Ocean Science
Discussions

Open Access

EGU

Interactive
comment

Interactive comment on "Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991–2013" *by* Jiping Xie et al.

Jiping Xie et al.

jiping.xie@nersc.no

Referee #1 This paper describes the result of a 23-year reanalysis (1991-2013) of the Arctic, as obtained from the TOPAZ4 coupled ocean and sea ice data assimilation system. The paper mainly provides a detailed comparison between the reanalysis and available observation datasets (sea level anomaly, sea surface temperature, in situ temperature and salinity profiles, sea ice concentration, sea ice drift and sea ice thickness). In its present form, the paper is essentially descriptive; it does not provide really new scientific ideas; and the method used to assess the ensemble reanalysis (statistics of the difference between ensemble mean and observations) is quite crude and not very original. Nevertheless, as I understand, this paper is meant to be the reference paper to an important new reanalysis product describing the Arctic Ocean (delivered by the Arctic component of the MyOcean system). As such, I think that this paper could

Printer-friendly version

Discussion paper

deserve publication in Ocean Science. I have however a few concerns concerning the manuscript that I believe should be taken into account in a revised version. 1. The main purpose of the paper is the assessment of the reanalysis using all available observations. However, to compare the reanalysis to observations, the authors just compute the average and RMS difference between the ensemble mean and observations. This method looks very crude to me, and does not make justice to the advanced method that is used to perform data assimilation. The ensemble data assimilation system provides a probability distribution for the reanalysis, which is described by an ensemble of model states. Why then assessing the reanalysis using the ensemble mean only? Probabilistic tools exist to perform an objective comparison between ensemble simulations and observations (see for instance Toth et al., 2003, or Candille et al., 2007). Why performing an ensemble reanalysis if the probabilistic information is discarded to study the performance of the system? Would it be possible to include some kind of probabilistic assessment, or at least explain better why using such a crude assessment method? Would it be possible to include some kind of probabilistic assessment, or at least explain better why using such a crude assessment method?

Reply: We would like to thank the reviewer for this constructive comment and suggestion. Our main purpose is to present and validate the official product of Copernicus CMEMS for the Arctic region, which is provided as a deterministic reanalysis product based on the ensemble mean, for consistency with other CMEMS reanalyses. However, we fully agree that validation of the quality of the ensemble is crucial to prove the ability of the reanalysis to make the best use of a heterogeneous observational network (spatially, temporally and various data sources); for example that we do not overfit one observational data set at the expense of the others. The reliability of a system is important as well for an EnKF-based data assimilation system like ours, since the efficiency of the system relies on adequate assumptions for model and observation errors. Unfortunately, our storage facility is insufficient to store the full ensemble of the daily averaged fields, and we only have at our disposal the ensemble statistics of the variables assimilated at each assimilation time (every week). In order to address the

reviewer comment, we propose to extend our validation work with a reliability analysis (e.g. Candille et al. 2007, Desroziers et al. 2005, Rodwell et al., 2016) of the observation network assimilated (SST, SSH, Ice concentration, T-S). This metric will be used to assess the behaviour of our assimilation system in space and in time.

2. In assessing the performance by computing the difference with observations, the paper implicitly (and sometimes explicitly) assumes that the closer to the observations, the better the reanalysis. This amount to completetly neglecting observation errors in the assessment of the reanalysis, which is usually not an appropriate approximation. This incorrect assumption is for instance made explicitly in: âĂć p. 13, l. 4, where the misfit to observations is called "error" on the reanalysis; âĂć p. 14, l. 21, where the reanalysis is said to be improved if difference to observations is smaller; âĂć p. 16, l. 6-7, where it is said that an RMSD with observations of 5% is good; whereas the accuracy of the observations is said to be about 10%. In my view, this just mean that the reanalysis is excessively close to observations. I think that it would be important to better explain the limitations of this simple approach for assessing the performance of the reanalysis; to explain why more sophisticated comparison metrics were not applied (see my previous comment) and avoid the misleading expressions listed above.

Reply: We agree with the reviewer and the above statements will be revised according to the reliability analysis. We will add when possible the quantity sqrt(obs_error+ens_spread) to ensure that we are not over-fitting observations and that the ensemble does not collapse.

3. In the introduction, the authors provide several arguments to support the idea that ensemble methods are an appropriate way to apply the dynamical model constraint in the estimation process. However, this is not discussed anymore in the assessment of the performance of the reanalysis. Only quantitative difference to observations are provided and analysed. I think that the quality of the paper would be enhanced if more explicit evidence of what is stated in the introduction was provided in addition to the simple description of the distance between reanalysis and observations.

Reply: The overall text will be revised according the change proposed above.

[OSD](#)

Interactive comment

[Printer-friendly version](#)

[Discussion paper](#)