Application of the Long-Short Term Memory neural networks to model bias correction: idealized experiments with the Lorenz-96 model

Arata Amemiya[1,2] , Shlok Mohta[3], and Takemasa Miyoshi[1,2]

RIKEN

1. RIKEN Center for Computational Science (R-CCS), Kobe, Japan
2. RIKEN Cluster for Pioneering Research, Kobe, Japan
3. Graduate School of Engineering, the University of Tokyo, Tokyo, Japan

Data Assimilation Team

Background

Model bias correction has been studied as an important subject in data assimilation. Conventional methods of model bias correction assumed a simple functional form such as a constant (Dee and Da Silva, 1998) or a linear dependence on model state variables (Danforth et al., 2007). In this study, an application of the Long-Short Term Memory (LSTM) neural networks on model bias correction problems is explored in the context of data assimilation using the Local Ensemble Transformed Kalman Filter (LETKF).

The proposed method is applicable to a model bias which is dependent on the current and past model states in an arbitrarily nonlinear manner. Spatial localization is found effective and is also implemented. The new method is examined by idealized numerical experiments using a multi-scale Lorenz-96 model (Lorenz 1996; Wilks 2005).

Experimental setup

"Nature run" : Lorenz96 model with an additional term

a) Constant additive bias
b) Advection term multiplicative factor (x1.2)
c) Coupled Lorenz96 (Lorenz, 1996; Wilks, 2005)
d) Shear Lorenz96 (Pulido et al., 2018)

$$\frac{d}{dt}x_k = x_{k-1}(x_{k+1} - x_{k-2}) - x_k + F - \frac{hc}{b} f_k(\boldsymbol{y})$$

$$\frac{d}{dt}y_j = cby_{j+1}(y_{j-1} - y_{j+2}) - cy_j + \frac{hc}{b} g_j(\boldsymbol{x})$$

$$f_k(\boldsymbol{y}) = \sum_{j=(k-1)J/K+1}^{kJ/K} y_j$$

$$g_j(\boldsymbol{x}) = \alpha\left(x_{\mathrm{int}(j/K)+1} - x_{\mathrm{int}(j/K)-1}\right)$$

Observation data is created from the nature run with random noise with standard deviation $\sigma = 0.01$
Location : all grid points Period: $\Delta t = 0.05$

Forecast model : imperfect Lorenz96 model

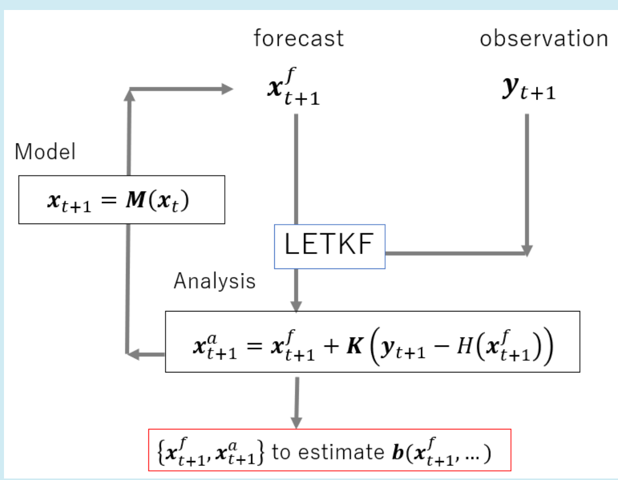
$$\frac{d}{dt}x_k = x_{k-1}(x_{k+1} - x_{k-2}) - x_k + F + \text{(Bias correction)}$$

Parameter values $K = 16, F = 8$ are common for all experiments.

Data Assimilation : LETKF

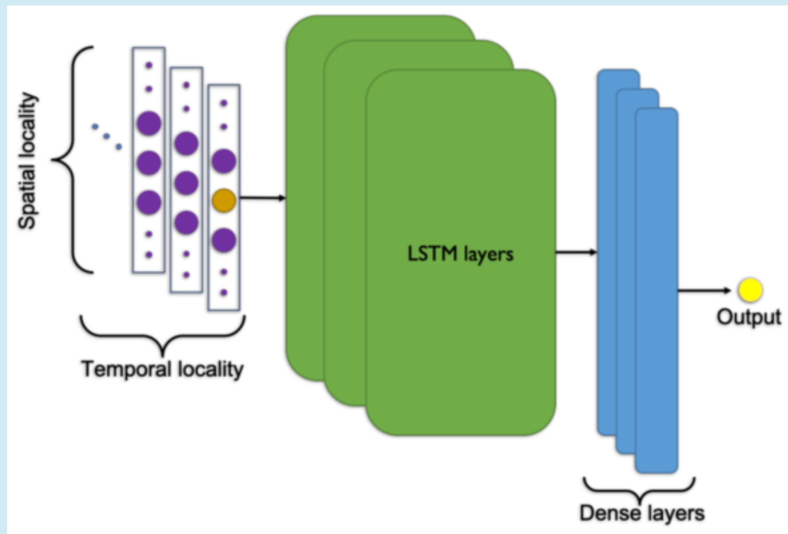
Common settings :
number of ensemble members: 20
Covariance inflation : multiplicative
Localization cutoff length: 5 grids

The estimation of the bias correction function is performed in an offline manner, using analyses x^a and biased forecast x^f as input data for training.



Bias correction by neural networks

The neural network consists of LSTM layers and dense (fully-connected) layers. As the bias correction function is spatially localized, the input data is x^f in a shape of [number of past step, number of localized grid], whereas the output is a single value of $x^a - x^f$. Network parameters are optimized by *Optuna*. This study used a network with 1 LSTM layer and 3 dense layers, and 5 past time steps to input.



Evaluation

Analysis RMSE

With the trained neural network, a set of data assimilation experiments is performed, with various multiplicative inflation factor values.

Extended ensemble forecasts RMSE

The ensemble forecasts are performed, starting from the analysis at a specific initial time in the analysis sequence. RMSE is calculated as a difference between Nature run and the forecast as a function of forecast time, and averaged over forecasts with 100 different initial times.

Comparison with other methods

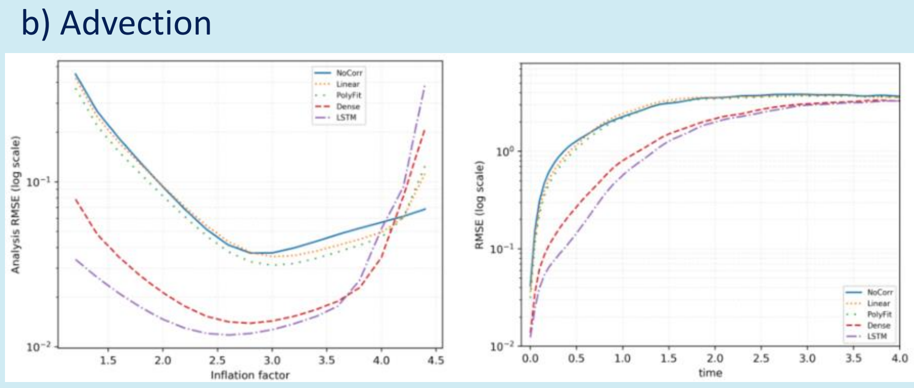
For comparison, the experiments above are also performed with other bias correction methods listed below:

1. No correction
2. Simple linear regression
3. 4th order polynomial regression
4. Dense neural networks without recurrent layer

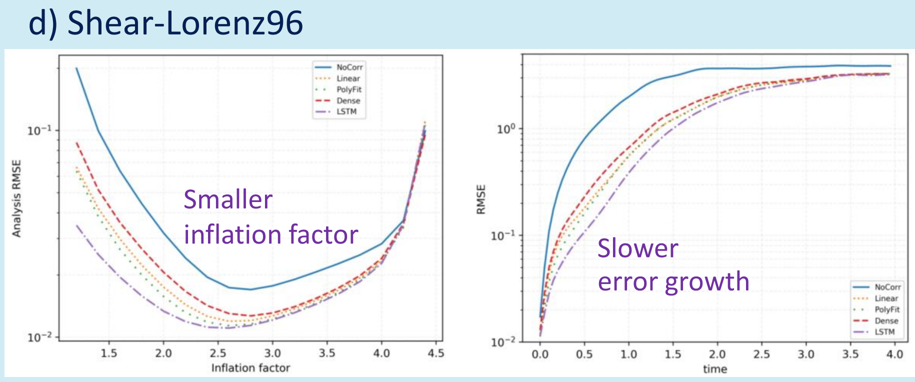
Results

The results of analysis and forecast RMSE are shown on the right, for b) Advection and d) Shear-Lorenz96 experiments.
The bias correction scheme by the LSTM (purple curves) showed better performance than that by polynomial fitting and dense neural networks. The LSTM reduced not only initial analysis RMSE, but also the error growth rate in the beginning of extended forecasts.

b) Advection



d) Shear-Lorenz96



References

Akiba, Takuya, et al. "Optuna: A next-generation hyperparameter optimization framework." *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2019.
Dee, Dick P. "Bias and data assimilation." *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 131.613 (2005): 3323-3343.
Danforth, Christopher M., Eugenia Kalnay, and Takemasa Miyoshi. "Estimating and correcting global weather model error." Monthly weather review 135.2 (2007): 281-299.
Li, Hong, et al. "Accounting for model errors in ensemble data assimilation." Monthly Weather Review 137.10 (2009): 3407-3419.
Lorenz, Edward N. "Predictability: A problem partly solved." Proc. Seminar on predictability. Vol. 1. No. 1. 1996.
Wilks, Daniel S. "Effects of stochastic parametrizations in the Lorenz'96 system." Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography 131.606 (2005): 389-407.