Ocean Science

Discussions

Open Access

Interactive comment on "How essential are Argo observations to constrain a global ocean data assimilation system?" *by* V. Turpin et al.

Anonymous Referee #2

Received and published: 10 August 2015

The study withholds Argo observations, which are the majority of in-situ real-time observations available for real-time global forecasting, to try to understand how important these are for 'constraining' water mass properties in order to produce more accurate global ocean forecasts. I have two main points for discussion. Firstly, the study mainly focusses on the impact on analysis error, which can be less meaningful than forecast error. This is because the observations can easily be overfitted to create an analysis with low errors and unphysical fields in the model, which can lead to a poor forecast which would imply dynamical adjustment indicative of an unbalanced analysis. Secondly, it usually takes of the order of years to spin-up a free model interior state to be relatively constrained to Argo data. Since the ocean interior changes slowly and is relatively immune to surface forcing at these timescales, there should be memory of the constrained water masses that persists for some time, even after when the ob-

servations are turned off. Here we would expect a spin-down time where the errors grow to something like double those from climatology and saturate around this level. This could actually have a similar timescale and I think this effect is probably embedded in the results of the experiments presented in this study. The improvements in the statistics presented suggest this to me. Whilst the study has merit, is only one of a few on the subject, and provides important information around the impact of Argo in ocean forecasting, it could do a more cleaner approach to either addressing the problem or framing the language around the experiments and results that is clearer on the limitations.

Further corrections/comments. P1147 L20: Worth summarising the conclusions of the studies that are cited, which have looked at the same problem. P1149 L2: ECMWF acronym wrong, change 'of to 'for'. L5-10: It would be good to get a clear idea about the sequential DA scheme, when analyses are done, what is the observation window in relation to the cycle, is it centred or asymmetric? L15-25: How were the observations processed prior to assimilation? What was done to account for measurement and representation error. Apart from the usual QC, were they converted into super-observations? With the in-situ data, how were they treated in the vertical to represent the model layers? P1150 L5: Some repetitive text regarding Argo P1151 L10-15: Regarding point 3, it seems like there were still other in-situ obs assimilated in this experiment, which would try to constrain the system. Were these sparse enough to have no impact on the results? Also, the experiments that assimilate SLA and SST without Argo still project information from the observations into the subsurface and influence the error. Was there any improvement in subsurface compared to the free run? If there was, this should be accounted for in determining the Argo impact. P1153 L14: In places there are mixed pronoun references, this needs to be made consistent. For example Antarctic Ocean and Southern Ocean, which are the same are used interchangeably. L20: How do we know that the salinity bias is not from a projection of SST and SLA into the model through the assimilation, rather than a model bias. Is the same bias in the noArgo as the free run? P1154: L12: Reduction of the misfits is obvious, shows that

the analysis is working. L22: The word 'current' is a pronoun and should be capitalised here and in other places in the text. Eg should be 'Aghulas Current' L25: Argo seems crucial to improve the model. . ..So far you have talked about improving the fitting of an analysis rather than improving the actual model. Also, all free models have substantial sub-surface errors when compared to Argo, so its essential that they are assimilated in order to improve the initial conditions for forecasting the ocean. Similarly as for other observations. It's the obvious problem of trying to forecast the weather without observations to initialise the state, its just not possible. P1155: L: The use of RMS in the analysis, which is more correctly written as RMSD, may not be as robust a measure as mean absolute deviation (MAD). To paraphrase a recent study 'RMSD tends to be dominated by a relatively small number of innovation elements with large magnitudes and may not accurately represent overall system performance in the whole domain. These elements may correspond to either less observed or more chaotic parts of the model, or be caused by observations with large errors – as the metric does not take into account the observation error.' L5: English P1157: L2-5: The term 'Heat Content' does not need to be capitalised as a Pronoun. L11-12: The error estimates change also as a function of the number of observations, which can make comparison tricky. L24: Think 'western boundary current' should be 'western boundary currents'. P1158: L21: What prevented the error stats being calculated in observation space rather the binning. If 2x2 degree boxes were done to make a spatial map of the error, it would be good to know how many observations went into each box in order to understand if there are sampling differences that may influence the interpretation of the result. P1159: L11: Remove 'region' after 'Southern Ocean'. L21-25: Forecast innovation error implies the calculation of the deviations in observation space using un-assimilated and independent observations. Its not clear from what is presented in the text that this is true. P1161: L1-5: As mentioned before it takes at least a year, but more likely several, to spin-up Argo into a global ocean modelling system. Are the ~1yr experiments long enough to get the right results? I would have expected to see errors go down from no Argo to full Argo greater than the overall 20% reduction. Usually the errors of all

variables are at least halved by data assimilation. P1162: L14: ACC - know what it is but it's not defined for the reader. L21: English

---