Review of: Updates to the Met Office's global ocean-sea ice forecasting system including model and data assimilation changes

by Davi Mignac, Jennifer Waters, Daniel J. Lea, Matthew J. Martin, James While, Anthony T. Weaver, Arthur Vidard, Catherine Guiavarc'h, Dave Storkey, David Ford, Edward W. Blockley, Jonathan Baker, Keith Haines, Martin R. Price, Michael J. Bell, and Richard Renshaw

December 10, 2024

Manuscript Synopsis

The article is a system definition for the lastest version of the Met Office FOAM ocean data assimilation system. As such, this article is an important reference for testament of the current configuration of this system, and needs to be published. In general, the manuscript is well written and complete in its scientific rigour. I have only a few major comments that need to be addressed before publication. In particular, since the balance relationship scheme employed by NEMOVAR plays a critical role in the changes to the system, I believe a review of this scheme, previously described in the litereraturе, is warranted in preparation for its role in the changes. Although, for the most part, the changes and rationale for the changes are well written and easily followed with the evidence (figures) given, I did not find this as convincingly so for the addition of the Brunt-Väisälä criteria described in Section 3.2. The evidence (figure 4) is not as unambigiously positive as the authors would seem to suggest, and I feel a more detailed walk through this evidence to better point the reader to the benefit (or at least worst of the alternatives) that this change provides. A similar charge might also be leveled with regards to the benefits the updated version provides to the overturning circulation (latter portions of Section 4.3). Again the evidence (figure 14) fails to convince me of the positive impacts, at least not without further description – although in this case, at least part of the problem may be due to figure presentation deficiencies.

My recommendation is Minor Revisions

Major Comments

1. A large portion of the justification for removing the longer covariance lengthscale from the balanced portion of temperature background covariances – and therefore sea surface height and salinity increments – is based on the balance operators used in the NEMOVAR assimilation scheme. I believe a more detailed reminder of what those balance operators are beyond the stated "physically-based balance relationship" (l. 146) is required. The balance relationships of dynamical height adjustment and maintenance of neutral density, are critical to understanding why the longer temperature covariance lengthscale may be sub-optimal, but the contribution to an unbalanced salinity increment might have lesser negative effects. Although this might be obvious to the authors, it is sometimes necessary to lower the conversation to the readers potential ignorance. While the Waters et al. [2015] reference (BTW: An addition of the DOI to the reference would be useful: https://doi.org/10.1002/qj.2388) does extensively discuss this, I believe a more explicitly stated reminder would be useful here, possibly further referencing some of the earlier linearization work. Particular emphasis should be placed on why the introduction of a large scale covariance on temperature might bleed corrections to the unbalanced, barotropic sea surface anomaly (SLA) into the sub-surface temperature. By the way, the statement (l. 188) that "large-scale error covariances have a 400 km length-scale for temperature, unbalanced

salinity and unbalanced SSH" is correct [Mirouze et al., 2016], but it is also confusing! With the start of that sentence discussing the shorter length scale convariances, this statement somehow suggests to me that unbalanced SSH also has two convariances length scales, when in reality, it only has the longer, 400 km length scale.

2. I believe Figure S4 should be elevated to the main article. Discussion of why the SSH balance relationship is now applied throughout the mixed layer in GOSI9 (or why it was turned off through the mixed layer in GO6), is quite brief, and seems quite conclusively demonstated by Figure S4. Furthermore, it is equally, if not more, important, than the subsequent discussion of then mitigating some of the negative effects of that change (along with the removal of the long temperature covariance lengthscale) via the Brunt-Väisälä criteria in subsection 3.2. What is not expanded on, however, is whether the change to applying the balance throughout the mixed layer is related to the change to a single temperature covariance length scale. (I.e. I presume previous versions of the system better performed when the this relationship was not applied in the mixed layer.)

3. I found the rationale for applying the Brunt-Väisälä criteria to determine if the increment should be retained as less than convincing, at least with the current discussion and presentation of Figure 4. Showing the change of mean state at 1200m is a little confusing, as judging from the mean error shown in the profile of Figure 4d, it is not immediatedly obvious that the mean state of GOSI9+BV at that level is better than GO6 – and it is only just improving from being worse even then GOSI9 without BV slightly higher in the water column. I would think it might be better to highlight the mean temperature somewhere around 500m, where you can show that a more accurate mean state is achieved by **not** retaining the increment – and thus possibly increasing slightly the RMSE, although this latter part of the argument depends somewhat on whether the OMB statistics displayed are from the IAU phase (assumed), or the trial phase. This then can also lead, by not instigating spurious deep convection, to better RMSE and mean error lower down (below 1000 and 1300m respectively) in the column. I do not believe it is completely accurate to describe the mean state as being improved from 250m and below (l. 247), as it is fairly obviously degraded at 750m, at least relative to GO6. Finally, the Figure S6 is perhaps more definite in showing the improvement – although here the cost (increased RMSE) of not retaining the increment at certain levels is not as obvious. I would keep with the current discussion of the improvement, when well described, in the Labrador Sea deep convection region, but note to the reader further evidence for the Mediterranean overflow region is shown in the supplementary material.

4. Finally, I found the discussion of GOSI9 improving the overturning circulation (AMOC) in the latter portions of Section 4.3 (ll. 493 – 509) also not particularly convincing, at least by how it is served by Figure 14. The principle conclusion of section 4.3 in general, is that GOSI9 is more stable to the lack of sub-surface observations, or more specifically, mitigates the negative consequences of sea level anomaly assimilation in the absence of correcting sub-surface observations. The discussion on the AMOC attempts to then extend this improvement when lacking sub-surface profiles to the AMOC. This is convincingly demonstrated for the AMOC at 26.6°N (Figure 14a), where the divergence of the non-profile assimilation run with regards to the underlaying RAPID observations is convincingly shown. This is less convincingly shown at 50°N (Figure 14b) in the absence of underlying observations, which could be rectified by instead considering overturning along the OSNAP section, for which observations do exist, and the FOAM system has previously been evaluated [Jackson et al., 2019]. This is even less convincingly demonstrated by the the AMOC at 30°S, where the GOSI9, no T/S profile assimilation (cyan) line appears to be the outlier instead of GO6 no profile (orange), and appears to be going in the unwanted direction of too small, and even negative southern hemisphere AMOC that resulted in FOAM not being used in the Mignac et al. [2018] analysis. I would encourage the authors to at least include an AMOC estimate along the OSNAP section as previously shown in Jackson et al. [2019], with accompanying observational estimates, which should exist for the examined 2019.

5. Furthermore, I disagree with the implicit statement – it is not actually stated, but just somewhat implied by the discussion – that GOSI9 might improve (ll. 552-556) the "discrepant" equatorial overturning FOAM demonstated in GO6, which is only somewhat, and not conclusively, demonstated through Figures 14d-h, or the "unrealistic" southern hemisphere AMOC suggested by Mignac et al. [2018], which is not demonstated at all in Figures 14d-h as the figure does not seem to go as far south

as 35°S in order to compare with Mignac et al. [2018]. Although as stated (ll. 501-503) some of this might await a proper evaluation of a FOAM reanalysis. .

Minor Comments

1. The manuscript mentions several assimilation method and ancillary input changes (ll. 71-74), along with the testing of new observational types (ll. 72-73). Most of the assimilation method and ancillary information changes have been implemented (or continue to be used) in GOSI9, but none of the mentioned new observation types are used – for the most part, due to the fact that they would not be operationally available. It might be interesting to return to this in the summary (i.e. could SSS assimilation [Martin et al., 2019] further allow removal of the longer salinity covariance scale; could U/V assimilation [Waters et al., 2024] improve equatorial transports, etc.).
2. **All** Figures: It is not mentioned whether the observation minus background statistics shown throughout the manuscript come from the trial run or from the IAU run. It is known FOAM calculates both, and it would be useful to know which of the two is shown throughout the manuscript – assuming it is not a mix of the two. This would help with the interpretation of figures such as Figure 4; it would be natural for RMSE of the IAU step to increase at levels where the increment is not being applied.
3. (ll. 230-232) The adjustment of melt pond variables to changes in SIC DA changes seems quite important to improvements/changes in sea ice for GOSI9 (ll. 393-403). It is also a rather ambiguous statement. Further details would be warranted/appreciated.
4. ll. 194-195, or more generally the whole discussion of 2 versus 1 covariance length scales (ll. 179-198). It would be useful to remind the reader why two covariance lengthscales was implemented, which was primarily to correct the near-surface drifts seen in the southern ocean (Figure 2e-f) and Mirouze et al. [2016].
   - In light of only near surface differences between the 1T-1S and 1T-2S covariance lengthscale schemes, could a near surface only (damped/cutoff with depth) long length scale S covariance work even better?
   - The only difference in T-profile statistics seems to be a very deep (> 1500m) increase in global RMSE of 1T-1S over 1T-2S (Figure 2a). This perhaps contradicts my statement immediately above. But more for the sake of curiosity than anything else, which region is responsible for this increase – as it is not the southern ocean.
5. ll. 332-333, Section 4.2. It likely should be mentioned that the SLA degradations found in coastal regions and enclosed seas are due to model changes (Figure S7c), although degradations in the central and eastern North Atlantic, or Mediterranean outflow regions would be due to data assimilation changes (Figure S7e).
   - The Figures 5c/d and S7e/f should be identical, but are not? Why? In particular, the degradations in the Mediterranean Outflow region in Figure 5d seem worse than those in Figure S7f.
6. The model changes include changes to the fundamental state variables of Temperature and Salinity. Therefore there are changes to the observation operators, and to the balance operators of NEMOVAR. Given that observation minus background (OmB) statistics are being used to show differences between the two versions:
   - Can OmB statistics be fairly compared? The GOSI9 model only SST changes seem universally degraded, especially in regions of large potential T vs conservative T differences (Figure S1).
   - Can changes to SLA OmB statistics possibly be attributed to changes in the dynamic height balance operator? Like in the coastal and enclosed sea regions?
7. Figures 6&7. A spatial map of OmB RMSE for depths between 300-700m might also be useful to visualize the improvements to the system upgrades. Note: Actual depth range was chosen somewhat arbitrarily – but differences spanning 500m depth seem to be relatively global in breadth.

8. In light of Mignac et al. [2022] I am curious as to why / disappointed that no evaluation of sea ice thickness was performed.

9. The no profile experiments answer the systems ability to operate in the pre-Argo period (1993-2005) with the presence of satellite altimeter and SST observations. However, it does not answer to the systems ability to operate in the absence of altimeter observations prior to 1993 – or a transition from no altimeter to altimeter observations in late 1992 (as for instance shown in Figure S3), particularly in the absence of correcting T/S profile measurements.

   - It may be necessary to caveat statements of the usefullness of the ocean reanalysis to a particular time period (1993 onward), as "GOSI9 will lead to more potential use of Met Office ocean reanalyses in climate studies" may have different envisioned time scales depending on the user (seasonal/decadal/century scale).
   - Nevertheless, it might be worth commenting on the possibile usefullness of this GOSI9 version of the Met Office FOAM system to long, century(?) scale reanalysis – or defer that discussion to a manuscript dedicated to an ocean reanalysis.

Presentation Comments

1. I do not find Figure 1 particulaly useful. A statement that the SLA observation error in GOSI9 is larger than that used in GO6, with the difference largely attributed to the use of 4cm measurement error might suffice. The figure itself could be demoted to the supplementary material. Personally, as mentioned above, I would rather see Figure S4 promoted to the manuscript.

2. Profile Figures. I believe it would be useful to have profile plots explicitly (in figure) labelled with a (largish) T or S (presumably in the white space normally found from mid-depth downward on the right side of most T and S profiles. This would rectify the necessity to refer to the caption all the time (although admittedly, those more familiar with the plots will immediately identify the differing shapes of T and S profiles).

3. Figure 2. I had difficultly, at least on the printed page, to differentiate the '2' and '1' in the labelling. You cannot improve my eyesight, but the labels could be comfortably enlarged somewhat. Improving even further beyond that might entail changing the shape of the font used 1T-2S (helvetica) may be more distinguishable than 1T-2S (computer modern).

4. Figure 3. One needs a Ph.D. in colour combinations to identify the regions. Caption note would seem suffice to warn the reader, but confusion will still remain. I have no suggestions, however.

5. Figure 14. Vertical span of transport strength is not sufficient to adequately distinquish the differing experiment lines. In general, the figure seems oddly arranged. I would think a 3 across set of time series above a 5 across set of Hovmüller diagrams would work better, allowing the vertical transport strength axis to be stretched to near the size of the horzontal time axis. For time series with observed data points, an accompanying set of correlation coefficients would be useful. I am aware such an r-coefficient would **only** distinquish the models ability to correctly follow the observed seasonal cycle – which in turn is largely Ekman wind driven (i.e. will largely be identical across all experiments, and the observation), but this may still be useful.

References

L. C. Jackson, C. Dubois, G. Forget, K. Haines, M. Harrison, D. Iovino, A. Köhl, D. Mignac, S. Masina, K. A. Peterson, C. G. Piecuch, C. D. Roberts, J. Robson, A. Storto, T. Toyoda, M. Valdivieso, C. Wilson, Y. Wang, and H. Zuo. The mean state and variability of the north atlantic circulation: A perspective from ocean reanalyses. *Journal of Geophysical Research: Oceans*, 124(12):9141–9170, 2019. doi: https://doi.org/10.1029/2019JC015210. URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JC015210`.

Matthew J. Martin, Robert R. King, James While, and Ana Barbosa Aguiar. Assimilating satellite sea-surface salinity data from smos, aquarius and smap into a global ocean forecasting system. *Quarterly Journal of the Royal Meteorological Society*, 145(719):705–726, 2019. doi: https://doi.org/10.1002/qj.3461. URL `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3461`.

D. Mignac, D. Ferreira, and K. Haines. South atlantic meridional transports from nemo-based simulations and reanalyses. *Ocean Science*, 14(1):53–68, 2018. doi: 10.5194/os-14-53-2018. URL `https://os.copernicus.org/articles/14/53/2018/`.

Davi Mignac, Matthew Martin, Emma Fiedler, Ed Blockley, and Nicolas Fournier. Improving the met office's forecast ocean assimilation model (foam) with the assimilation of satellite-derived sea-ice thickness data from cryosat-2 and smos in the arctic. *Quarterly Journal of the Royal Meteorological Society*, 148(744): 1144–1167, 2022. doi: https://doi.org/10.1002/qj.4252. URL `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4252`.

I. Mirouze, E.W. Blockley, D.J. Lea, M.J. Martin, and M.J. Bell. A multiple length scale correlation operator with application to ocean data assimilation. *Submitted to Tellus*, 2016.

Jennifer Waters, Daniel J. Lea, Matthew J. Martin, Isabelle Mirouze, Anthony Weaver, and James While. Implementing a variational data assimilation system in an operational 1/4 degree global ocean model. *Quarterly Journal of the Royal Meteorological Society*, 141(687):333–349, 2015. doi: https://doi.org/10.1002/qj.2388. URL `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2388`.

Jennifer Waters, Matthew J. Martin, Michael J. Bell, Robert R. King, Lucile Gaultier, Clément Ubelmann, Craig Donlon, and Simon Van Gennip. Assessing the potential impact of assimilating total surface current velocities in the met office's global ocean forecasting system. *Frontiers in Marine Science*, 11, 2024. ISSN 2296-7745. doi: 10.3389/fmars.2024.1383522. URL `https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2024.1383522`.