Something about ENKF

Chu Zhuyiheng

OrangeUser

2024.11.30

Key References

- Houtekamer, P. L., & Mitchell, H. L. (1998). *Data Assimilation Using an Ensemble Kalman Filter Technique*. Monthly Weather Review, 126(3), 796–811. Available at: Link to Article
- Houtekamer, P. L., & Mitchell, H. L. (2001). *A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation*. Monthly Weather Review, 129(1), 123–137. Available at: Link to Article

Experimental Environment

Using radiosondes and satellites to observe atmospheric state, using domain averaged streamfunction error squared north of 208N at 50 kPa to measure of the error variance.

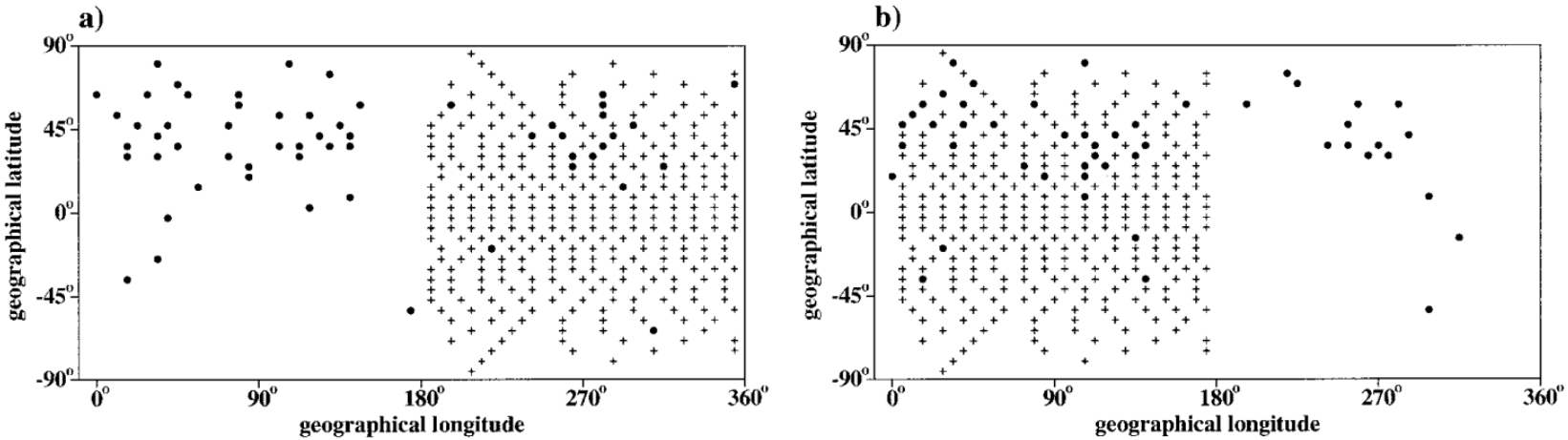
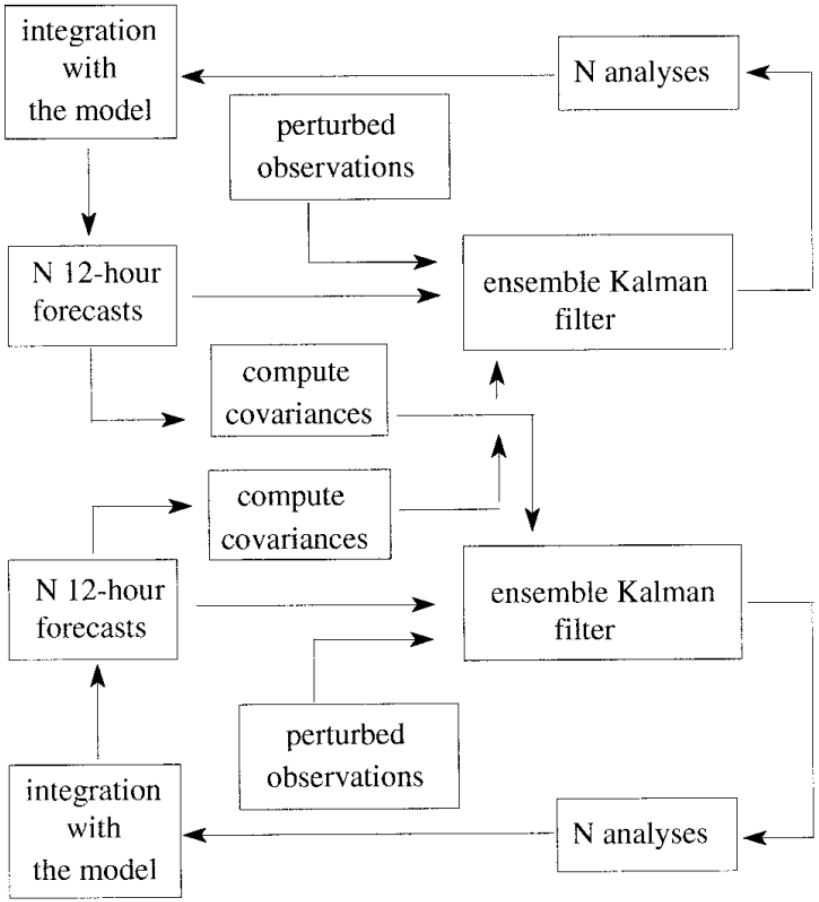


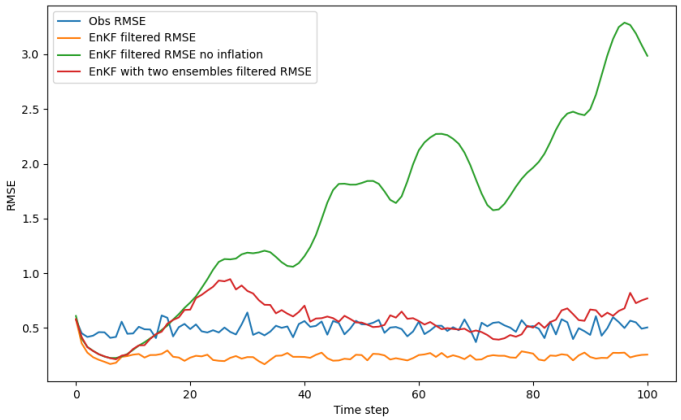
FIG. 1. (a) The observational network at 0000 UTC. Dots indicate the positions of radiosonde observations. The plus signs indicate the positions of satellite observations. (b) As in (a) but at 1200 UTC.

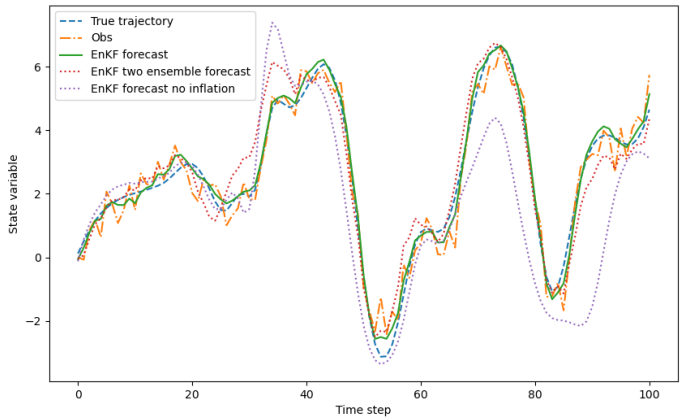
How to get the true state and what the stream function is are not important for understanding the subsequent ENKF part, so we will skip it.

Setup


integration with the model
N analyses
perturbed observations
N 12-hour forecasts
ensemble Kalman filter
compute covariances
compute covariances
N 12-hour forecasts
ensemble Kalman filter
perturbed observations
integration with the model
N analyses

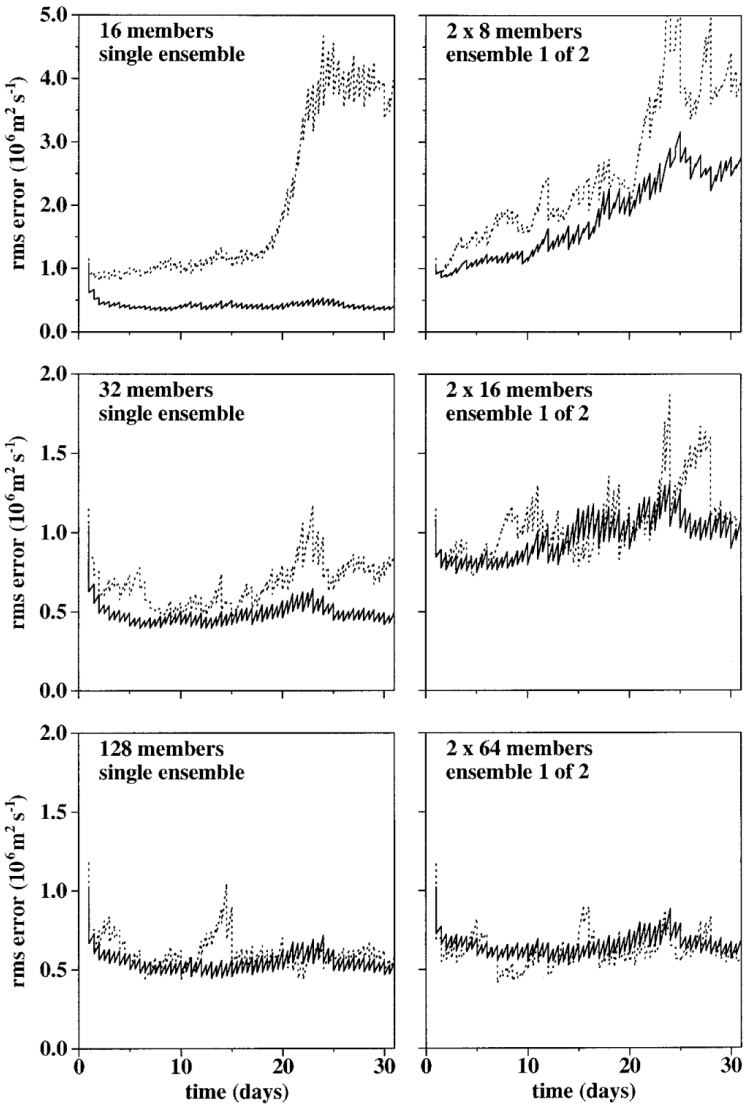
Try to use it in Project3





The effect is not as good as using inflation, but better than not using inflation.

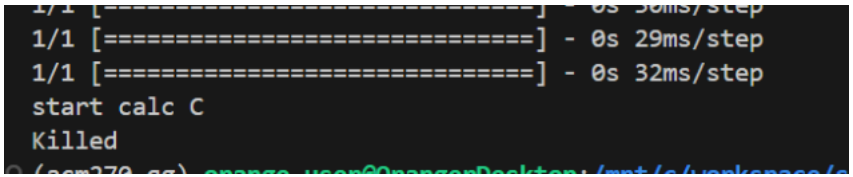
Why it works



The solid line on the graph is the rms error, and the dotted line is the rms spread.
Some guesses:Perhaps when it is hard to carefully select the inflation parameter, 2 ensembles EnKF would also be a possible option?

Problem in Project2

I guess all students have encountered similar situations in project2?



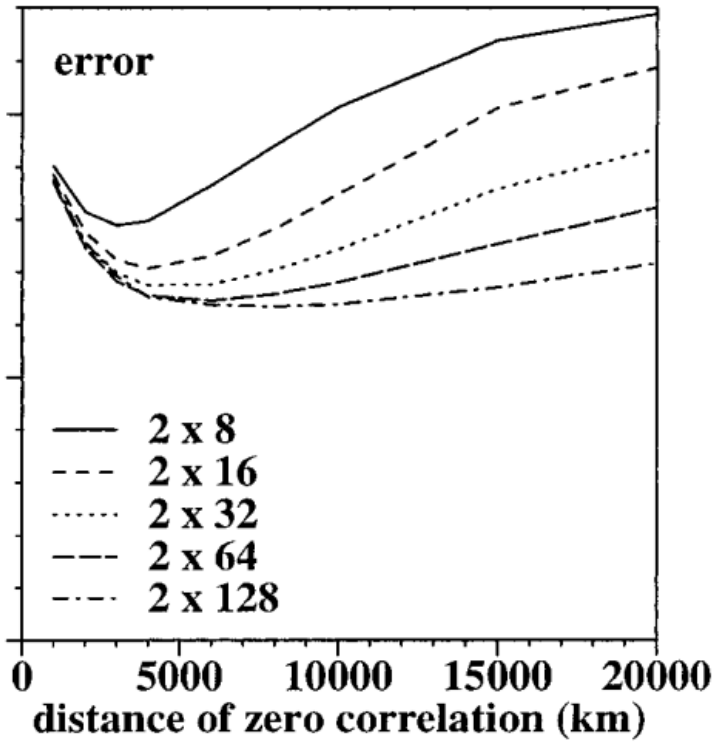
This is why we need to use localization to make the matrix sparse.

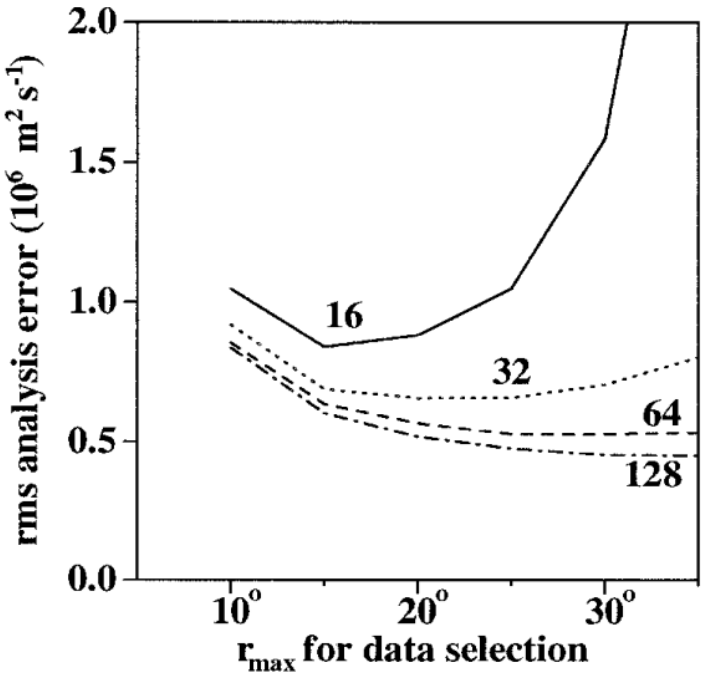
Cutoff method and Schur product method

- Cutoff method:For each vertical column of analysis points, all the data within a given horizontal distance, r_{max}, and only that data, is used.
- Schur product method:

$$\mathbf{K}_j = [\rho \circ (\mathbf{P}_j^f \mathcal{H}^{\mathrm{T}})][\rho \circ (\mathcal{H}\mathbf{P}_j^f \mathcal{H}^{\mathrm{T}}) + \mathbf{R}]^{-1},$$

Their results are similar





My result

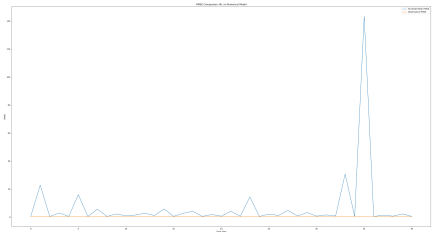


Figure 1: $r_{max} = 2$

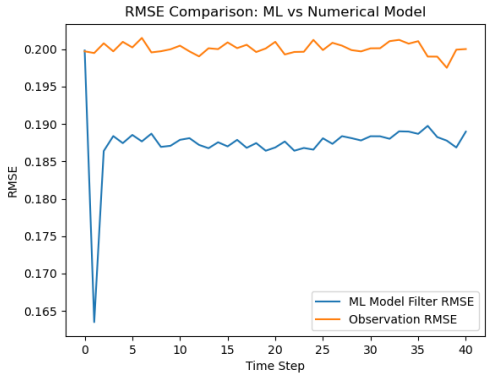


Figure 2: $r_{max} = 3$

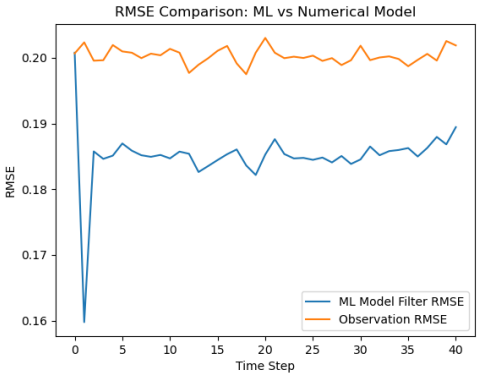
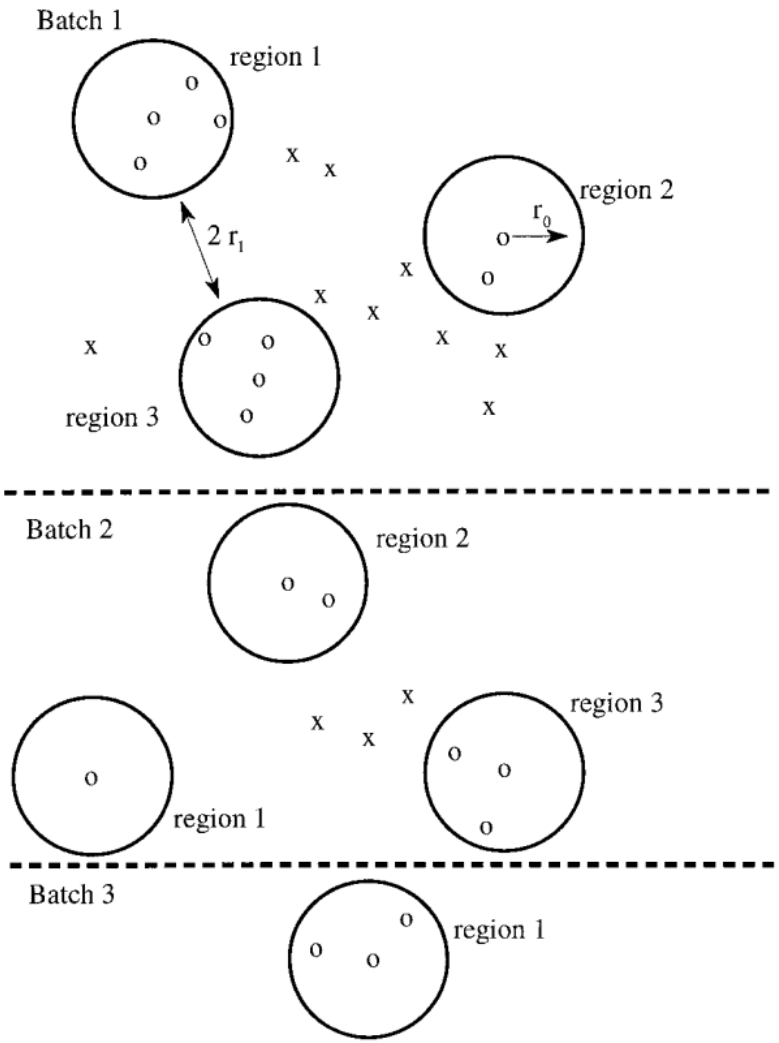


Figure 3: $r_{max} = 4$

My computer can't run for larger data, but the existing data and the trends mentioned in the original article are similar.

Sequential processing of batches



In each batch, multiple regions that are far away from each other are selected for parallel assimilation, and the assimilation of a batch of observations is completed in multiple batches.
It looks interesting, but I don't know how to implement it yet.
The observation method in the project is not consistent with that in the paper.