

# Review report for npg-2022-19 “Using orthogonal vectors to improve the ensemble space of the EnKF and its effect on data assimilation and forecasting”

## Main concerns

The idea of improving the performance of an EnKF by artificially increasing the ensemble size appears sensible in general. Nevertheless, there appears to be a major weakness in the current manuscript, that is, the technical idea is not fully supported/explained by the numerical experiments, and I think more thorough numerical investigations would be required. In this regard, I'd like to invite the authors to consider the following questions/suggestions:

1. If it's beneficial to artificially increase the ensemble size, why did the authors only consider to increase the size by one or two in their implementation? In general, how many more pseudo-members can the authors include before the EnKF's performance saturates (or perphas deteriorates)?
2. The authors considered a few different ways (EMV, IESV, RSV) to add ensemble members and concluded that EMV appears to be better. Any reason to explain the better performance of EMV?
3. There are many factors, such as inflation factor, frequency and density of the observation operator, the variance of observation errors, the length of the assimilation time window and even random seeds, which could potentially affect the performance of an EnKF. I's suggest that the authors conduct an extensive investigation on the impacts of these factors on the performance of the EnKF.
4. In particular, I'd suggest the authors repeat one experiment for a number of times (each time with a different random seed) to reduce the effect of statistical fluctuations. In the current manuscript, I got the impression that the authors did each experiment only once, which does not make much sense for the L96 model, since it's quite cheap to run. If the same experiment is repeated multiple times, then I'd further suggest that the authors use a table

to summarize the results (RMSE) in the form of mean  $\pm$  STD, which would make the performance comparison statistically more meaningful.

## Minor issues

1. L18, P1: The statement “The BEC is sampled by the background ensemble” does appear precise, since one cannot sample a covariance (perhaps you want to say “estimate” instead?)
2. L22, P1: In linear algebra, the statement “cause the ensemble space to be underestimated” also appears problematic. In which sense is a (linear) space underestimated? My guess is that here by “ensemble space” the authors possibly mean ensemble variance instead. If so, it seems that there are many similar instances later in which the terminology “ensemble space” is not correctly used (e.g., as in “apply SVD to the ensemble space” in L63, P3).
3. The authors stated that using Eq. 3 can preserve the ensemble mean and spread. It’d be good to provide a proof for this statement.
4. L92, P3: Rephrase the text “as a demonstration of proof of concept”.
5. In Caption of Figure 4, “according” is duplicated.