Review of "Contribution of a constellation of two Wide-Swath Altimetry Missions to Global Ocean Analysis and Forecasting" by Mounir Benkiran, Pierre-Yves Le Traon, and Gérald Dibarboure

The manuscript is generally well written and understandable. The authors present results of several ocean prediction OSSEs to simulate the impact of two swath altimeter satellites in combination with three nadir altimeter satellites. The results are well presented. The results are significant and will be of interest to the science community for determining the importance of future investment. The quality of the work is high. The results are based on realistic ocean forecast systems being applied operationally. These results build on many prior studies on the impact of satellite and in situ observations. Most the comments I have are for improving the clarity of text and presentation. There are a few points that may require further consideration to resolve, though it is certainly possible I did not understand the intent of the text clearly, and in this case it may be relatively easy to resolve. The significant points are highlighted in yellow below.

7: "and has convinced more than thirty thousand expert services and users worldwide" – Do you mean that there are 30,000 users worldwide?

18: "so there is a strong diminishing return associated with classical altimetry" – we have seen a linear improvement in spatial correlation in properties associated with fronts. Surface divergence and frontogenesis forcing in particular increase linearly from 1 to 4 nadir altimeters. So the statement on diminishing returns depends on which properties are examined. For ssh, the statement is certainly correct. Clarification would be helpful.

P3, 17 – "it is imply impossible to resolve the time scales of the small-scale features" That is a bit of a strong statement that could use some context. It applies to submesoscale instabilities in the open ocean. Arguably, tidal variability in coastal areas will be resolved in space, and though aliased in time SWOT observations should provide new information on small scale tides. Help the reader understand the particular ocean features to which the statement applies.

P3, 20 – "can be observed with a mean revisit time of 5 days or less" – What is the source of the 5 day limit? Energetic areas such as the Gulf Stream or Kuroshio exhibit large changes even over 3 day time scales.

P3, 26 – "good enough for tidal aliasing despite being sun-synchronous" – That depends on application. It is not necessarily good enough for diurnal tides.

P3, 35 – "spectral noise floor of 2 cm²/cycle/km" – The SWOT requirement should be stated clearly. Leaving the statement as this leaves the reader to assume this applies in the along track and cross track directions and that the noise is uncorrelated. The correlated error levels are very important. These have specific correlation shapes in the across track direction with spectral power that is not white noise in the along track direction.

P4, 14 – "In our study, we use the same NEMO 1/12° resolution model but with different configurations and forcings" – It would be helpful for readers to have a few additional details on the model system at this point. What is the vertical resolution? This would help inform the reader as to the realism of the vertical structure of features in the nature run. What surface forcing is used at what temporal frequency? Is tidal forcing used in the system? What is the turbulence closure representation? Since the nature run is a free run, what is the realism of the circulation and energetics?

P4, 25 – "or Sentinel 6, which will use the same orbit" – verb tense. Sentinel-6A already is in the same orbit.

P5 – 23 – "OSSE2 (not presented here) is similar to OSSE1 except that it assimilated Sea 23 Surface Height (SSH) from two Wide-Swath Altimeter (2S) datasets instead of nadir altimeter data." – Why not show results from OSSE2? They would be valuable to put in context of the other results. Are the results very similar to those of one of the other OSSEs? Please point out which if that is the case.

P6 – 8 – "The SSH variance in the NR compares very favorably with real altimeter observations" – It would be helpful to see the ratio of the NR variance to the observed variance. Fig. 2 of Benkiran 2021 is very small, and there are areas where the NR has greater variability than DUACS and areas of less variability. The North Atlantic in the Irminger and Labrador Seas certainly is quite different. Rather than a qualitative statement of comparison, a plot of the ratio would be a quantitative comparison.

P6 – 11 – "SSH variance error" – This might be better worded as "variance of the SSH error" or "SSH error variance". I see the notation changing throughout the paper. It would be good to maintain consistency throughout.

Table 1 – The VAR* is the VarError divided by the variance of the NR if I understand correctly. However, the OSSE1 and OSSE3 values seem to be normalized differently. For example, 21.2/15.6=1.359, 14.1/10.1=1.396, 30.4/24.8 = 1.225, 21.3/17.0 = 1.253. Apparently I do not understand how these were computed. The text on page 6 should be more clear.

P6 – 38 – "compares the error variance of the different OSSEs for wavelengths smaller than 200 km" – This is a good way to evaluate the change in skill depending on feature sizes. My question is how these scales were extracted from the OSSE and nature runs. Was it by a spatial convolution filter, and if so what are the properties? I am wondering how sharp the filter cutoff may be and also how fields near land are filtered. Similarly, what filtering was used to separate time scales less than 20 days?

P7 – 21 – "Box D in Figure 5B" – There is not a box labeled as D on Figure 5b. There is a box labeled D on Figure 3. Boxes C and D in Fig 3 are located in the North Atlantic near the Iceland Basin and Irminger Sea. Figure 2 of Benkiran 2021 compares the nature run (left below) to Duacs (right below) showing much more variability in the nature run than in the observations in the areas of boxes C and D of Figure 3. If the model is over energetic, the impact of the satellite observations could be over estimated.
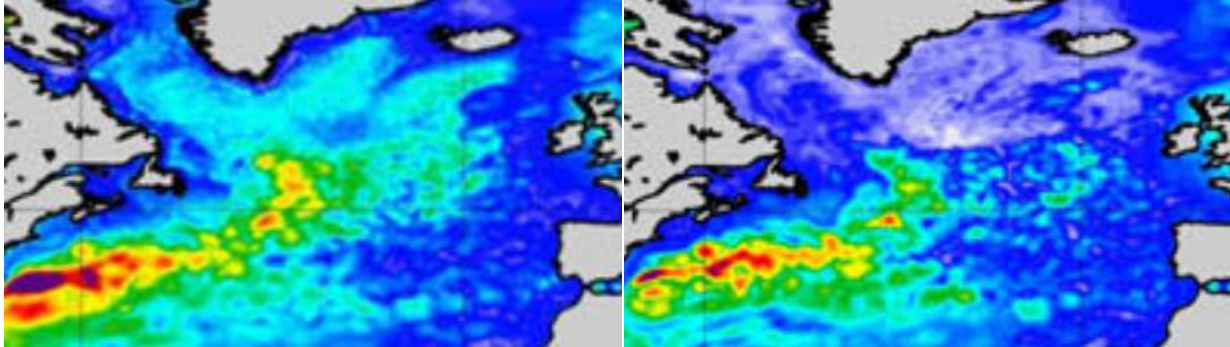
P 7 – 27 – "This is mainly due to the weak signal in these regions and the limited space/time sampling of the nadir altimeter constellation at these wavelengths." – I do not expect this is correct. Region A in Figure 3 is near the equator where spatial scales are typically quite large. The primary features are large eddies that are typically well resolved by 3 nadir altimeters.

P 7 – 30 – "The contribution of the swath altimeter data contributes to a clear reduction in the error spectra in all these regions for wavelengths larger than 50 km." – I do not reach that conclusion from Figure 10. Figure 10 area A indicates the 3N+2S deviating from the FR around 120 km scales. Area B is near 80 km, area C is near 50, and area D is near 70 km.

P 7 – 33 – "The reduction of the error at the different wavelengths (ERspec) is defined as the percentage of the error with respect to FR (OSSE0)." The comparison of error reduction from the FR to OSSE1 and OSSE3 is somewhat misleading. There certainly is reduction, though the FR error variance should be larger than the NR variance. The predominant processes affecting SSH variability are eddy instabilities. The processes that are deterministically forced by winds would be expected to be a much smaller fraction of the variance. If the eddy features in the FR are stochastically positioned relative to the features in the NR, the FR error variance can be up to twice the NR variance. This can mislead interpretation. For example, if an OSSE reduced error variance to half the error variance of the FR relative to the nature run, the OSSE could have no skill. The OSSE error variance could be equal to the NR variance.

It would be helpful to have maps of the FR and OSSE error variance divided by the NR error variance. That would more clearly demonstrate where the OSSE has skill and would still convey the same information on the errors in the OSSE relative to the FR. The plots of coherence in Figure 11 do help alleviate this problem to an extent. It should be addressed in with more clear plots.

Figure 11 plots C and D should have the horizontal axis range extended to show the point at which the black lines cross 0.5.

P 8 – 17 – "The calculation of this coherence was based on filtered SSH fields of scales greater than 500 km to avoid the impact of large-scale and high frequency signals on the results" – Does this imply scales greater than 500

km were removed before computing the time coherence? On first reading I interpreted it the opposite manner. More clear wording would help the reader.

P 8 – 29 – "This error is significantly reduced by assimilating the nadir altimeter data (black profile) compared 29 to the free model (FR, orange profile)" – Again, the variance of the FR – NR is misleading and will be larger than the variability of the NR. It would be useful to plot the NR variance of temperature and salinity on these. A similar comment pertains to Figure 14 of the velocity errors. The NR variance should be plotted on Figure 14 as well.

Returning to an earlier comment, this is a point at which the details of the assimilation are very important. It is not explained how the SSH translates to a subsurface temperature and salinity. If the covariances are not accurate, this can be a source of the lack of impact of the swath data. As was done with the SSH in Figures 10, 11, and 12, the spectra of the temperature and salinity are valuable to understand the scales at which skill is gained. The general science community will be more interested in these variables than the SSH.

P 9 – 25 – "Surface current forecast errors should be equivalent to today's surface current analysis errors or alternatively will be improved (variance error reduction) by 30% at the surface and 50% for 300 m depths." – This is a confusing statement. Why should the forecast errors be the equivalent to today's? The results show the reductions. We should expect reduction in forecast current errors.

Figure 15 – the Titles at the top of the panels seem to have an error. Three of the panels are titled "U Variance Error".