1 Case2

We will study two types of simple model error with Lorenz 96 model. We will introduce constant deterministic error to the standard Lorenz 96 model and try to use ensemble Kalman filter on an augmented system to improve the performance of data assimilation.

1.1 Notation

$$
\begin{aligned}
\text{True model:} \quad & m \\
\text{Forecast model:} \quad & \hat{m} \\
\text{True state:} \quad & \boldsymbol{x}_n \\
\text{Model state:} \quad & \hat{\boldsymbol{x}}_n \\
\text{Observation Operator:} \quad & H : \boldsymbol{x}_n \mapsto \boldsymbol{y}_n
\end{aligned}
$$

The Lorenz96 equation is:

$$
\dot{x}_i = (x_{i+1} - x_{i+2})x_{i-1} - x_i + F, \quad i = 1, 2, ..., N \tag{1}
$$

We will use the common choice of $N = 40$ and $F = 8$ with periodic boundary conditions. For readability, we write the whole system as:

$$
\dot{\boldsymbol{x}} = \boldsymbol{L}(\boldsymbol{x}) \tag{2}
$$

1.2 Types of error

We consider three types of model error, where the true systems are as follows:

$$
\begin{aligned}
\dot{\boldsymbol{x}} =& \boldsymbol{L}(\boldsymbol{x}) + \boldsymbol{\zeta} & \text{(Type \quad I)} & \tag{3} \\
\dot{\boldsymbol{x}} =& \boldsymbol{L}(\boldsymbol{x} + \boldsymbol{\xi}) & \text{(Type \quad II)} & \tag{4} \\
\dot{\boldsymbol{x}} =& \boldsymbol{L}(\boldsymbol{x} + \boldsymbol{\xi}) + \boldsymbol{\zeta} & \text{(Type \quad III)} & \tag{5}
\end{aligned}
$$

The first type error is a constant bias. The second type of error models the situation where the true dynamics and the dynamics used for forecasting are associated with a coordinate transformation. For simplicity, we assume the transformation is just a shift of the forecast state to the true state. We combine the two types of error to get the third type of truth bias.

For simplicity, we assume that $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$ in the equations above are both constant in time. Some have suggested that when the error is arguably small, they exhibit behaviors similar to those of the system. Accordingly, we set

$$
\begin{aligned}
\zeta_i =& A \sin(2\pi \frac{i-1}{N}) \tag{6} \\
\xi_i =& B \sin(2\pi \frac{i-1}{N}), \quad i = 1, ..., N \tag{7}
\end{aligned}
$$

where A and B are constants.

1.3 Model

The true model m used for generating the truth is integrating eq.(3,4,5), respectively in each setting, from time t_{n-1} to time t_n. The forecast model

$\hat{m}$ used for evolving the model state is integrating eq.(2) from time t_{n-1} to time t_n.

In the numerical experiments below, to avoid the potential influence of error introduced by numerical integration, we do not use a fixed step method for integration. Instead, we use an error correction method with a specified error tolerance of 10^{-7}. One common method of such type is fourth order Runge-Kutta method, with a fifth order Runge-Kutta method used for error estimation. In Matlab, it's ode45.

1.3.1 Model 1

To address the Type I model error, we first define the model bias:

$$
\boldsymbol{b}_n = m(\boldsymbol{x}_{n-1}) - \hat{m}(\boldsymbol{x}_{n-1}) \tag{8}
$$

We therefore incorporate $\boldsymbol{b}_n$ to the augmented system:

$$
\begin{aligned}
\boldsymbol{x}_n^f =& \hat{m}(\boldsymbol{x}_{n-1}^a) + \boldsymbol{b}_n^f \tag{9} \\
\boldsymbol{b}_n^f =& f_b(\boldsymbol{x}_{n-1}^a, \boldsymbol{b}_n^a) \tag{10}
\end{aligned}
$$

In the case where $\boldsymbol{\zeta}$ is constant, we assume $\boldsymbol{b}_n$ is constant too, and then

$$
\boldsymbol{b}_n^f = f_b(\boldsymbol{x}_{n-1}^a, \boldsymbol{b}_{n-1}^a) = \boldsymbol{b}_{n-1}^a \tag{11}
$$

To see how this model associates with the Type I error, we suppose the model state and true state are close, i.e. the bias is small, and then take the difference between the true equation (3) and the model equation (2),

$$
\frac{d}{dt}(\boldsymbol{x} - \hat{\boldsymbol{x}}) = \boldsymbol{L}(\boldsymbol{x}) + \boldsymbol{\zeta} - \boldsymbol{L}(\hat{\boldsymbol{x}}) \approx \boldsymbol{\zeta} \tag{12}
$$

then, integrating eq.(12) for the time interval $t_{n-1} \le t \le t_n$ with the initial condition $\boldsymbol{x}(t_{n-1}) = \hat{\boldsymbol{x}}(t_{n-1}) = \boldsymbol{x}_{n-1}$, to obtain the following

$$
\boldsymbol{b}_n = \boldsymbol{x}(t_n) - \hat{\boldsymbol{x}}(t_n) \approx \int_{t_{n-1}}^{t_n} \boldsymbol{\zeta} dt = \boldsymbol{\zeta} \Delta t \tag{13}
$$

Since $\boldsymbol{\zeta}$ is constant in time, if the data assimilation scheme works, then the augmented state $pmbb_n$ should converge to a value that is very close to $\boldsymbol{\zeta} \Delta t$. In the numerical experiments below, we will therefore compare time averaged $pmbb_n$ and $\boldsymbol{\zeta} \Delta t$ to see if the bias estimation is successful.

1.3.2 Model 2

For the second type of error, we use the analysis to model $\hat{\boldsymbol{x}}_n$ instead of the true state $\boldsymbol{x}_n$. We define the bias as

$$
\boldsymbol{c}_n = m(\boldsymbol{x}_{n-1}^t) - \hat{m}(\boldsymbol{x}_{n-1}^m) = m(\boldsymbol{x}_{n-1}^t) - \hat{m}(\boldsymbol{x}_{n-1}^t - \boldsymbol{c}_{n-1}) \tag{14}
$$

The augmented system then becomes

$$\boldsymbol{x}_n^f = \hat{m}(\boldsymbol{x}_{n-1}^a) \tag{15}$$

$$\boldsymbol{c}_n^f = f_c(\boldsymbol{x}_{n-1}^a, \boldsymbol{c}_n^a) \tag{16}$$

For simplicity, we assume we know $\boldsymbol{\xi}$ is constant, which means $\boldsymbol{c}_n$ should be constant, and then

$$\boldsymbol{c}_n^f = f_c(\boldsymbol{x}_{n-1}^a, \boldsymbol{c}_{n-1}^a) = \boldsymbol{c}_{n-1}^a \tag{17}$$

Note that in such setting, the analysis will model the model trajectory $\hat{\boldsymbol{x}}_n$ instead of the true trajectory $\boldsymbol{x}_n$, and therefore the observation should be:

$$\boldsymbol{y}_n = H(\hat{\boldsymbol{x}}_n + \boldsymbol{c}_n) \tag{18}$$

To see how the augmented state $\boldsymbol{c}$ relates to the model error $\boldsymbol{\xi}$, we differentiate $\boldsymbol{c}$ to get

$$\begin{aligned} \frac{d\boldsymbol{c}}{dt} &= \frac{d\boldsymbol{x}}{dt} - \frac{d\hat{\boldsymbol{x}}}{dt} \\ &= \boldsymbol{L}(\boldsymbol{x} + \boldsymbol{\xi}) - \boldsymbol{L}(\hat{\boldsymbol{x}}) \\ &= \boldsymbol{L}(\boldsymbol{x} + \boldsymbol{\xi}) - \boldsymbol{L}(\boldsymbol{x} - \boldsymbol{c}) \end{aligned} \tag{19}$$

Apparently, $\boldsymbol{c} = -\boldsymbol{\xi}$ solves the equation, and therefore it's reasonable to believe if our data assimilation works, $\boldsymbol{c}$ will converge to $-\boldsymbol{\xi}$. In the numerical experiments below, we will thus compare a time averaged $\boldsymbol{c}$ and $-\boldsymbol{\xi}$ to see if the bias estimation is successful.

1.3.3 Model 3

When we combine the two types of the model errors and try to address them together, we have the following scheme:

$$\boldsymbol{x}_n^f = \hat{m}(\boldsymbol{x}_{n-1}^a) + \boldsymbol{b}_n^f \tag{20}$$

$$\boldsymbol{b}_n^f = f_b(\boldsymbol{x}_{n-1}^a, \boldsymbol{b}_{n-1}^a, \boldsymbol{c}_{n-1}^a) \tag{21}$$

$$\boldsymbol{c}_n^f = f_c(\boldsymbol{x}_{n-1}^a, \boldsymbol{b}_{n-1}^a, \boldsymbol{c}_{n-1}^a) \tag{22}$$

Again, for simplicity, we assume c_n and b_n are constant, i.e.,

$$\boldsymbol{b}_n^f = f_b(\boldsymbol{x}_{n-1}^a, \boldsymbol{b}_{n-1}^a) = \boldsymbol{b}_{n-1}^a \tag{23}$$

$$\boldsymbol{c}_n^f = f_c(\boldsymbol{x}_{n-1}^a, \boldsymbol{c}_{n-1}^a) = \boldsymbol{c}_{n-1}^a \tag{24}$$

Same as in Model 2, the observation operator becomes

$$\boldsymbol{y}_n = H(\hat{\boldsymbol{x}}_n + \boldsymbol{c}_n) \tag{25}$$

Since model 3 is simply a combination of model 1 and 2, it's natural to reason that $\boldsymbol{b} \approx \boldsymbol{\zeta}\Delta t$ and $\boldsymbol{c} = -\boldsymbol{\xi}$ still hold. Indeed, if we differentiate $\boldsymbol{b}$, we have

$$\frac{d\boldsymbol{b}}{dt} = \boldsymbol{L}(\boldsymbol{x} + \boldsymbol{\xi}) + \boldsymbol{\zeta} - \boldsymbol{L}(\boldsymbol{x} - \boldsymbol{c}) \tag{26}$$

which gives us solution

$$\frac{d\boldsymbol{b}}{dt} = \boldsymbol{\zeta} \Rightarrow \boldsymbol{b} = \boldsymbol{\zeta}\Delta t \tag{27}$$

$$\boldsymbol{c} = -\boldsymbol{\xi} \tag{28}$$

which is consistent with our reasoning for model 1 and 2.

1.4 Numerical Experiments

1.4.1 Experiment Setup

In each setting of error, we will use ensemble Kalman Filter with perfect model assumption, i.e. without an augmented system, and with Model 1, 2, and 3, to estimate the solution. We will compare the performance of these different methods.

In order for ensemble Kalman filter to work, we will introduce uncorrelated Gaussian noise with zero mean and $\sigma_{obs} = 0.09$ at each step to the observation operator.

We will use time-averaged root-mean-square of the analysis error to measure the performance of the different methods, where the analysis error of Model 1 is defined as

$$\bar{\boldsymbol{e}}_a^{(n)} = \bar{\boldsymbol{x}}_n^a - \boldsymbol{x}_n, \tag{29}$$

and the analysis error of Model 2 and 3 is

$$\bar{\boldsymbol{e}}_a^{(n)} = \bar{\boldsymbol{x}}_n^a + \bar{\boldsymbol{c}}_n - \boldsymbol{x}_n, \tag{30}$$

where $\bar{\boldsymbol{x}}_n^a$ is the ensemble mean of the analysis at each step and $\bar{\boldsymbol{c}}_n$ the ensemble mean of the bias $\boldsymbol{c}_n$. And root-mean-square of the analysis error is denoted as $\langle \bar{\boldsymbol{e}}_a^{(n)} \rangle$, the time-average $\langle\langle \bar{\boldsymbol{e}}_a^{(n)} \rangle\rangle$.

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

$$\boldsymbol{P}_n^a \rightarrow \boldsymbol{P}_n^a + \frac{\mu\Lambda}{K}\boldsymbol{I}_K \tag{31}$$

where $\boldsymbol{P}_n^a$ is the analysis error covariance matrix, μ is an inflation coefficient, whose value is critical to the performance of data assimilation, which we will discuss below, $\Lambda = \text{Trace}(\boldsymbol{P}_n^a)$, K is the dimension of the augmented space for each model, and $\boldsymbol{I}_k$ is the K dimensional identity matrix.

1.4.2 Effect of Randomness

The experiments are in some sense not deterministic, since, first, the observations are generated with white noise, second, we introduce randomness in the initial state of the ensemble and we randomly perturb the observations at each step. Some have suggested using averaged results over trials with various random seeds instead of results of an individual trial to avoid the possibility that, in some extreme cases, the randomness might lead to abnormal behavior of the results. Nevertheless, after many experiments with different random seeds and everything else fixed, we found no significant differences worth mentioning in the results. Therefore we will simply use the results of a single trial, in each of the settings below.
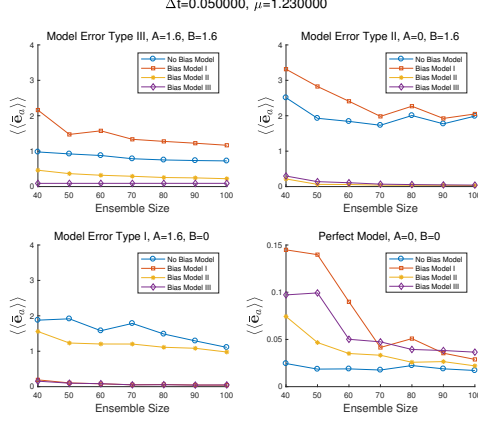
2

Figure 1: Time-averaged rms analysis error, $\langle\langle \bar{\boldsymbol{e}}_a \rangle\rangle$, versus ensemble size M, for the different cases of model error, each with all 4 models.

Figure 2: Time-averaged rms analysis error, $\langle\langle \bar{\boldsymbol{e}}_a \rangle\rangle$, versus step size between observations Δt, for the different cases of model error, each with all 4 models.

1.4.3 Effect of Ensemble Size

Considering the dimension of the problem $N = 40$, to also use $M = 40$ as ensemble size seems a rational choice. On the one hand, it's natural to assume the reasonable choice of ensemble size is proportional to the dimension of the problem. On the other hand, to make the ensemble size larger than or equal to the dimension of the system largely avoids the rank deficiency of the covariance matrix. After some experiments, we found that, in each setting of model error, for the model that best captures the model error, the ensemble size does not significantly affect the performance of data assimilation. Meanwhile, for the models that do not capture the model error well and the model without bias estimation, increased ensemble size might improve the performance of data assimilation. Fig.(1) demonstrates the relation between the performance of data assimilation and the choice of ensemble size.

Some research has suggested choosing different ensemble sizes for each of the models. For example, choose ensemble size proportional to the dimension of the augmented space, i.e. M for perfect model, $2M$ for model 1 and 2, and $3M$ for model 3. However, such approach would potentially mix the effects of successful bias estimation with increased ensemble size. Also, since ensemble size is not a significant factor in increasing performance, we will choose an ensemble size of $M = 40$ for all the models in the experiments below.

1.4.4 Effect of Step Size

Many sources have suggested using $\Delta t = 0.05$ as the step size between two observations. Nevertheless, it's interesting to inspect how the step size Δt influence the performance of data assimilation. In-
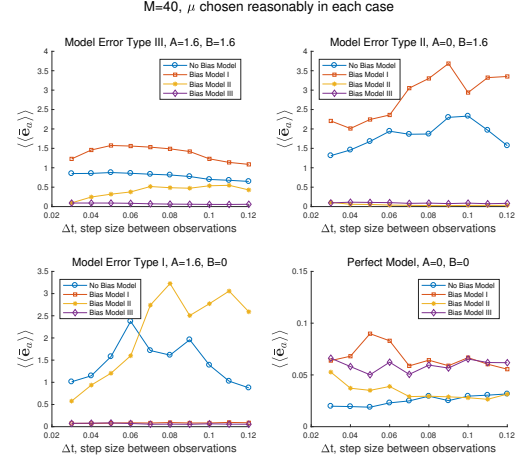
tuitively speaking, the more frequent we make the observations, the better the performance should be. As shown in Fig.(2), the performance is not significantly affected by step size, similar to the case of ensemble size discussed above. We can also see that in some cases, the increase in step size even make the data assimilation worse.

Therefore, we will simply use the suggested $\Delta t = 0.05$ for the experiments below.

1.4.5 Effect of Inflation Coefficient

Throughout the experiment, we realize that the most significant factor affecting the performance appears to be the inflation coefficient μ, as defined in eq.(31). In each of the model error settings, for each of the models, there seems to be an optimal value for μ to achieve the best performance. Moreover, the relation between performance and the inflation coefficient seems complicated. The optimal values of μ in each situation for each model are quite different. In some cases, a poorly chosen μ could make the better model perform worse than other models. Therefore, in the discussion below, we will mainly focus on the relation between the variance inflation parameter μ and the time averaged rms error $\langle\langle \bar{\boldsymbol{e}}_a \rangle\rangle$, as a measure of the performance of data assimilation.

1.4.6 Perfect Model

We first test our bias models for the case of a perfect forecast model, i.e. for the case when $A = 0$ in eq.(6) and $B = 0$ in eq.(7), and the true model used for evolving the true state is simply eq.(2), the same as the model used for forecast.
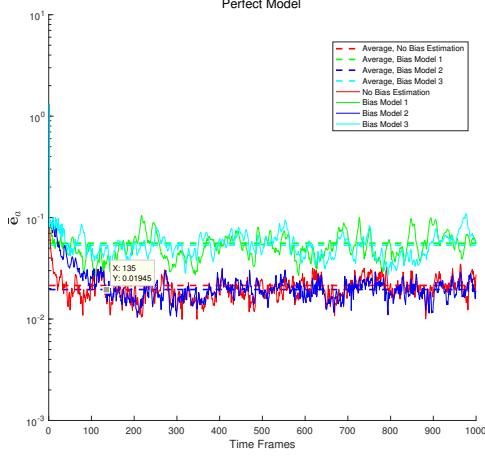
3

Figure 3: Time-averaged rms analysis error, $\langle\langle \bar{\boldsymbol{e}}_a \rangle\rangle$, and rms analysis error at each time step $\bar{\boldsymbol{e}}_a^{(n)}$, for the case without model error, i.e. $A = 0, B = 0$, with all 4 models.

Settling time of error Since the perfect model is, in some sense, the simplest model, it's natural to assume that the rms analysis error will stay around a constant value, after a short period of time. Fig.(3) shows that that is indeed the case. The error settles down with roughly 100 time steps for all four models. We then use 200 as number of time steps in the experiment below, and whenever we refer to "time-averaged" we mean time averaged over time step 100 to 200, a rather stable period for the error.

Performance with μ As mentioned earlier in section 1.4.5, the most significant factor affecting the performance is the variance inflation coefficient μ. Fig.(4) shows the relation between error and μ. We can see that, generally speaking, the simple model performs best. This is reasonable, since all other models, in some sense, introduce unnecessary parameters, and the model with the the most unnecessary parameters, i.e. model 3, performs the worst. One interesting phenomena is that for the bias models, the optimal values for μ are roughly speaking the same, i.e. $\mu = 0.378$, and if μ is larger than this value, the error will increase remarkably.

Bias estimation In this case, the "biases" are simply zero. In Fig.(5), we can see that the bias models fail to capture this trivial "bias" very well. Especially for the estimation of $\boldsymbol{c}$, model 2 roughly stays around zero, which is the true value, but model 3 exhibits some unwanted oscillating behavior.
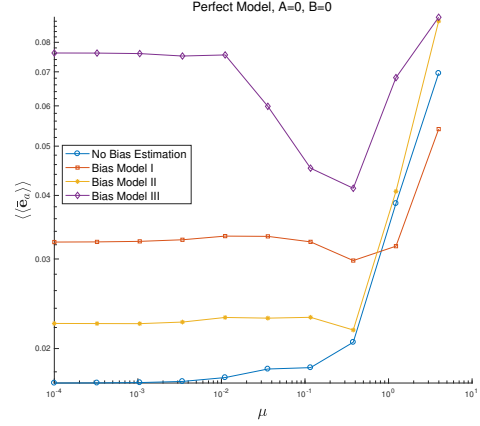

Figure 4: Time-averaged rms analysis error, $\langle\langle \bar{\boldsymbol{e}}_a \rangle\rangle$, versus the variance inflation parameter μ, for the case without model error, i.e. $A = 0, B = 0$, with all 4 models.
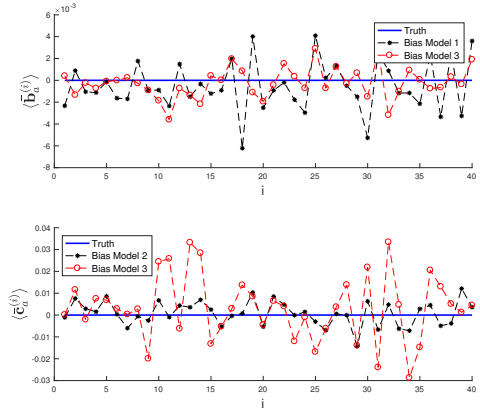

Figure 5: Time-averaged bias estimate at each location for the case without model error, i.e. $A = 0, B = 0$. The truth $\zeta_i \Delta t$ and $-\xi_i$ are shown as the solid curve.
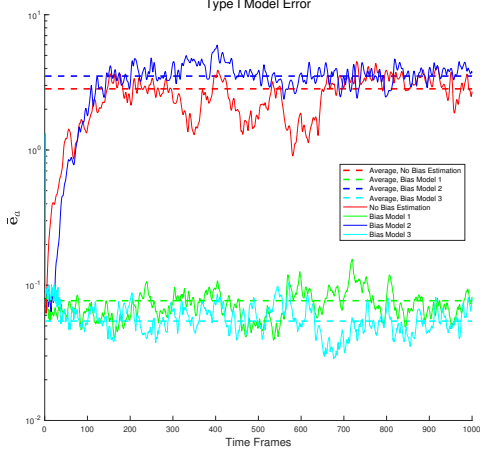
4

Figure 6: Time-averaged rms analysis error, $\langle\langle\bar{e}_a\rangle\rangle$, and rms analysis error at each time step $\bar{e}_a^{(n)}$, for the Type I model error, with all 4 models.

1.4.7 Type I Model Bias

In this experiment, we perform data assimilation using the model without bias estimation with unaugmented states i.e. eq.(2), and the three models with augmented space defined in subsection 1.3, when the true state is evolved with Type I model error, defined in eq.(3), with $A = 0.2F = 1.6$ in eq.(6).

Settling time of error The behavior of error with respect to time in this case is quite different from that of a perfect mode. Fig.(6) shows that the error for model 1 and 3 converges rather rapidly. However, the errors for the simple model and model 2 are initially small, and then it takes a long time for the errors to stay stable around a value that is considerably large. One way to explain this is that the way model 2 is formulated does not account for this type of model error, and therefore it will not behave very well. We will use 200 as number of time steps in the experiment below, and time step 100 to 200 as the period for time averaging.

Performance with μ Fig.(7) shows the relation between error and μ. Just as we expect, the models that are formulated to account for the Type I model bias perform better than those that do not account for this type of model bias. Model 1 performs the best, and model 3 performs a little bit worse, since model 3 introduces unnecessary parameters $\boldsymbol{b}$ in this case. We can see that for model 1 and 3, the value of μ does not affect the performance much, especially for model 1. However, for the other two models, larger value of μ significantly improves the performance, and when $\mu = 4$, the model without bias estimation even beats model 3.
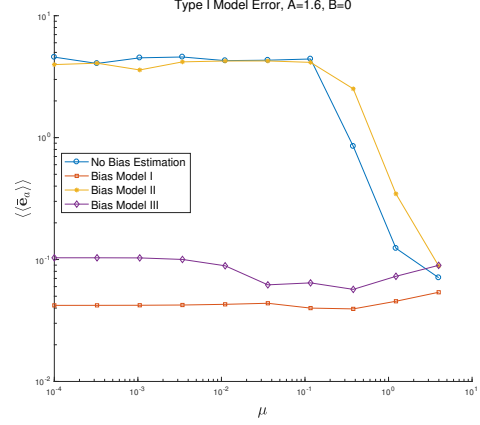

Figure 7: Time-averaged rms analysis error, $\langle\langle\bar{e}_a\rangle\rangle$, versus the variance inflation parameter μ, for the Type I model error, with all 4 models.
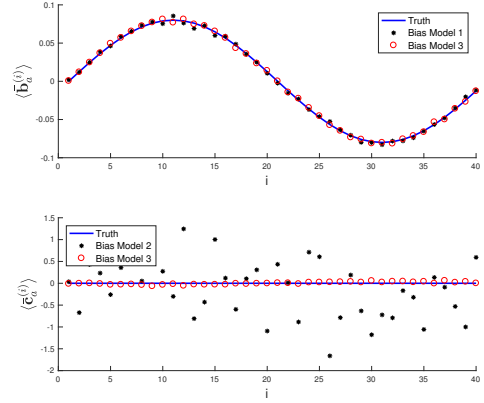

Figure 8: Time-averaged bias estimate at each location for the case with Type I model error. The truth $\zeta_i\Delta t$ and $-\xi_i$ are shown as the solid curve.

Bias estimation Fig.(8) shows the results for bias estimation. Both model 1 and 3 capture the bias very well, with very few noticeable deviance from the truth. Moreover, unlike the case of perfect model, the estimation for the zero "bias" is successful for model 3, although miserable for model 2. This is very interesting. Heuristically speaking, model 3 is closer to the truth than model 2, and therefore estimates the zero bias better.

1.4.8 Type II Model Bias

In this experiment, we perform data assimilation using the model without bias estimation with unaugmented states i.e. eq.(2), and the three models with augmented space defined in subsection 1.3, when the true state is evolved with Type II model error, defined in eq.(4), with $B = 0.2F = 1.6$ in eq.(7).
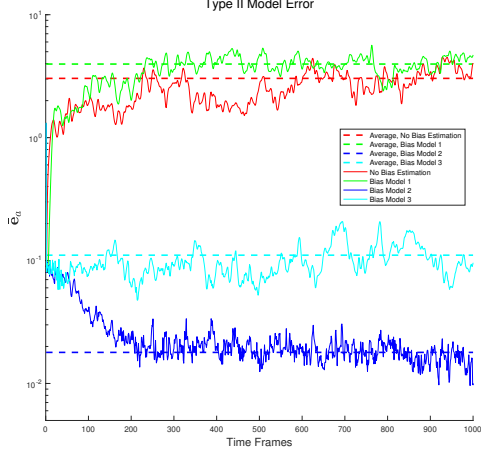
5

Figure 9: Time-averaged rms analysis error, $\langle\langle\bar{\boldsymbol{e}}_a\rangle\rangle$, and rms analysis error at each time step $\bar{\boldsymbol{e}}_a^{(n)}$, for the Type II model error, with all 4 models.

Figure 10: Time-averaged rms analysis error, $\langle\langle\bar{\boldsymbol{e}}_a\rangle\rangle$, versus the variance inflation parameter μ, for the Type II model error, with all 4 models.

Settling time of error The behavior of error with respect to time in this case is similar but still different from that of Type I model bias. Fig.(9) shows that the error for model 3 converges rather rapidly. However, the error for model 2, which best captures the dynamics, stabilizes after a rather long time. Some has suggested using a larger initial spread will reduce the settling time, but the improvement is small. We will accordingly use 400 as number of time steps in the experiment below, and time step 200 to 400 as the period for time averaging.

Performance with μ Fig.(10) shows the relation between error and μ. Similar to the case of Type I model bias, the models that are formulated to account for the Type II model bias perform better than those that do not account for this type of model bias. Model 2 performs the best. Model 3 performs worse than model 2, but still significantly better than the others, probably since model 3 introduces unnecessary parameters $\boldsymbol{c}$ in this case. We can see that for model 2, we can safely say there exist an optimal value for μ around 0.4, which is very similar to the case in the perfect model. For the other models, larger value of μ seems to improve the performance, especially for the perfect model and bias model 1.

Bias estimation Fig.(11) shows the results for bias estimation. The analysis of the results is basically the same as that of Type I model bias. Both model 2 and 3 capture the bias very well, with very few noticeable deviance from the truth. And model 3 also provides good estimation for the zero "bias", whereas model 1 fails to do so.
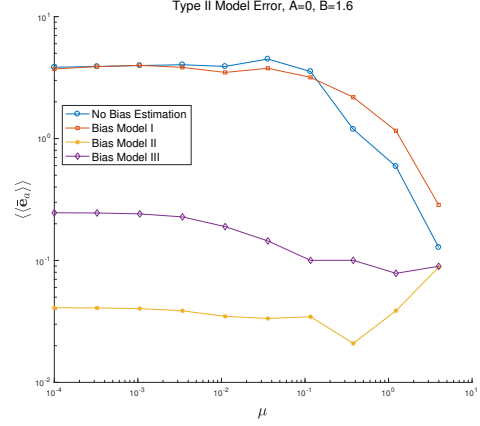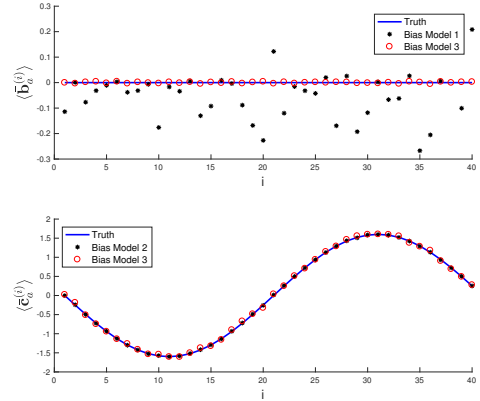

Figure 11: Time-averaged bias estimate at each location for the case with Type II model error. The truth $\zeta_i\Delta t$ and $-\xi_i$ are shown as the solid curve.
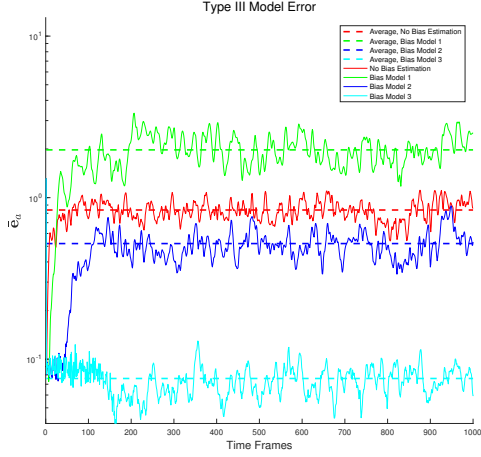
Figure 12: Time-averaged rms analysis error, $\langle\langle \bar{\boldsymbol{e}}_a \rangle\rangle$, and rms analysis error at each time step $\bar{\boldsymbol{e}}_a^{(n)}$, for the Type III model error, with all 4 models.

1.4.9 Type III Model Bias

In this experiment, we perform data assimilation using the model without bias estimation with unaugmented states i.e. eq.(2), and the three models with augmented space defined in subsection 1.3, when the true state is evolved with Type I model error, defined in eq.(5), with $A = 0.2F = 1.6$ in eq.(6) and $B = 0.2F = 1.6$ in eq.(7).

Settling time of error The behavior of error with respect to time in this case is really similar as the case of Type I model bias. Fig.(12) shows that the error for model 3 converges rather rapidly, and for other models, which do not capture some aspect of the model error, the error is initially small, takes a long time to increase and finally stay stable around a large value. We will use 200 as number of time steps in the experiment below, and time step 100 to 200 as the period for time averaging.

Performance with μ Fig.(13) shows the relation between error and μ in this case. We notice that the plot is quite different from those of Type I and II model bias. For a range of values for μ, model 3, which is supposed to capture the dynamics the best, performs significantly better than other models. However, as we increase μ, the performance of all models is improved, and eventually, at $\mu = 4$, the advantage for model 3 is not as apparent, though still a little bit better than the model without bias estimation. Also, the minimum error achieved is considerably larger than those of the previous cases. One way to explain this is: although model 3 is formulated such that it includes all the bias parameters, the actual effect of mixture of the two types of model error is complicated than
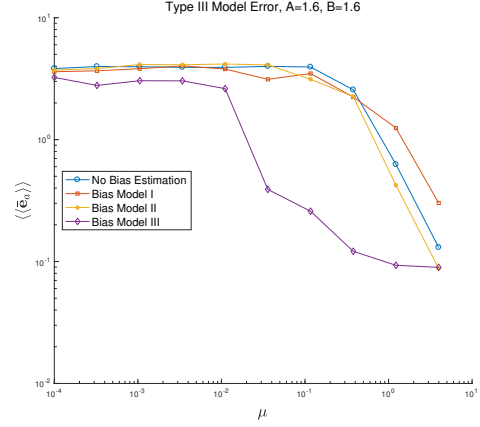

Figure 13: Time-averaged rms analysis error, $\langle\langle \bar{\boldsymbol{e}}_a \rangle\rangle$, versus the variance inflation parameter μ, for the Type III model error, with all 4 models.
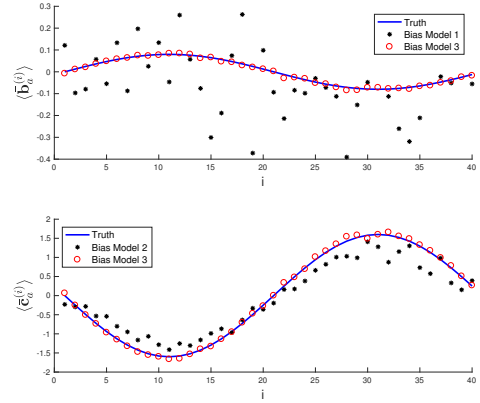

Figure 14: Time-averaged bias estimate at each location for the case with Type III model error. The truth $\zeta_i \Delta t$ and $-\xi_i$ are shown as the solid curve.

we initially think.

Bias estimation Fig.(14) shows the results for bias estimation. Although model 3 does not significantly outperform the other models in terms of time-averaged rms error, it does capture the bias remarkably well. The estimation essentially overlaps the truth. Meanwhile, model 1 completely fails to estimate $\boldsymbol{\zeta}$, and the estimation of $\boldsymbol{\xi}$ by model 2 is also much worse than that by model 3. This might suggest that model 3 is still potentially the best model for Type III model error, and we need to devise a better algorithm, or a different form of bias model, based on model 3, to improve the performance.

7

1.5   Conclusions

We considered three bias models for use in state space augmentation strategies to account for the three types of model error.

We find out the performance of data assimilation strongly depends on the specific type of model error and the specific model used for state augmentation. Generally speaking, if a model is formulated to account for certain type of error, than in the case with that type of error, the model will outperform other models. However, if the model error is complicated, then the advantage is not as apparent.

Moreover, we find out that the inflation coefficient μ plays an important role in determining the performance. Generally speaking, the less confidence we have for the model, or the farther from the truth the model is, or the larger the model error is, the larger μ we should use. Nevertheless, if the model is really close to the truth, or if the model captures the dynamics really well, then we should choose μ carefully; in many cases, the optimal μ exists in a relatively small range.