Improving Background Error Covariance with the Convolution Neural Network (CNN) in the Gain form Ensemble Transform Kalman Filter (GETKF)

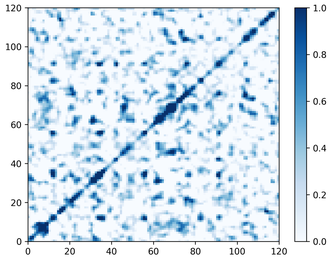
Thiruvengadam Padmanabhan[1], Xuguang Wang[1,2], Yongming Wang[1,2]

1. **M**ulti-scale data **A**ssimilation and **P**redictability (**MAP**) Laboratory, University of Oklahoma 2. **C**onsortium for **A**dvanced **D**ata **A**ssimilation **R**esearch and **E**ducation (**CADRE**)

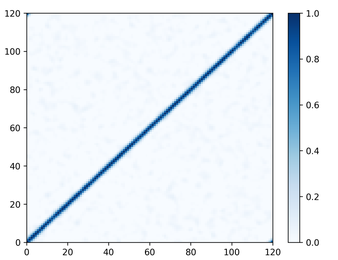
Motivation

Sampling error from small ensembles limits the accuracy of background error covariance in ensemble data assimilation.

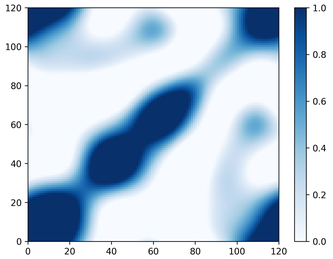
(a) Small scale raw covariance



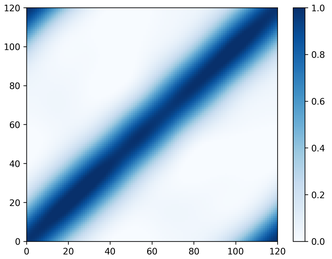
(b) Small scale 2000 members covariance



(c) Large scale raw covariance



(d) Large scale 2000 members covariance



Existing Approach

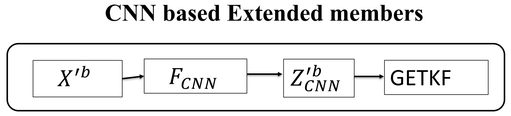
- Ensemble expansion with Gaspari-Cohn (GC) in GETKF and multiscale variant (MLGETKF) reduces spurious sampling error correlations.

Limitation

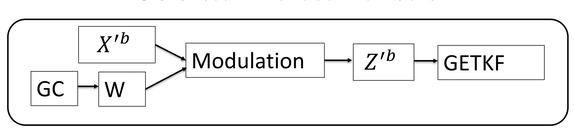
- Do not correct variance errors; need extra inflation.
- GC Fixed cutoff lacks adaptiveness to sampling error variability.
- Higher computational cost for small-scale features; more expanded members needed.

Proposed method: CNN Ensemble Expansion

CNN based Extended members



GC based Extended members



X'^b : Raw perturbation; Z'^b : GC Extended perturbations; Z^b_{CNN}: CNN extended perturbations W: Square roots of localization; F_{CNN} : Trained CNN model.

Experiment Design

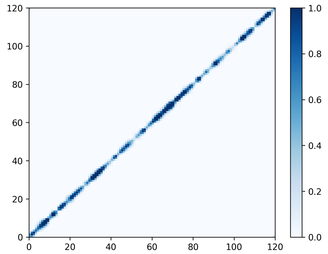
CNN was tested at small and large spatial scales within GETKF.

- **Equal Cost**: Same number of expanded members as GC to compare accuracy.
- **Reduced Cost**: Far fewer CNN expanded members to assess efficiency and accuracy

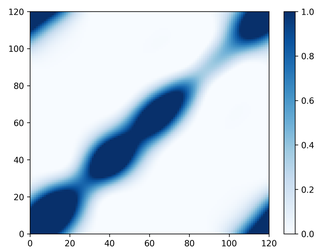
Results

Evaluation of Covariance Accuracy: Equal cost

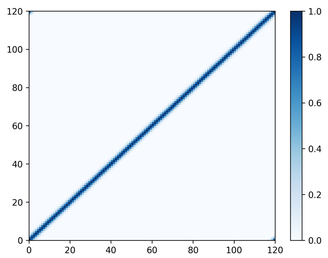
(a) Small scale GC covariance



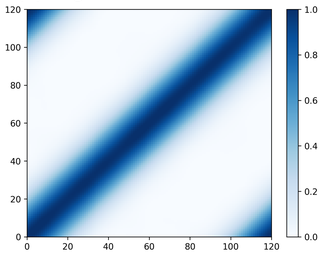
(b) Large scale GC covariance



(c) Small scale CNN covariance

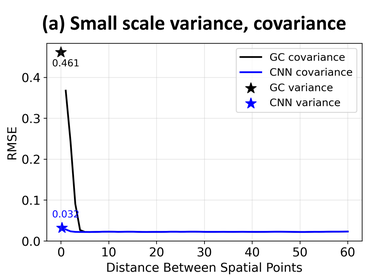


(d) Large scale CNN covariance

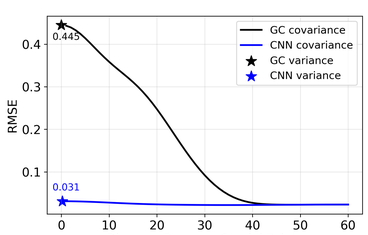


- CNN reduces both variance and off-diagonal errors and its correlation curve is flatter and lower
- GC shows a large variance in RMSE and shows a sharp, short-range correlation error peak due to the rapid decay of localisation weights.

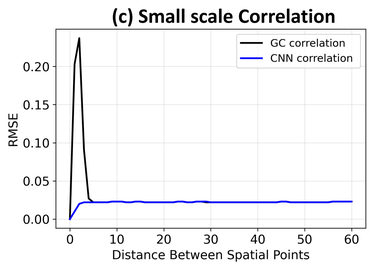
(a) Small scale variance, covariance



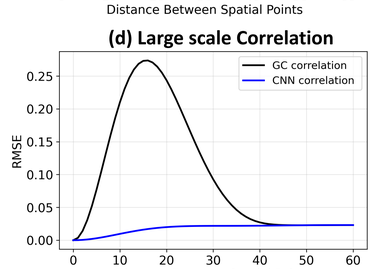
(b) Large scale variance, covariance



(c) Small scale Correlation

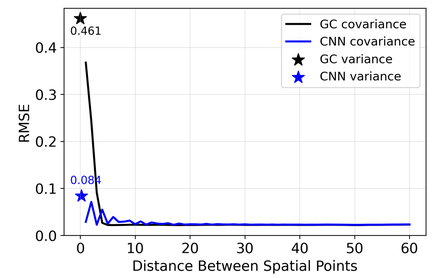


(d) Large scale Correlation

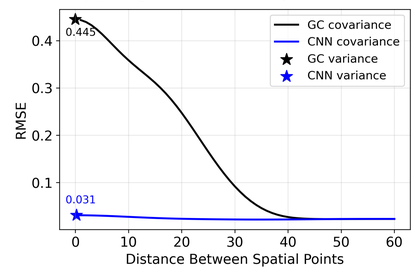


Evaluation of Covariance Accuracy: Reduced cost

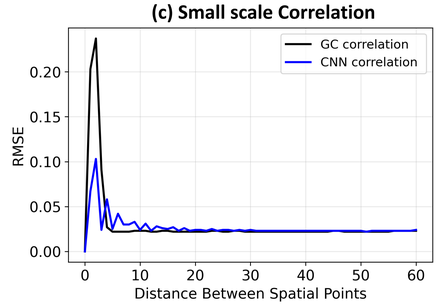
(a) Small scale variance, covariance



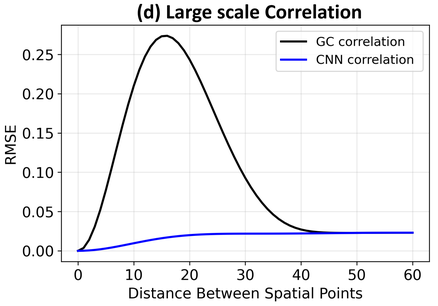
(b) Large scale variance, covariance



(c) Small scale Correlation

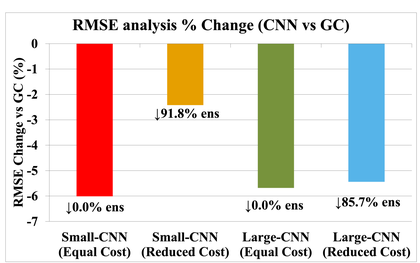


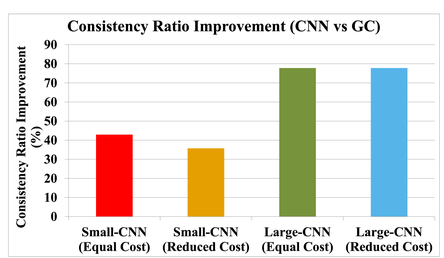
(d) Large scale Correlation



- Small scale reduced cost CNN (–91.8% members) still outperforms GC at short separations (≤3 grid points), with lower variance, covariance, and correlation RMSE.
- Beyond 3 points, CNN is slightly worse, but near-range improvements (≤3 grid points) outweigh longer-range losses.
- In the case of a large scale reduced cost CNN (-85.7% members), CNN maintains covariance and correlation structures even with reduced ensembles.

Comparison of GETKF analysis





CNN consistently reduces analysis RMSE and improves ensemble reliability (CR) compared to GC, at both equal and reduced expanded ensemble sizes.

Conclusion

- CNN based ensemble expansion outperforms GC at both small and large scale by correcting both variance and correlation errors, lowering covariance RMSE by up to ~ 89% and reducing analysis RMSE by ~6% at equal cost.
- Even with ~ >85% fewer members, CNN still improves accuracy and ensemble reliability, offering a more computationally efficient solution for data assimilation.

References

Bishop, C. H., Whitaker, J. S., & Lei, L. (2017). Gain form of the ensemble transform Kalman Filter and its relevance to satellite data assimilation with model space ensemble covariance localization. *Monthly Weather Review*, *145*(11). https://doi.org/10.1175/MWR-D-17-0102.1

Wang, X., Chipilski, H. G., Bishop, C. H., Satterfield, E., Baker, N., & Whitaker, J. S. (2021). A multiscale local gain form ensemble transform Kalman filter (MLGETKF). *Monthly Weather Review*, *149*(3). https://doi.org/10.1175/MWR-D-20-0290.1