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.

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:

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

Types of error

$$
\dot{\boldsymbol{x}} = \boldsymbol{L}(\boldsymbol{x}) + \boldsymbol{\zeta} \tag{3}
$$

$$
\dot{\boldsymbol{x}} = \boldsymbol{L}(\boldsymbol{x} + \boldsymbol{\xi}) \tag{4}
$$

$$
\dot{\boldsymbol{x}} = \boldsymbol{L}(\boldsymbol{x} + \boldsymbol{\xi}) + \boldsymbol{\zeta} \tag{5}
$$

For simplicity, we assume that $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$ 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

$$
\zeta_i = A \sin(2\pi \frac{i-1}{N}) \tag{6}
$$

$$
\xi_i = B \sin(2\pi \frac{i-1}{N}), \quad i = 1, ..., N \tag{7}
$$

where A and B are constants.

Model

Model 1

To address the first type of 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:

$$\boldsymbol{x}_n^f = \hat{m}(\boldsymbol{x}_{n-1}^a) + \boldsymbol{b}_n^f \tag{9}$$

$$\boldsymbol{b}_n^f = f_b(\boldsymbol{x}_{n-1}^a, \boldsymbol{b}_n^a) \tag{10}$$

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

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

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

$$\boldsymbol{c}_n = m(\boldsymbol{x}_{n-1}^t) - \hat{m}(\boldsymbol{x}_{n-1}^m) = m(\boldsymbol{x}_{n-1}^t) - \hat{m}(\boldsymbol{x}_{n-1}^t - \boldsymbol{c}_{n-1}) \tag{12}$$

The augmented system then becomes

$$\boldsymbol{x}_n^f = \hat{m}(\boldsymbol{x}_{n-1}^a) \tag{13}$$

$$\boldsymbol{c}_n^f = f_c(\boldsymbol{x}_{n-1}^a, \boldsymbol{c}_n^a) \tag{14}$$

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

$$\boldsymbol{c}_n^f = f_c(\boldsymbol{x}_{n-1}^a, \boldsymbol{c}_{n-1}^a) = \boldsymbol{c}_{n-1}^a \tag{15}$$

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:

$$\boldsymbol{y}_n = H(\hat{\boldsymbol{x}}_n + \boldsymbol{c}_n) \tag{16}$$

Model 3

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

$$\boldsymbol{x}_n^f = \hat{m}(\boldsymbol{x}_{n-1}^a) + \boldsymbol{b}_n^f \tag{17}$$

$$\boldsymbol{b}_n^f = f_b(\boldsymbol{x}_{n-1}^a, \boldsymbol{b}_{n-1}^a, \boldsymbol{c}_{n-1}^a) \tag{18}$$

$$\boldsymbol{c}_n^f = f_c(\boldsymbol{x}_{n-1}^a, \boldsymbol{b}_{n-1}^a, \boldsymbol{c}_{n-1}^a) \tag{19}$$

Again, for simplicity, we assume $\boldsymbol{c}_n$ and $\boldsymbol{b}_n$ are constant, i.e.,

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

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

Same as in Model 2, the observation operator becomes

$$\boldsymbol{y}_n = H(\hat{\boldsymbol{x}}_n + \boldsymbol{c}_n) \tag{22}$$

Numerical methods

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 at each step to the observation operator.

We will use 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

$$\boldsymbol{e}_n = \bar{\boldsymbol{x}}_n^a - \boldsymbol{x}_n, \tag{23}$$

and the analysis error of Model 2 and 3 is

$$\boldsymbol{e}_n = \bar{\boldsymbol{x}}_n^a + \bar{\boldsymbol{c}}_n - \boldsymbol{x}_n, \tag{24}$$

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 $\boldsymbol{c}_n$