Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC2
https://doi.org/10.5194/gmd-2022-91-RC2, 2022

Comment on gmd-2022-91

Anonymous Referee #2

Referee comment on "An EnKF-based ocean data assimilation system improved by adaptive observation error inflation (AOEI)" by Shun Ohishi et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-91-RC2, 2022

The authors propose to use a novel extension for the use of ensemble Kalman filters (EnKF) in a pre-operational ocean reanalysis product. The adaptive observation error inflation, previously introduced for satellite data assimilation, reduces assimilation increments by automatically inflating observational errors. The results show this automatic inflation as improvement compared to static observational errors. These results hold especially at Ocean frontal zones, where a large vertical diffusion can be observed with a static covariance. In general, this idea is relevant to improve data assimilation/reanalyses with ensemble Kalman filters, and the manuscript is well-written. Nevertheless, the manuscript needs a revision in its current form, at least with more and longer discussions, especially in relation to the number of figures. Also, the manuscript is not totally self-contained.

1) Whereas the ensemble Kalman filter, its assumptions, and its equations, are well-known, adaptive observation error inflation is quite unknown in the literature. Although the authors state and shortly explain the relevant equations, the explanations for this technique are too short. Its assumption and when we would expect that it works well remains totally unknown. Based on Desroziers et al., 2005 (often cited as 2005 and not 2006, which might confuse an informed reader), the relationship within the innovation statistics assumes Gaussian background and observational errors as in the ensemble Kalman filter, but what happens if these assumptions are violated?
In addition, a crucial assumption is the correct representation of the background error covariance with the ensemble; only then, Equation (1) represents a correct observational error covariance. The heavy use of relaxation to prior perturbations (RTPP) shows difficulties with the ensemble spread and I wonder if the ensemble spread is correctly tuned, especially in the Ocean frontal zones. The use of the maximum between estimated covariance and prescribed covariance lessens possible problems with these assumptions, but nevertheless, they should be named and discussed in the manuscript.
The equation from Desroziers et al. is only valid in expectation of the errors. For me, it remains unclear if and how this expectation is built in the data assimilation system. If no expectation is used, then its consequences and its connection to quality control and robust assimilation should be discussed, e.g., what happens in different innovation magnitude regimes (smaller or larger than the expected

innovation magnitude)? In total, the method part of the adaptive observational error inflation needs to be revised.

2) The results show an improvement with adaptive observation error inflation compared to a static observational error assumption. The static observational errors results into too large assimilation increments and, thus, to a strong vertical diffusion at the Ocean frontal zones. As the static observational error covariances are important for increments, its magnitudes are very important. Although the numbers are stated, their sources remains unknown. Because of the missing sources, the reader is unable to know if the prescribed uncertainties come only from the uncertainties of the observational products or if they also include other uncertainty sources like the observation operator or the representation error. The results indicate a larger representation error at the Ocean frontal zones than included in the observational error. A usual approach would be thus to generally inflate the observational errors or to withheld observations in these zones. Consequently, I would wish for a comparison experiment with an inflated observational error (e.g., 2 times the stated observational error) to see if a proper tuning of the errors would lead to better scores and how this might help in the case of the frontal zones. The authors have stated that they have only a limited computational budget, and a proper tuning of the observational errors and/or a comparison experiment might be too expensive. It might be therefore also enough to explain more in detail the advantages and disadvantages of adaptive observation error inflation compared to a tuned observational error, which can be again related to the discussion in point 1 of this review. Although the results seem to be good, the reader could be generally tempted to believe that the results are only caused by a non-tuned assimilation system.

3) In general, the results part would profit a lot on concentrating on the most important parts of the study. Although well-written, the amount of figures compared to the discussion makes it difficult to follow the red line in the results part. Sometimes, similar information is shown twice (e.g., Figure 6-8) and could be condensed into a single figure. Caused by the difficulties to follow the red line and a rather loose summary section, the main message of the manuscript remains also slightly unclear for me. On the one hand, this study tries to show how the static observational error induces problems with the vertical diffusion. On the other hand, it promotes of how adaptive observational error inflation can help. As discussed in section 2 of this review, the sensitivity experiments might be not enough to promote adaptive error inflation and to cancel out difficulties with the static observational error.

I like how the authors explain their evaluation in detail within the results part, but in its current form, it distracts from the main results and is too long. I would recommend to give here only concise explanations of the evaluation and to move specific equations and details into the appendix.

Smaller comments:

As an advice, the chosen colormaps might be generally misleading and inaccessible for colour-blind persons. In addition, the same colours are used for different meanings in subfigures (e.g. Figure 5) , which can be also very misleading for the reader.

I would be interested into a comparison experiment without any data assimilation, except for example SST and SSH nudging as done for the spin-up phase. Currently, it remains unclear for me if the noisier pattern in the SST fields compared to observations are caused by the data assimilation or if this is a "natural feature" of the model. This could be even shortly stated in the results part and then simply shown in a supplementary material or if this was discussed in the other manuscript, then the authors could simply point this fact to the other submission. In this sense also the naming of the experiments is a little bit confusing as the "control" run is usually an open-loop run without data assimilation whereas here it describes the baseline EnKF experiment, I would rename it into EnKF or STATIC.

The authors frame the introduction as there are only two previous works on the EnKF for the Ocean. It might be correct that there are only two reanalysis products based on the EnKF but there is surely more work on the EnKF for the Ocean.

In line 133, the authors state that they use covariance localisation. This term might be misleading, as they seem to use observational (covariance) localisation. I would rename it into R-matrix localisation as normally used in ensemble Kalman filter literature. In line 136, the use of incremental analysis updates (IAU) is indicated. The sentence links the use of IAU to ensemble inflation, which is not its normal use in ensemble Kalman filters. I would thus split the sentence with IAU and RTPP into two sentences. In addition, it is unclear how IAU is applied, if for example the increments are applied before and after the original time point or only after etc.

In line 143, "the" SSS nudging is named, what is "the" SSS nudging? Is it the same nudging as used for the spin-up phase? If yes, please state this explicitly.

Other, smaller, issues could be resolved after a revision round.