An Efficient Ensemble Data Assimilation Approach To Deal With Range Limited Observation

A. Shah[1,2], M. E. Gharamti[3], L. Bertino[1], M. Borup[4]
[1]Nansen Environmental and Remote Sensing Center (NERSC), [2]University of Bergen (UIB)
[3]National Center for Atmospheric Research (NCAR), [4]Technical University of Denmark (DTU)

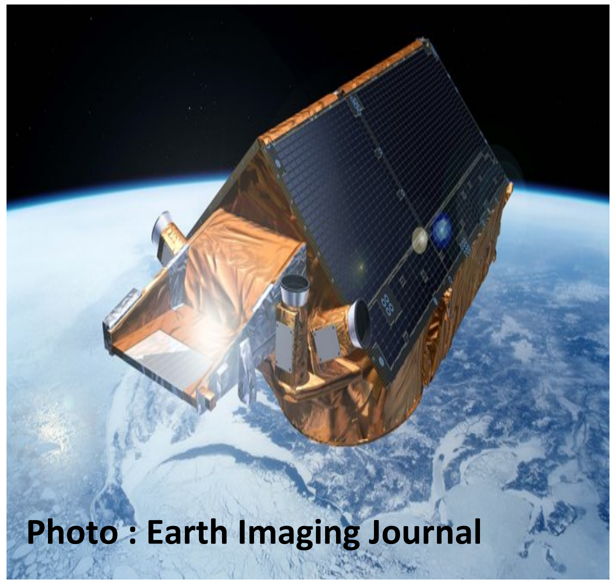
Objective :

To develop a Data Assimilation algorithm under the framework of stochastic EnKF for assimilating range limited observation so as to use the qualitative information from it.

1. Range Limited Observation (RLO) :

Observations obtained via any source exist only in certain interval of its range, due to limitation of measuring gauge and etc.

Examples of range limited observations



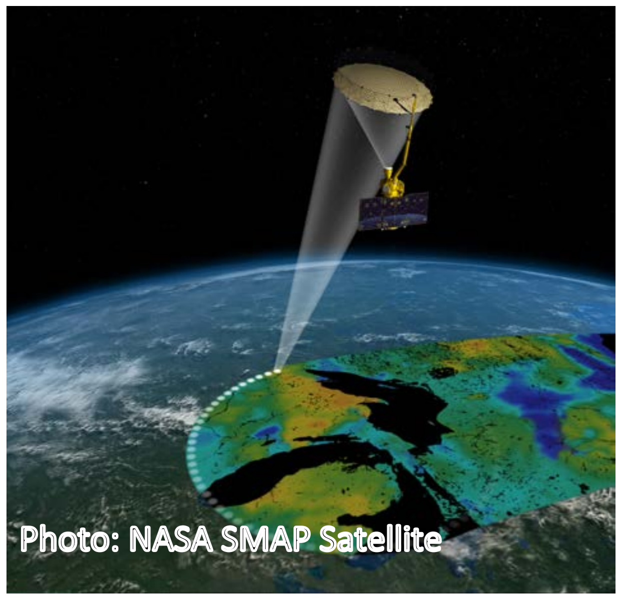


Figure 1.a: Illustrative example of sea-ice thickness which is available only up till 50cm in real time from SMOS satellite and figure b: SMAP satellite measure soil moisture for top 5cm

2. Methodology and Algorithm

Bayesian Rule : $$p(\boldsymbol{x} \mid \boldsymbol{y}) = \frac{p(\boldsymbol{x})\,p(\boldsymbol{y} \mid \boldsymbol{x})}{p(\boldsymbol{y})}$$

Concept of Ensemble partial updating (EnPU) is used for RLO, which allows to use qualitative information of the data i.e., posterior will be

$$p\left(\mathbf{x}_k \mid \mathbf{y}_{\text{quant}}, \mathbf{y}_{\text{qual}}\right)$$

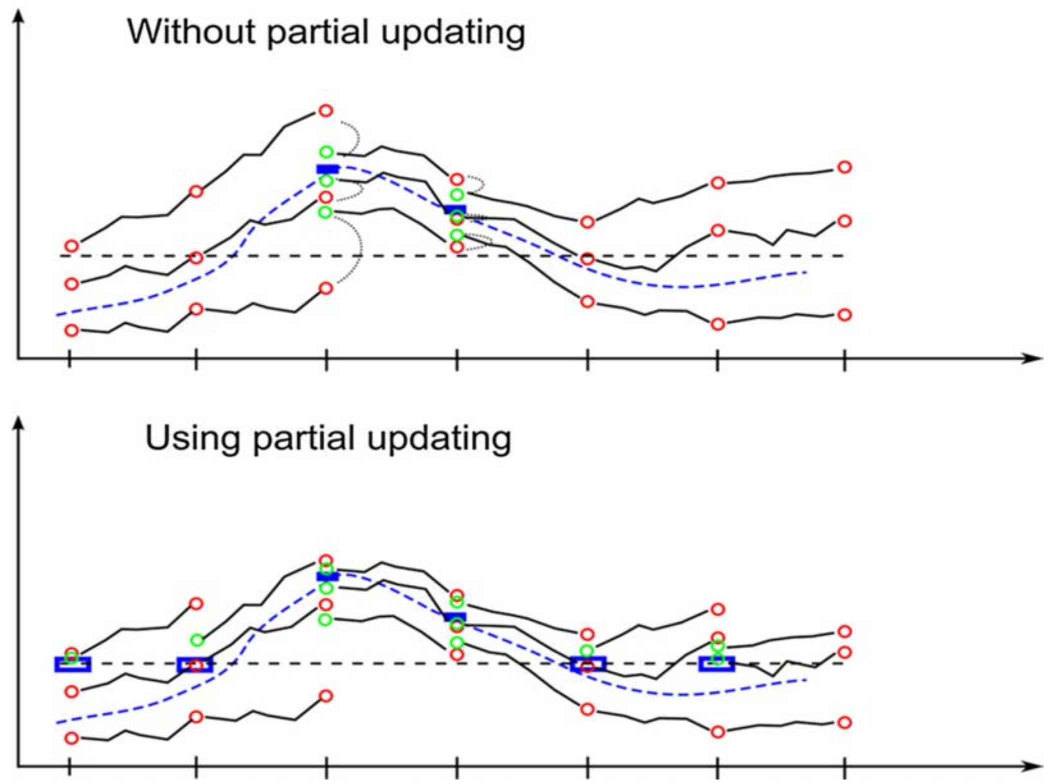


Figure 2: (Borup et. al., 2015) With and without partial updating when the measurement gauge has lower observation limit

Partial Ensemble Kalman Filter (PEnKF)

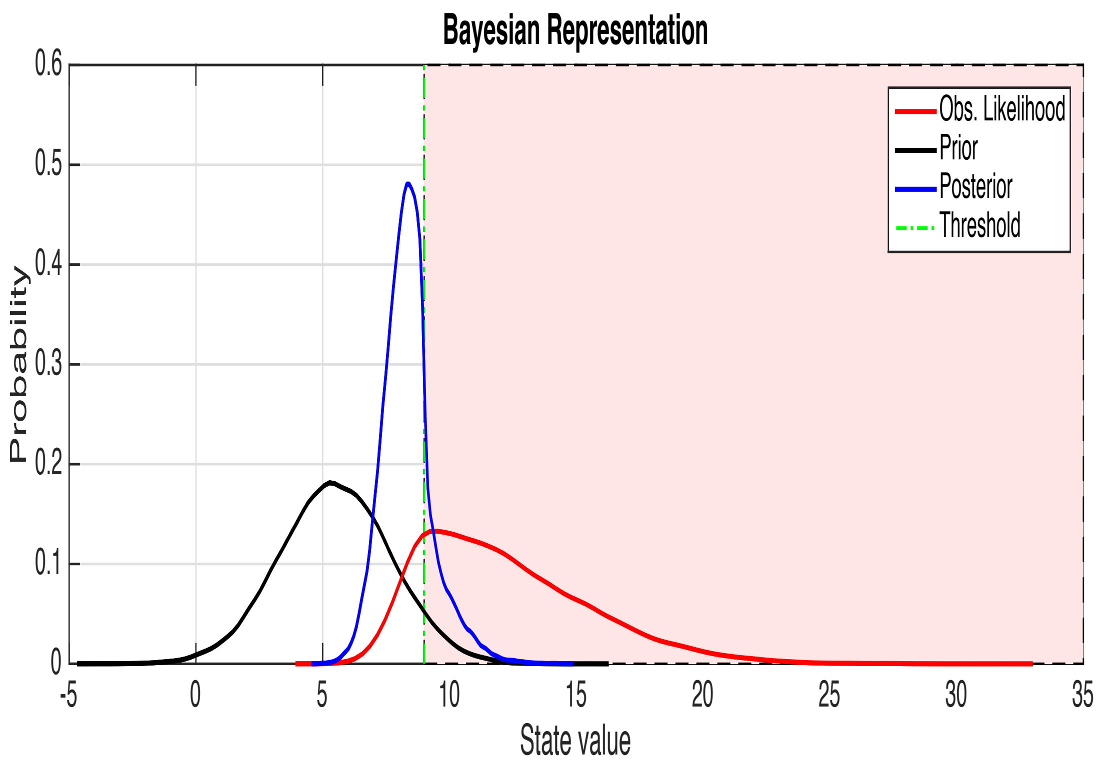
- The forecast and update equations remains the same as in Stochastic Ensemble Kalman Filter (Evensen, 1994)
- Update only those members which lies within the observable range
- For out of range observations, creating a virtual observation at threshold limit
- Using Two Piece Gaussian distribution (Fechner's Kollektivmasslehre, 1897), perturbing observations where one of the observation error standard deviation will be

$$\sigma_{or} = p * \left(\mathbf{H}\bar{\mathbf{x}}^{f}\right)$$

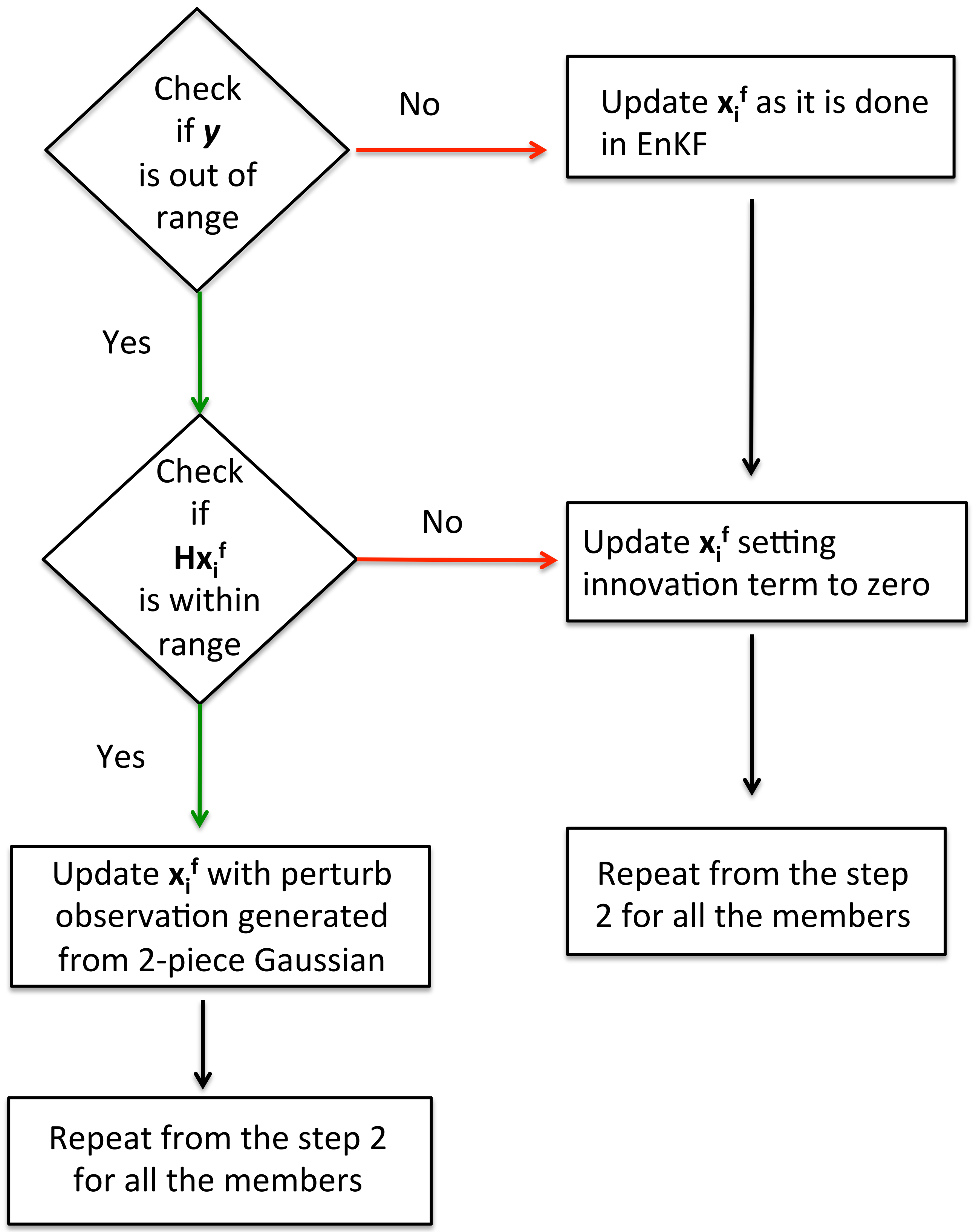
where p is real number between (0,1]

Posterior when the prior is in-range

$$p\left(\mathbf{x}_k \mid \mathbf{y}_{quant}, \mathbf{y}_{qualit}\right) \propto \begin{cases} p(\mathbf{x}_k)p(\mathbf{y}_{quant} \mid \mathbf{x}_k) \\ p(\mathbf{x}_k)p(\mathbf{y}_{qualit} \mid \mathbf{x}_k) \end{cases}$$



Algorithm : The flow Chart for the PEnKF



3. Numerical Experiments

- ✧ EnKF, PEnKF, and EnKF-Ignored (observation out of range ignored), DA methods are tested under the framework of twin experiment.
- ✧ Model – Lorenz'96
- ✧ Sensitivity analysis with number of observations, observation frequency, threshold limit, model error and etc., was performed

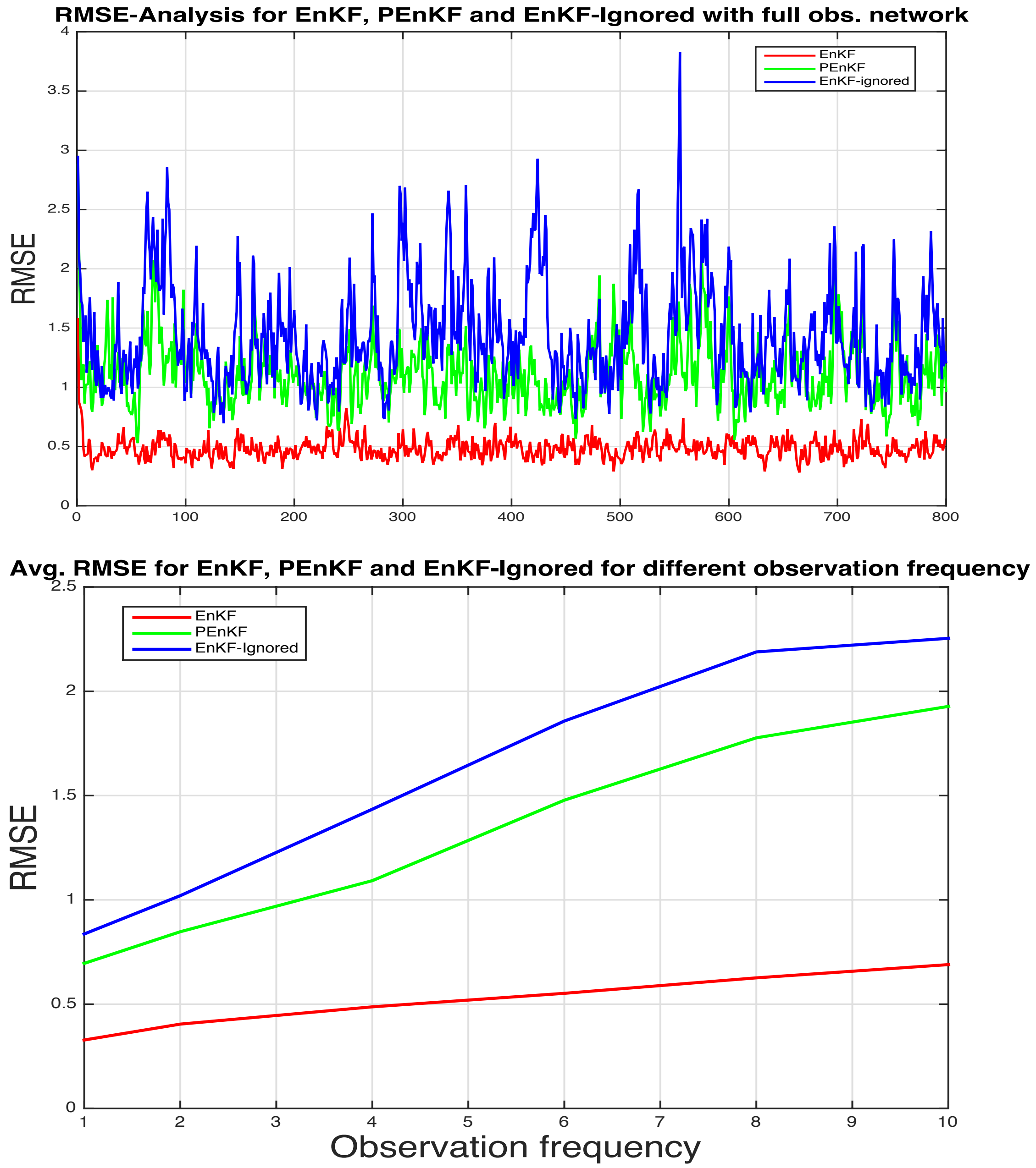


Figure 3a: RMSE for analysis in time where 75% of observations are out of range on an average for total time of integration. 3b: RMSE for analysis for different observation frequency

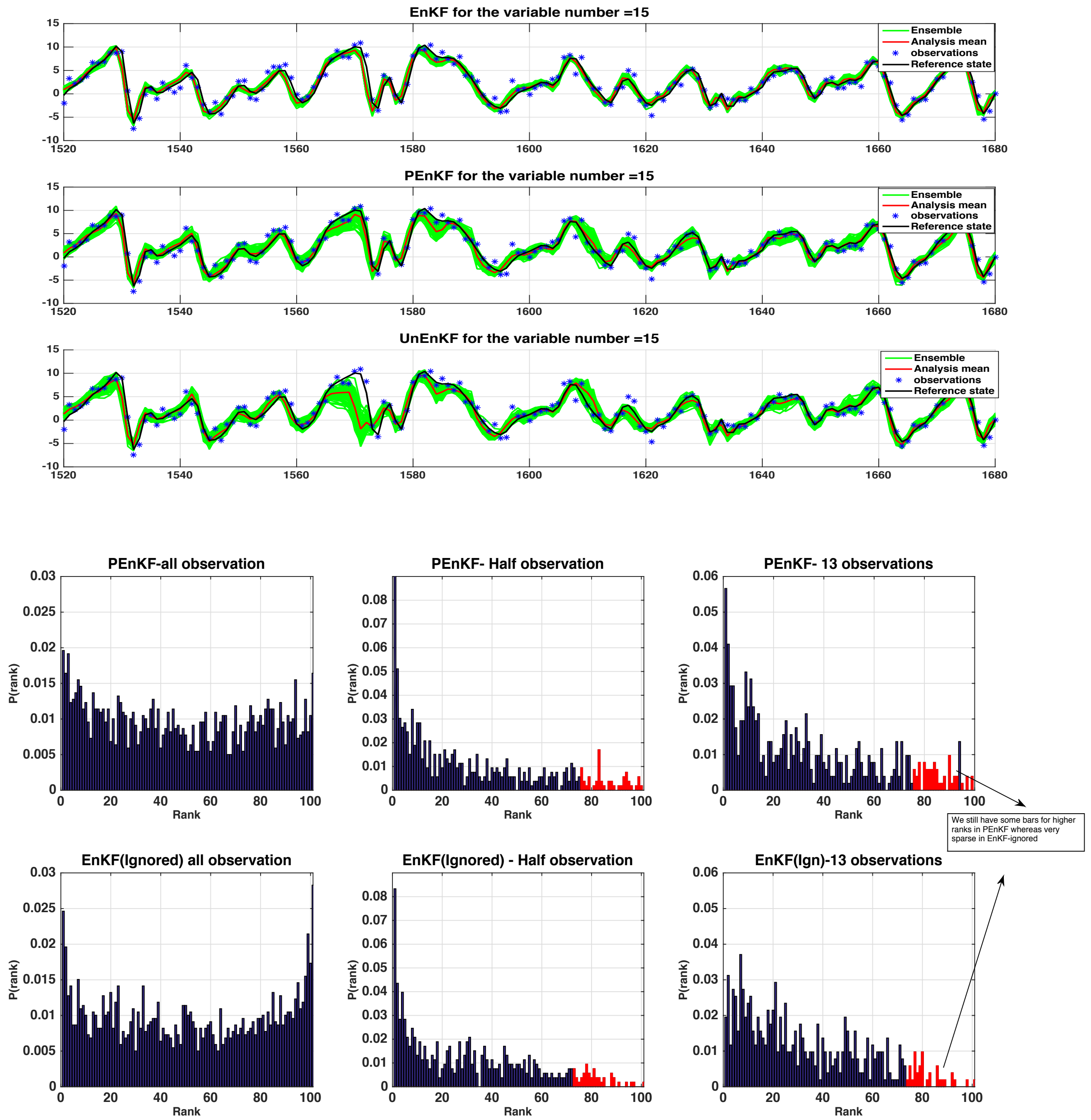


Figure 4a: snap shot of time series of a particular state variable. b) Rank histogram(reliability) for all, half and quarter observations.

4. Conclusion

- ✓ Improves quality of forecast
- ✓ Reduce uncertainty
- ✓ Improves reliability of forecast

Future work : Implementation with real data set on small scale real world problem

Acknowledgement : Research is funded by project **EmBLA** (Ensemble-based data assimilation for environmental monitoring and prediction) Nordic Center of Excellence for ensemble based DA