

Comment on [essd-2021-194](https://doi.org/10.5194/essd-2021-194)

Anonymous Referee #1

Referee comment on "Next generation of Bluelink ocean reanalysis with multiscale data assimilation: BRAN2020" by Matthew A. Chamberlain et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-194-RC1>, 2021

The manuscript presents a detailed description and validation of an ocean reanalysis, and it is very appropriate to accompany the dataset itself. The paper is well written, and explains very honestly the strengths and weaknesses of such dataset. I enjoyed reading it. As such, I recommend the paper for publication, but I ask the authors to better discuss, illustrate and explain several issues, summarized below. I think having a dedicated section "Discussion" before the Conclusions is a good option.

Thank you for reading the manuscript and the useful comments. The manuscript is modified to address the issues raised, please see below for details. Extra details have been added to the introduction which give further context to the work, and in a new subsection to the results.

General points (for which I suggest including a more detailed discussion).

- Neglecting sea-ice modeling in the reanalysis seems to me a weak point, considering the effects of sea-ice induced circulation around e.g. ACC, which is certainly an area of interest for the authors. A deeper discussion about what they expect to miss, and if there are plans for including a sea-ice model will benefit the paper.

A paragraph has been added to the introduction to give better context to the motivation and applications of the work presented. The development of BRAN is in support of operational ocean forecasting around the Australian region, and as such, dynamics close to Antarctica influenced by sea ice processes have not been a priority.

Previous versions of BRAN have been the basis of many studies, some of which are now listed in the introduction as well. The Bluelink Project does intend to include sea ice in future versions of BRAN to increase the utility of the product for research.

- The validation mostly focus on observation diagnostics. In my opinion a reanalysis is unique in the sense that can capture integrated diagnostics (OHC trends, transports, etc.) which are to some extent unobserved. It seems from the presentation of the work that climate applications are not the focus at all of this dataset. Also, the heterogeneous assimilated data (many datasets switching from delayed to real-time mode) may also compromise the low-frequency variability. Any thoughts about this, to include in the Summary/Discussion?

Previous versions of BRAN have indeed found many applications, e.g. to study transports around the Australian region (Schiller et al 2008, and, Divakaran and Brassington, 2011), extreme temperature events (Schiller et al. 2009, and, Oke and Griffin 2011), as well as providing boundary conditions to regional models (Steven et al. 2019).

This is now mentioned in the introduction of the paper.

This new version of BRAN is also suitable to the same applications, with the advantage of reduced biases.

As a demonstration, a new subsection is added to the paper comparing boundary currents around Australia in the new reanalysis with previous estimates, and finds that the results are entirely consistent,

The impact of heterogeneous datasets is unavoidable in the production of long reanalyses.

Under the "Analysis of innovation" section and discussion of Fig. 2, there is commentary regarding the impact of new observations entering the reanalysis.

The most striking feature is the reduction of mean absolute innovations for subsurface temperature and salinity once coverage of Argo data becomes global.

Also noted is the improvement in SST as new satellites and sensors become available; with VIIRS in particular in 2012.

- The DA system is detailed in a companion paper, not available at the moment (in review for Ocean Modeling). Because of that, some aspects of the DA system are probably not explained, so the reader

may have troubles to understand the formulation, if there is some lag between the two publications (see also below specifically).

This paper describing the DA system is now accepted (Chamberlain et al. 2021, doi:10.1016/j.ocemod.2021.101849) and publicly available. Some of the key results are emphasised again here. In particular, the impact of multiscale data assimilation at separate scales, which is demonstrated in this companion DA paper with a clean set of experiments.

Minor points

Line 4: it is said up to 2019. It will be useful for the interested readership to specify the plans for update (near real time, once in a while, or never?).

Yes, the intention is to update BRAN2020 to within months of real-time while it is our most current configuration. Comments have been added to the abstract as suggested.

Line 7; "for some variables" please be specific in the abstract.

Text is clarified here to specify that it is the subsurface temperature and salinity as the variables that are most improved.

The spinup period seems short (3 months). Do you have evidence that is enough for stabilizing most low-frequency variability indexes, or was chosen more for practical reasons?

We find that after 3 months or 30 DA cycles there is no further improvement or decrease in the innovations calculated, indeed, most of the decrease is within the first month. This is now clarified in the text.

We also avoid a long spinup to reduce the amount of drift and bias to build up in the ocean state before data is assimilated.

Line 72 vertical resolution seems quite low compared to most other reanalyses. It is also not clear (Line 142) how can the system assimilate both night and day time SST data, since the diurnal cycle will be for sure underestimated. Or maybe I am missing something in the explanation (lines 141-142 are not very clear to me).

The diurnal cycle is essentially removed from both the background and the observations. The background from the model is a daily average, so even the dampened diurnal cycle in the model is removed. SST data density is high, and the process of calculating super-observations averages both day and night observations. In addition, as stated, the "sea surface temperature at 0.2m" is assimilated rather than "skin temperature," also reducing any diurnal cycle in the data.

Line 74: Forcing fields masked if sea-ice. Not clear how do you use them if sea-ice occurs? Which fluxes would you use instead?

A short explanation is added, "values in the atmospheric reanalysis fields are replaced with values expected below sea ice..." This is applied in a preprocessing step before running the ocean model.

Line 108: it is clear from the text below that the ensemble anomalies in the EnOI are "climatological" (i.e. not flow-dependent). Better to specify here for clarity.

At line 108, it is clarified that there are two separate ensembles. In the text below it stated the coarse ensemble contains climatological anomalies; the paragraph now also states the high-resolution ensemble is built with seasonal-scale anomalies.

The multiscale Data assimilation formulation seems sub-optimal: no proper scale separation is used, and the use of the same observational data between the two steps implies non-zero cross-covariances between observations and background (in the second step). This seems theoretically sub-optimal and should be mentioned. Another issue is that the time dimension does not change between the two steps. It would be reasonable to assume that longer (broad scale) dynamics is associated to longer time scales, while here the 3-day time scale applies to both. Any thought about this? Again, maybe this is included in the manuscript submitted to Ocean Model., but without it being published it is worth to discuss.

In the paper that has now been accepted by Ocean Modelling, the method is shown to effectively apply corrections at different scales, since the anomalies in each ensemble have different scales. Features at all scales are present in the observations. The scales in the increments calculated at each step are determined when the DA

system projects the observation-model innovations onto ensemble members.

This is now discussed in the text at the end of section 2.3.2.

Broad scale dynamics might act on longer time scales. However, in practice, it was beneficial to apply the coarse DA correction at each analysis cycle due to drift and biases accumulating noticeably when the coarse DA step was not applied every cycle.

MDT (lines 155-157): I understand this is computed as Mean SSH from a free-running- model run. This means that long-term mean barotropic-dominated transports (e.g. in ACC) in the reanalysis should look very similar to the control experiment, by construction. Some comments about this will be beneficial, as most other reanalyses use other strategies (either an "observed MDT" or one with assimilation of in-situ data)

You are correct and comments are added to the text as suggested. This will be reconsidered in future versions.

Table 1. TEM vs TEM2 and SAL vs SAL2: the difference is not explained in the text.

Details added to text, "observation types with higher uncertainties, such as XBT and sea mammals, are assigned larger errors in types TEM2 and SAL2 in Table 1."

Line 180: Doesn't sound better/Isn't more common to say "super-obbing" with double "b"?

This short-hand term has been replaced with a fuller description, "to build super-observations."

Table 2,3,4. I really like the efforts in quantifying the error growth. Perhaps reporting all those values also in the figure 7,8 (in the profile panel) will help to better see the error growth and the differences with BRAN2016

The values in these Figures are calculated separately from those shown in the Tables. The error growth in Table 4 are differences in the magnitudes of forecast/background and analysis absolute errors from the DA step of the analysis cycle, whereas the profiles were calculated

separately based on observation-model differences from daily averaged ocean model output. These profile panels are already busy, and adding an error growth profile calculated from the DA cycle would also be inconsistent with the figure.

However, there is depth information still in the Tables with a breakdown of error growth near the surface (<50m), moderate depths (50-500m) and deeper (>500m). Values are consistent with the profiles, and are sufficient to compare and contrast error growth in subsurface temperatures and salinity.

Also, better to say immediately which observations are used to validate and form the statistics: are those assimilated in BRAN2016 only or also those supplemented in BRAN2020 (like sea mammals, etc.)? Both are possible choices in my opinion, but the interpretation of the results will be different.

The values in the tables are based on the observation databases assimilated into the respective versions.

This is now noted in the text in section 3.1.

Linked to this: lines 265-269 are not very clear to me. I don't understand why differences in skill scores improvements between surface and sub-surface data should lead to the authors' conclusions? It is because of the same atmospheric forcing/ingested surface data? Better to state clearly.

This last of paragraph of section 3.1 is rewritten to clarify.

We are attributing the improvement of the subsurface data assimilation to the multiscale technique based on results in Chamberlain et al. (2021a, doi:10.1016/j.ocemod.2021.101849), that is now published.

The comparison between BRAN2020 and BRAN2016 is complicated by other differences applied to the set up and observations.

However, in Chamberlain et al (2021a), a clean comparison was done where the only change was the addition of the second data assimilation step, and this showed similar improvements in the subsurface to those described in BRAN2020.

Line 323 typo in "reanalysed"

Fixed.

Line 331: SPINUP-EI is a misleading name. Perhaps change to CTRL or similar

Spinup-EI is one of the existing Bluelink spinup experiments that is also available from the NCI data catalogue, and has been used in other applications, so the nomenclature will be kept in this case.