Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC1
https://doi.org/10.5194/gmd-2022-40-RC1, 2022

Comment on gmd-2022-40

Anonymous Referee #1

Referee comment on "An ensemble Kalman filter system with the Stony Brook Parallel Ocean Model v1.0" by Shun Ohishi et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-40-RC1, 2022

Summary: This paper documents the tuning of relaxation parameters for an ocean DA based on EnKF and IAU. The methods used in this study are well established in other DA practice. Since ocean DA faces the challenge from dynamical imbalance with shorter cycling periods, the use of relaxation with IAU is a good approach and tuning results are meaningful for the ocean prediction community. I've found several issues in experimental design and result presentation, which I believe the authors should be able to address before the paper can be accepted.

Major Issues:
1. Are cases with alpha<0.9 tested for RTPP/RTPS without IAU? As RTPP+IAU approaches NO INFL+IAU results in both balance and accuracy, I guess the RTPP cases will also approach NO INFL as alpha decrease. If you have tested several points (maybe RTPP05 and RTPP07) it would be interesting to add them in the plots. For example, could there be an alpha value for RTPP that beats RTPP09+IAU?

2. The choice of inflation factor in MULTI is more problematic. Since the multiplicative inflation is applied throughout the domain, it is more sensitive the the rho value. The relaxation methods have build-in spatial variations in inflation so I think it is not a fair comparison between MULTI and RTPP/RTPS methods. In regions where analysis increments are smaller (fewer observations) the inflation of spread can accumulate over time exponentially. Ideally using a spatial varying inflation (such as in adaptive MULTI algorithms) can help. So, if you choose to show MULTI results here the exact value of rho is very important. Could you estimate an equivalent rho from the best RTPP/RTPS cases? You can averaged the (1-alpha) + alpha*prior_spread/posterior_spread over the domain and time (for RTPP) to estimate the equivalent rho, is it near 1.05 or much smaller?

3. Tuning of relaxation can also be case-dependent, you also need to consider the density of observations and localization radius. In the method description maybe you should state more clearly how you tuned localization with this observing network (can you also show a

map of observation density for reference?), and the results from tuning alpha in relaxation would likely change if one use another set of observations with different density and localization radius. A discussion in the conclusion would be nice.

4. I found time evolution of imbalance and errors to be important in this particular case. Since you used fixed values in inflation schemes, it is not guaranteed that the performance will be steady in time. Does the imbalance gradually increase or decrease over time for a chosen alpha value? A time series of spatially averaged delta NBE could be more convincing that the performance is steady. I would be also curious about how long the initial spin up period is for DA solutions to become steady.

Minor Issues:
Line 153: abs denotes taking the absolute value, please use standard notation |x|.

Line 160: the same term IR is used for both RMSE and NBE?, maybe add a suffix to distinguish.

Line 220, Table 3: gross error check not "growth error check"?

Line 237: SSS nudging: could you provide more details of this approach, maybe a reference or technical report?

Line 246: is every experiment tested against NO INFL for significance of improvement/degradation? If so, you should state this more clearly.

Line 264: this imbalance is substantially improved => reduced.

Line 276: I guess the RTPP09 and RTPS09 cases are also tested against NO INFL for significance?

Figures 1, 3, 4, 7 and 8: you used hollow/solid circles to denote significant/non-significant improvements, but for degradation you used "x" which cannot show hollow/solid differences, maybe use another symobal (triangle?) so you can be consistent.

Line 551: confidence limit: do you mean confidence level (p value < 0.01)? and no significant difference has p>0.01?, if you used t-test just state the p value threshold to be

clear.