider the case of Type C truth bias and augment the state using Bias Model III, non-zero α_b diffusion (with $\alpha_c = 0$) can achieve a large decrease in the settling time, from 800 time steps to 300 time steps.

All of the decreases in settling time we have described come without significant increase in the time-averaged rms analysis error $\langle\langle \mathbf{e}^a \rangle\rangle$. We find that using diffusive evolution on bias estimates actually decreases $\langle\langle \mathbf{e}^a \rangle\rangle$ in some cases. Figure 16 shows the time asymptotic analysis errors as a function of α_b for the case of Type C truth bias and Bias Model III assimilation. We see that a small amount of diffusion, in addition to shortening the settling time, also improves the time asymptotic performance of the assimilation. This improvement is not seen in the experiments with Bias Model I and II.

Finally, we note that if diffusion [eqs. (53) and (54)] is added in the perfect forecast model case (Section 3.2), all three cases of augmentation have the same values of rms analysis error as that of the unaugmented case. That is, curves corresponding to the four cases in Fig. 6 have the same minimum values with appropriate amounts of diffusion. Evidently diffused evolution of the bias estimate allows the estimate to converge to the truth faster, and also reduces the rms analysis error of the state augmented estimates.

3.7. A simple state-dependent model error

In this section, we introduce a simple model error, γx_i^2, which is proportional to the square of the state variable. That is, the

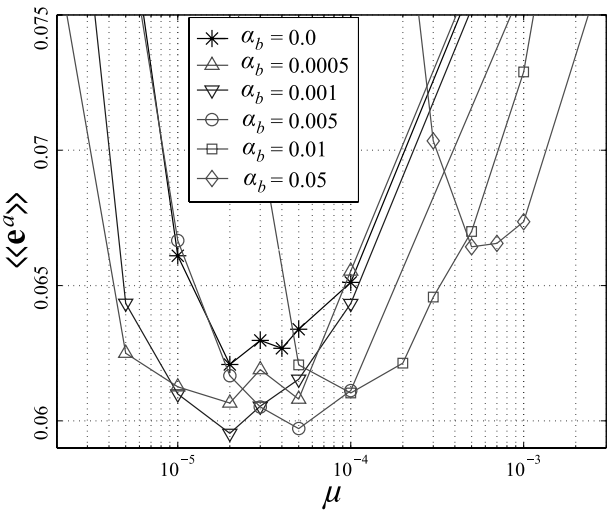

Fig. 16. Time-averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, versus μ with various α_b and $\alpha_c = 0$ for the case of Type C truth bias corrected using Bias Model III: Small diffusion improves the performance of the assimilation up to $\alpha_b = 0.01$.

dynamic equation for the true system is as follows:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = (x_{i+1} - x_{i-2})x_{i-1} - x_i - \gamma x_i^2 + \Theta, \quad i = 1, \ldots, N, \tag{55}$$

where $\gamma = 0.05$ and $\Theta = 10.0$. Here, Θ is increased to maintain chaotic behavior of the true dynamics (introducing γx_i^2 without changing Θ from its previous value of $\Theta = 8.0$ results in dominance of time periodic behavior of the true dynamics). For the forecast model, we use eq. (25) with $\Theta = 10.0$. Through the numerical experiment, we obtain the time average $\langle -\gamma x_i^2 \rangle \approx -0.95$ at each point i while $-\gamma x_i^2$ itself has large temporal fluctuations ranging from 0 to around -10. The task we undertake here is to successfully estimate the time mean effect of the model error, $\langle -\gamma x_i^2 \rangle$, with our bias estimation schemes. Using analysis similar to that in eqs. (41) and (42), we obtain

$$\left\langle b_i^t \right\rangle \approx \left\langle -\gamma \int_{t_{n-1}}^{t_n} (x_i^t)^2 \, \mathrm{d}t \right\rangle \approx -\gamma \left\langle \left(x_i^t\right)^2 \right\rangle \Delta t, \tag{56}$$

and hence $\langle b_i^t \rangle \approx -0.048$ for each location i.

In Fig. 17, we plot the rms analysis errors of the numerical experiments. Without bias estimation the minimum rms analysis error $\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.186$ is obtained at $\mu \approx 1.0$. Among the three bias models, Bias Model I produces the best result, $\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.171$ at $\mu \approx 0.05$. In terms of the difference with the error in the perfect model case ($\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.057$), this represents a 12% improvement toward the perfect model performance. If we employ diffusive evolution (with $\alpha_b = 0.05$), we can further decrease the rms error to obtain $\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.163$ at $\mu \approx 0.09$ and 18% improvement toward the perfect model performance. In both cases (with and without diffusion), we obtain $\langle \bar{b}_i^a \rangle \approx -0.05$, which is a good

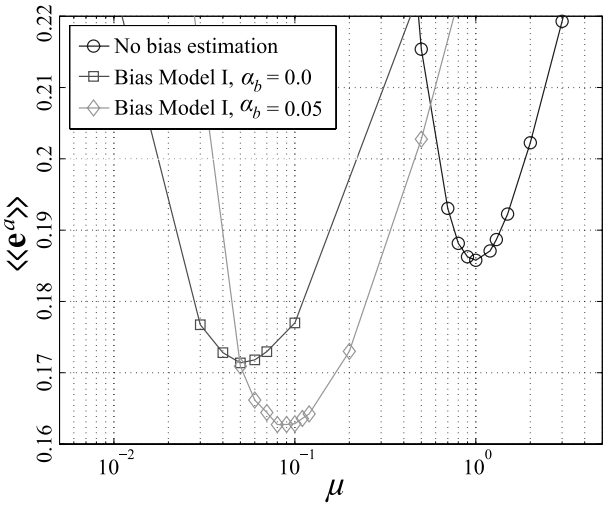

Fig. 17. Time-averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, versus μ with the simple state-dependent model error: With Bias Model I and diffusion process, we can improve the performance in terms of the rms analysis error.

estimate of $\langle b_i^t \rangle$, with smaller spatial variations when we use diffusion.

We also find that the performance of the assimilation with Bias Model I is not sensitive to the selection of Θ in the model equation while the performance of the assimilation without bias estimation is sensitive to the selection of Θ. We can also achieve the same performance with Bias Model III with appropriate diffusion ($\alpha_b = 0.05, \alpha_c = 0.2$), but we cannot achieve it with Bias Model II. We conjecture that the reason is because the form of the bias in eq. (55) is closer to Type A truth bias than Type B.

4. Conclusions and discussion

In this paper, we considered three bias models for use in state space augmentation strategies to mitigate the effects of model biases on forecasts.

(i) Bias Model I is based on the assumption that the best background information is obtained when the initial condition of the short-term forecast that provides the background (the analysis at the previous assimilation time) is as close to the truth as possible.

(ii) Bias Model II is based on the assumption that there exists a transformation from orbits on the attractor of the forecast model to orbits on the attractor of the true system.

(iii) Bias Model III combines Bias Model I and Bias Model II.

While Bias Model I was considered by others in earlier papers for schemes other than the LEKF, Bias Model II and Bias Model III are (to the best of our knowledge) first introduced in the present paper.

To evaluate the performance of the proposed bias models for use in augmented ensemble Kalman filtering, we carried out experiments with the Lorenz-96 model. While we used the original model equation of Lorenz and Emanuel (1998) to evolve the model state, we employed altered versions of Lorenz and Emanuel's equation to generate sets of time-series of the true states. Each alteration of the equation corresponded to distinctly different types of model biases. The main results of these numerical experiments are the following.

(i) The effectiveness of the different bias models strongly depends on the actual form of the true model bias. In our numerical experiments it was found that when the bias model was suited to the bias of the forecast model in modeling the true dynamics, then good results were obtained. However, when this was not the case, the results were not improved by the model bias correction scheme. This suggests that serious consideration of the choice of the bias model may be crucial in obtaining a successful scheme for model bias mitigation.

(ii) For the bias models we considered, Bias Model III performed as well or better than the other bias models in terms of average analysis error, at the expense of requiring a larger ensemble and in some cases increasing the settling time. In most cases, the inclusion of parameters that were not present in the model bias did not yield improved performance. However, in the case of Type A truth bias, Bias Model III did outperform Bias Model I, due to the fact that the model bias was added to a continuous time forecast model, while the bias correction was applied at discrete times (see Section 3.3).

(iii) We found (Section 3.6) that the model bias correction scheme took many more iteration steps to converge than in the case in which no model biases are present. The settling time strongly depends on the actual model bias, and on the bias model employed. As a result of the possibility of long settling times, one might anticipate that use of these model correction schemes may become problematic. We can, however, dramatically reduce the settling time by increase of the initial ensemble spread of the bias estimates. Moreover, when the model bias is slowly varying in space, we demonstrated that choosing a diffusive evolution of the model bias can also reduce the settling time.

(iv) In a case with state-dependent, and thus time-varying, additive model error (Section 3.7), Bias Model I estimated a model bias that was close to the time average of the model error. In this sense, it found the best estimate of the model error within our constant-in-time parameterization. The improvement in performance compared to no bias estimation was modest but significant.

Finally, we note that state space augmentation is not the only way to account for the effect of model errors in the state estimation process. As we mentioned earlier, variance inflation (both additive and multiplicative) can improve the resilience of Kalman filter schemes to the effects of model errors. Promising results were achieved by employing additive variance inflation schemes

(Houtekamer et al., 2005; Hamill and Whitaker, 2005) and by using hybrid Ensemble/3DVAR schemes (Etherton and Bishop, 2004) first proposed by Hamill and Snyder (2000). Nevertheless, we believe that optimal parameterization of model errors in the data assimilation process is a very promising means of mitigating model error and that much remains to be done along these lines. We view our paper as a contribution only to the initial phase of the quest for efficient model error parameterization algorithms.

5. Acknowledgments

This work was supported by the Army Research Office, by a James S. McDonnell 21st Century Research Award, by a National Oceanic and Atmospheric Administration THORPEX grant, by the NPOESS Integrated Program Office (IPO), by the Office of Naval Research (Physics) and by the National Science Foundation (Grants nos. 0104087 and ATM 034225).

References

Anderson, J. L. 2001. An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.* **129**, 2884–2903.

Annan, J. D. and Hargreaves, J. C. 2004. Efficient parameter estimation for a highly chaotic system. *Tellus* **56A**, 520–526.

Baer, F. and Tribbia, J. J. 1977. On complete filtering of gravity modes through nonlinear initialization. *Mon. Wea. Rev.* **105**, 1536–1539.

Bell, M. J., Martin, M. J. and Nichols, N. K. 2004. Assimilation of data into an ocean model with systematic errors near the equator. *Quart. J. Roy. Meteor. Soc.* **130**, 873–893.

Bishop, C. H., Etherton, B. J. and Majumdar, S. J. 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.* **129**, 420–436.

Carton, J. A., Chepurin, G. and Cao, X. 2000. A simple ocean data assimilation analysis of the global upper ocean 1950–1995. Part I: Methodology. *J. Phys. Oceanogr.* **30**, 294–309.

Cohn, S. E. 1997. An introduction to estimation theory. *J. Meteor. Soc. Japan* **75**, 257–288.

Daley, R. 1992. The effect of serially correlated observation and model error on atmospheric data assimilation. *Mon. Wea. Rev.* **120**, 164–177.

Dee, D. P. and Da Silva, A. M. 1998. Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.* **124**, 269–295.

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

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

Etherton, B. J. and Bishop, C. H. 2004. Resilience of hybrid ensemble/3DVAR analysis schemes to model error and ensemble covariance error. *Mon. Wea. Rev.* **132**, 1065–1080.

Evensen, G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99C5**, 10 143–10 162.

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

Ghil, M., Cohn, S., Tavantzis J., Bube, K. and Isaacson, E. 1981. Applications of estimation theory to numerical weather prediction. In: *Dynamic Meteorology: Data Assimilation Methods* (eds. A. L. Bengtsson, M. Ghil, and E. Kallen) Springer-Verlag, New York, 139–224.

Hamill, T. M. and Snyder, C. 2000. A hybrid ensemble kalman filter-3D variational analysis scheme. *Men. Wea. Rev.* **128**, 2905–2919.

Hamill, T. M., Whitaker, J. S. and Snyder, C. 2001. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.* **129**, 2776–2790.

Hamill, T. M. and Whitaker, J. S. 2005. Accounting for the error due to unresolved scales in ensemble data assimilation: A comparison of different approaches. *Mon. Wea. Rev.* **133**, 3132–3147.

Hansen, J. A. 2002. Accounting for model error in ensemble-based state estimation and forecasting. *Mon. Wea. Rev.* **130**, 2373–2391.

Houtekamer, P. L. and Mitchell, H. L. 1998. Data assimilation using an ensemble kalman filter technique. *Mon. Wea. Rev.* **126**, 796–811.

Houtekamer, P. L. and Mitchell, H. L. 2001. A sequential ensemble kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.* **129**, 123–137.

Houtekamer, P. L., Mitchell, H. L., Pellerin, G., Buehner, M. and Charron, M., 2005. Atmospheric data assimilation with an ensemble kalman filter: Results with real observations. *Mon. Wea. Rev.* **133**, 604–620.

Jazwinski, A. H. 1970. *Stochastic Processes and Filtering Theory.* Academic Press, San Diego.

Lorenz, E. N. and Emanuel, K. A. 1998. Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.* **55**, 399–414.

Lynch, P. and Huang, X.-Y. 1992. Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.* **120**, 1019–1034.

Lynch, P. 1997. The Dolph-Chebyshev window: A simple optimal filter. *Mon. Wea. Rev.* **125**, 655–660.

Machenhauer, B. 1977. On the dynamics of gravity oscillations in a shallow water model with applications to normal mode initialization. *Contrib. Atmos. Phys.* **50**, 253–271.

Margolin, L. G., Titi, E. S. and Wynne, S. 2003. The postprocessing Galerkin and nonlinear Galerkin methods—A truncation analysis point of view. *SIAM J. Numer. Anal.* **41**, 695–714.

Martin, M. J., Bell, M. J. and Nichols, N. K. 2002. Estimation of systematic error in an equatorial ocean model using data assimilation. *Int. J. Numer. Meth. Fluids* **40**, 435–444.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V. and Kostelich, E. J. 2004a. A local ensemble Kalman filter for atmospheric data assimilation. *Tellus* **56A**, 415–428.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V. and Kostelich, E. J. 2004b. Estimating the state of large spatio-temporally chaotic systems. *Phys. Lett. A* **330**, 365–370.

Szunyogh, I., Kostelich, E. J., Gyarmati, G., Patil, D. J. and Hunt, B. R. 2005. Assessing a local ensemble Kalman filter: Perfect model experiments with the NCEP global model. *Tellus* **57A**, 528–545.

Whitaker, J. S. and Hamill, T. M. 2002. Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.* **130**, 1913–1924.

Whitaker, J. S., Compo, G. P., Wei, X. and Hamil T. M. 2004. Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.* **132**, 1190–1200.

Zupanski, D. 1997. A general weak constraint applicable to operational 4DVAR data assimilation systems. *Mon. Wea. Rev.* **125**, 2274–2292.