

Ocean Sci. Discuss., referee comment RC2
<https://doi.org/10.5194/os-2021-89-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on os-2021-89

Anonymous Referee #2

Referee comment on "Formulation and demonstration of an extended-3DVAR multi-scale data assimilation system for the SWOT altimetry era" by Zhijin Li et al., Ocean Sci. Discuss., <https://doi.org/10.5194/os-2021-89-RC2>, 2021

Review of "Formulation and demonstration of an extended-3DVAR multi-scale data assimilation system for the SWOT altimetry era" by Li et al

Summary

This paper claims to introduce a DA system suitable for assimilating wide swath altimetry data (SWOT). I found this a disappointing paper. First, the paper misleads the reader by giving the impression that assimilation of SWOT data will be explored or tested. In fact this is not the case and the discussion is insufficient to convince the reader that this system will be capable of assimilating SWOT. In particular, a system should be capable of assimilating data with the particular error characteristics of SWOT data. Particularly of concern are correlated errors in the swaths due to phase, roll and timing errors. Second, the experiments performed are unilluminating. The paper is really about demonstrating extended time window 3DVAR. This needs to be compared to a control experiment doing non-extended 3DVAR otherwise the results are of little interest. At the moment I cannot tell if the extended 3DVAR makes things better than your non-extended 3DVAR.

The paper, in my view, goes into too much detail in aspects that were not novel and glosses over aspects that were potentially novel. An example of this is spending a lot of time on the characteristics of the existing observing network for example SST and at the same failing to go into sufficient detail on how the "time sampling" error is determined. Similarly the discussion of the results is superficial.

I recommend rejecting this paper for the above reasons. I give some more detailed comments below.

More detailed comments

The title is misleading it is not convincing that this system addresses the assimilation of SWOT data. Just removing "for the SWOT altimetry era" from the title would avoid misleading the reader. By "extended" you mean "extended time window" it would be clearer to write that.

In the abstract you write your system "is first to assimilate routinely available observations" is misleading. Is the first for your system? It is certainly not the first in world to do this.

Line 93: Error in the citation Li et al (2019). There is a Li et al (2019a) and Li et al (2019b) which one is it?

It is highly misleading to call the system MSDA-SWOT when you don't appear to have tested it with SWOT data. I feel strongly that it is not acceptable to do this.

Figure 1b with the Google Earth background is very unclear.

Section 2.3 First sentence. Use of gridded and along track altimeter would mean you are assimilating the same data twice because the along track data is used to generate the gridded data. This needs clarification here.

Figure 2. explain in the caption what the ring of 6 black circles at about 36 N 125 W

Line 148. Gridded products will also distort the positions of eddies. They are produced using along track altimeter data. Indeed they may be more likely to do so since they do not use a model to evolve the eddy positions from a previous analysis.

Figure 4. A plot of the observations is really unnecessary here. The reader will be thinking "when am I going to see something about SWOT data"? The answer of course is never.

Section 2.4 line 178. You should write what you do in this paper that is not done in Li 2019b. It is unclear at this point whether the data assimilated or just used for validation in the work presented here and is it IR SST and/or Microwave SST (line 180).

Section 3. lines 191-95. It is not clear why a 1 km or finer resolution is necessary to assimilate SWOT. I suppose you might say a km scale model is needed to take full advantage of SWOT but it is not necessary to assimilate SWOT data. Anyway your model does not meet the claimed requirement having a resolution of ~3 km.

Section 3. lines 212-215. and Section 3.1.2. It is a good idea to account for the sampling time error but note this is not a new approach (see below). It should be noted first that errors associated with this are not random or Gaussian since they are associated with not accounting for model evolution. Second the impact of this will be to reduce the weight given to observations far from the analysis time which may consequently have only a small impact on the analysis. Assimilating such data if not carefully done may degrade the analysis in some cases.

Section 3.1.2. The sampling time error is far from a new type of observation error - Oke et al (2005) and Martin et al (2007) both increase the observation error with based on the difference of observation and analysis time. Additionally, since this is one of key novel aspects in your system this section requires much more detail. Questions not answered are what is the form of the time sampling error how does it vary with the analysis observation time difference? How are any parameters determined?

Section 4.1. Second sentence. Delete "Here we give a description". Not necessary since you are giving a description in the next sentence.

Section 4.2. There seems to be a fundamental misunderstanding in what the background error is. It does not relate to the observation scales instead it is the scales in the model forecast (the background) error. I would concede this is affected by previous observations assimilated in previous assimilation cycles, but not in the straightforward way you claim here. It is not clear what you mean by in lines 311-313. This should be clarified and perhaps explained with a specific example.

Section 4.2 line 325. It would help the reader to describe in words what the terms in equation 7 mean. What is the justification for (7) particularly including the measurement/instrumental error in the calculation of the sampling time error? This error is due to not including the model time evolution in the calculation of $H(x)$ and this should not be affected by the instrumental error. I suppose there is a possible argument for time sampling errors being associated with the representativeness error since this relates to how well the model equivalent represents the observations. This may be higher where the model is more variable and therefore where bigger errors would be due to assuming the model is the constant in the cost function. This whole discussion could do with more detail since this is what makes this work different to previous work you have done. Describing exactly how λ_k is determined would be interesting.

The paragraph lines 336-342 is very unclear. I think my problem relates to my fundamental issue with this whole section. I do not see that you can just specify the background error correlation scales based on baseline requirements from SWOT it should related to what the actual background error is in your system.

Table 2 is confusing. You list observations, but then it appears that many of the types are not assimilated in the experiments here. In that case they should not be a table with the caption "observations assimilated".

Table 2. It does not seem good practice to assimilate maps created from along track SSH and then assimilate the along track SSH data. You are in affect assimilating the same thing twice since the along track SSH data was used to generate the map. You should at least highlight potential issues with this approach not least overfitting to the observations.

Section 5. To convincingly demonstrate the utility of extended-3DVar but you need to run a control experiment with non-extended-3DVar. It would also be useful to have another experiment to assess the impact of the optimisations you mention versus a control run.

Section 5.1. Assessing the assimilation analysis against observations it has assimilated to produce the analysis is not very useful. The results may look good but you may in fact just be overfitting to the observations and their associated errors.

Figure 6. Describe the significance of the SWOT baseline line compared to the nadir altimeters observations. It has more power at larger wavenumbers/ smaller wavelengths. Is this signal or noise? Is there a significance where the lines cross over? Why is the power so much less at smaller wavenumbers/longer wavelengths?

Figure 7. Anomaly correlations are much more interesting that correlations this largely just tells me that the mean fields match. This needs another experiment to compare to. Explain why the RMSE grows with time. At the moment I might think that you started your DA experiment with an analysis from your old system and made the results worse with the extended-3DVar. Perhaps seasonal effects are causing the increase in error with time but I cannot tell from the results you show.

Line 408 A change from 2.6 cm to 2.8 cm does not seem substantial to me.

Line 436. Saying the domain average RMSD is as large as 10.0 cm is imprecise language just say what it is averaged over the comparison time period.

Line 440. "As much as 14%" (is it 14% or not?)

Figure 8. Again what is the little black ring of circles in each plot. It should be stated in the caption in each figure it appears in.

Figure 9. I really think you need to give anomaly correlations for SST since it is quite easy to match the climatology and achieve a very high correlation. The (non-anomaly) correlations will be very high even for a non-assimilating run. This plot also needs another experiment to compare to. Again the error grows with time so with no other experiment the reader may conclude that your changes may be making the results worse.

Figure 10. Not really discussed in detail. What is the significance of the shape of the histogram, for example? It again could really do with another experiment to compare the results with. Also the anomaly correlation should be used as the correlation is not very useful to SST as I explained previously.

Section 5.4 I think this idea of the "campaign area circulation" is a potentially interesting one. But the exploration of it here is superficial. I think this would be a good place to add more figures. It would be interesting to show an example of this and how your work has reduced such errors. It seems a bit trivial to say if you assimilate other observation types and keep your analysis close to the truth then the increments from campaign observations will be smaller. An illustration of this with results (perhaps showing surface currents) from a run where you exclude another data type so that analysis is not as close along with your campaign data would be useful to see.

Lines 475 I'm not really sure the comparison to no data assimilation is particularly interesting it is unlikely that someone would assimilate campaign data and fail to assimilate other observational data.

Line 492 Briefly explain here again what the DA Cal experiment is. What does "Cal" stand for?

Figure 11. You are comparing against observations you assimilate I'm not sure this plot is particularly useful since it is quite easy for a DA system to fit data it is assimilating. It certainly doesn't illuminate on the "campaign area circulation" idea.

Figure 12. This figure is a mess of overlapping lines there doesn't appear to be any consistent differences between the experiment results. The discussion of this figure does not help in at all to explain what the reader is supposed to conclude from this figure.

References

Oke, P.R., Schiller, A., Griffin, D.A. and Brassington, G.B. (2005), Ensemble data

assimilation for an eddy-resolving ocean model of the Australian region. *Q.J.R. Meteorol. Soc.*, 131: 3301-3311. <https://doi.org/10.1256/qj.05.95>

Martin, M.J., Hines, A. and Bell, M.J. (2007), Data assimilation in the FOAM operational short-range ocean forecasting system: a description of the scheme and its impact. *Q.J.R. Meteorol. Soc.*, 133: 981-995. <https://doi.org/10.1002/qj.74>