Response to first reviewer (Graham Quartly)

<u>Response:</u> Thank you for your review, we have complied with most of the changes you've requested, except for reassessment of the S3B drift and providing an uncertainty assessment on crossover approach. They will both be evaluated in further studies and will therefore be the object of other future publications.

Kind regards, Rémi Jugier

Sea level rise, and especially it's global average (GMSL) has been seen as one of the totemic measures of anthropogenic climate change, with accurate observations of the trend key to distinguishing between different scenarios. Satellite altimeters are the only means to provide a fully global picture, with similar trends noted from most missions (after much detailed analysis to understand instrumental effects!) The switch to delay/Doppler (or "SAR altimetry") for recent missions has raised questions about the consistency of trends from different satellites.

This paper provides some new ideas on how to assess the relative trends of missions, although it cannot identify which is the most accurate in an absolute sense. The method advocated is to interpolate each mission's data on a lat-lon-time grid, and then apply suitable area-based weighting to find the differences in trends. The main results are not new i.e. that S3A and S3B differ and show different trends to Jason-3. The new part is the discussion of methodology and the uncertainties.

It is unfortunate timing that this paper was submitted a few weeks before Dinardo revealed the cause of S3B's anomalous trend (incorrect application of USO correction). As a simple daily correction has been provided to undo the USO error, it would seem sensible to implement this and test whether there is now negligible drift between S3A and S3B. It could help if you mention the S3MPC overview paper (Quartly et al., 2020) noting the telemetered USO correction (Fig. 4b) and the other independent means used to monitor drift for S3A and S3B.

In the revised version of the paper, for S3B, we now explain the main cause of the drift (incorrect application of the USO correction), and link to Quartly et al. 2020. However, we

do not reassess the drift with the proper USO correction. As it was this study was the one (or one of) that initially managed to assess this drift, we think it is better to leave it as it is for proper tracking of the scientific process, and the assessment of the S3B drift with the correct application of the USO correction will be the object of future studies.

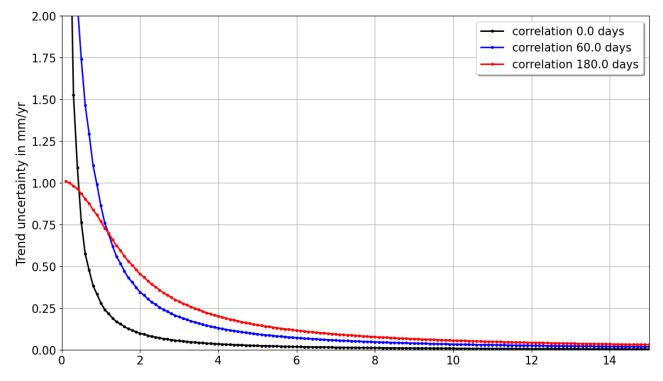
An alternative way to assess trend differences between just two missions e.g S3A and Jason-3) is to consider all dual crossovers < 10 days apart (which the authors briefly mention at the very end of the paper). I would like to see more discussion of the pros and cons of each method, along with a comparison of the uncertainties in each. It is not immediately clear that this new method is better. Indeed, we can consider crossovers < 10 days, or even < shorter periods. This approach is also very relevant : it allows for improving the spatial correlation between the measurements of the 2 missions but at the cost of fewer measurements being averaged. At this stage we have not yet evaluated the uncertainty of this approach to detect sea-level relative drifts between two altimeter missions. We plan to do this work on future studies with a similar statistical method to the one described in this paper. We would like to update this paper with a more general study that compares different cross-calibration methods together.

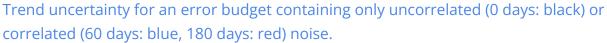
The discussion of how uncertainty varies with duration of the common period is very useful. In order to achieve accuracy in accord with GCOS requirements should the space agencies aim at 15-year missions rather than a series of satellites with 5-7 year lifetimes?

For Cal/ValL activities like the sea-level drift assessment presented in this paper, we need to have missions that are concurrently flying on a stable orbit track for the longer period possible. So, having short lifetime missions limits our ability to detect drifts, which can be partially offset if more missions are flying concurrently. With this point of view, it would have a great interest to get space missions with a longer lifetime. We did not add this recommendation in the paper because we do not believe it is one of these objectives.

For the case when there is no consistent long-term error, but only errors with time scales of <2 months, I would expect the uncertainty in trend to vary with (Duration)^(-3/2). No mathematical form is provided, so is this the right scaling? Clearly with extra ITRF and GIA trend uncertainty, the value will tend asymptotically to ~0.2 mm/yr. Yes, the trend uncertainty indeed evolves as alpha * (Duration_adim)^(-3/2) for uncorrelated errors (white noise), and it is also true for correlated errors if Duration >> correlation period. However the 'alpha' value is higher for correlated errors than for uncorrelated errors due to the loss of degrees of freedom (see figure below). It is however interesting, from a purely theoretical standpoint, to note that for durations <= correlation period, as can be seen for 180 days period on the figure, the error stops rising at some point. This is

explained because, as the errors are completely correlated for all values, they are affected by any error in the same direction and it therefore does not affect the trend. We do not include this mathematical formula for the asymptotic behaviour because the period that mostly concerns us is the [0,6] years as missions rarely have overlapping periods for longer than this.





For a good part of this period SARAL/AltiKa was in a drifting orbit: is that why data are gridded at 2-monthly intervals rather than monthly? More information on the choice of processing options would be appreciated.

There is maybe a misunderstanding: for all comparisons, we explain in section 3.1 that data is always gridded at 10 days intervals, not 1 or 2 months. Furthermore, the same periods are chosen for all missions, based on the Jason cycles. We actually did a lot of work that could not be included in the paper because it is a bit too methodological and tedious to explain, but we've shown that it is very important to compute grids from 2 missions on the same exact time periods, subtract those grids, and then compute GMSL. If you simply take 2 GMSL timeseries from different missions, you must interpolate on a common sampling, and you lose a lot of HF variation doing that, which affects our method by giving rise to an underestimation of the trend uncertainties. Directly subtracting 2 GMSL timeseries also does not ensure that the source data is matched spatially. So, using the method we described, we evaluated that 10 days is enough to have a global representation (Jason cycles) and keep enough of the HF variations.

Generally the paper was very clearly written, such that I only find a few minor errors worth mentioning.

Graham Quartly

Suggested corrections

I. 17 Possibly "maximum detectable" should be "minimum detectable"? Indeed. Corrected in the paper.

I. 41 Change 'data is' to 'data are'. Yes. Corrected in the paper.

I. 44 'C3S' should be expanded at first use. Yes. Corrected in the paper.

I. 88 Please expand on why data used for calibration purposes are discarded, as you do later use the calibration phase of S3B. Following the recommendations of S3A Cal/Val team, we have excluded the beginning phase of S3A due to a lot of gaps and changes in instrument modes. Concerning S3B, we did not observe as many gaps in its calibration phase as for the S3A calibration phase. Since the period was short, we decided to use the maximum data available.

I.101 Useful to also cite Frery et al., 2020. Yes, thank you for bringing it to our attention.

Fig. 1 There are common short-term variations for red and grey curves suggesting that much of the short-term variability is due to AL. Is this due to it being in a drifting orbit and thus does not provide complete global coverage on monthly timescales?

Indeed, it appears that we can observe correlated short-term variations between the red (J3/AL) and gray (S3A/AL) curves suggesting that it originates from AL. We did not investigate this short-term signal in this study because it is considered as a source of uncertainty in our error budget approach to estimate sea level drift uncertainties. In short, some of the short-term variations comes from the error between the measurements of the 2 altimeters, and some comes from the intrinsic variability of the ocean that is not observed in the same way by the different satellites. The fact that AL is on a drifting orbit could indeed contribute to a particular observation of the ocean variability but at this stage it is not possible to determine whether it comes from the drifting orbit and variation in the ocean sampling, or from a signal in one of the geophysical correction or range.

I. 172-175 Is it correct to assume that errors in SSB will be mainly sub-annual? For many regions the wave field has a strong annual signal (not just the Atlantic, but parts of the

Indian Ocean where winds will have significantly different fetches according to phase of monsoon). In the paper we look for all altimeter-induced drift, so all long term (> 1 year) errors of Range, SSB, or ionospheric correction, are precisely what we aim to detect. Therefore, we do not include it in the error budget, which only contains errors that are preventing us from measuring accurately the altimeter-induced drift.

In the future, it would be relevant to be able to better characterize the long-term errors of the SSB, in order to isolate the long-term errors coming only from the altimeter range. However to date, the characterization of the SSB errors is not mature enough to be taken into account in our error budget approach.

We have modified the text at the end of section 2 to make it clearer that we consider all altimeter-induced drift i.e. not only the range: 'However, in this study, we choose to focus on altimeter induced drift which affects the altimeter range and SSB, as well as the ionospheric correction.'

I. 217 Change 'splitted' to 'split'. Yes. Corrected in the paper.

I. 261 Should this be "Aublanc, 2020'? Yes. Corrected in the paper.

I. 262 Delete first instance of 'correction'. Yes. Corrected in the paper.

I. 278 Should be 'S3B'. Yes. Corrected in the paper.

I. 280-281 Needs revising in light of Dinardo's findings. Yes, we updated this paragraph to say that those results were not available at the time of this study and briefly say that it was an inverted sign in the implementation of the USO correction that was the cause of this issue, and reference to Dinardo et al. 2021.

I. 292 I do not understand the point being made 'The knowledge of the statistical behaviour of the error is a difficult task.' Please reword or remove. Yes, you are right, this sentence is unnecessary. Corrected in the paper.

I. 296 Suggest replace 'this time' with 'instead'. Yes. Corrected in the paper.

I. 317 In the light of Dinardo's findings on error in USO correction, please comment here on whether the differences are still significant or now understood. As explained in lines 180-181, we have updated the paper referencing the finding from Dinardo et al. 2021 on error in USO correction. But we made the choice to not update the results with the correct application of the USO correction in this paper : it might be performed in further studies.

I. 376 Change 'march' to 'March'. Yes. Corrected in the paper.

I. 377 In the light of Dinardo's findings, please revise, comment or remove this sentence. Yes, like for lines 180-181, we've referenced Dinardo et al. 2021 and briefly state that it is due to an incorrect implementation of the USO correction.

I. 381 I think 'up to' should be replaced with 'over'. Yes. Corrected in the paper.

I. 394 Need to use superscript (twice). Yes. Corrected in the paper.

References: Details are missing for Ablain, 2018; Aublanc, 2020; Jettou amd Rousseau, 2020; Meyssignac, 2019; Roinard and Michaud, 2020. Also for citing OSTST presentations (i.e. Poisson, 2019), one should give an address where they can still be accessed and the date that you last did so. Yes. Corrected in the paper.

Frery, M.-L.; Siméon, M.; Goldstein, C.; Féménias, P.; Borde, F.; Houpert, A.; Olea Garcia, A. Sentinel-3 Microwave Radiometers: Instrument Description, Calibration and Geophysical Products Performances. Remote Sens. 2020, 12, 2590. https://doi.org/10.3390/rs12162590

Quartly, G.D.; Nencioli, F.; Raynal, M.; Bonnefond, P.; Nilo Garcia, P.; Garcia-Mondéjar, A.; Flores de la Cruz, A.; Crétaux, J.-F.; Taburet, N.; Frery, M.-L.; Cancet, M.; Muir, A.; Brockley, D.; McMillan, M.; Abdalla, S.; Fleury, S.; Cadier, E.; Gao, Q.; Escorihuela, M.J.; Roca, M.; Bergé-Nguyen, M.; Laurain, O.; Bruniquel, J.; Féménias, P.; Lucas, B. The Roles of the S3MPC: Monitoring, Validation and Evolution of Sentinel-3 Altimetry Observations. Remote Sens. 2020, 12, 1763. https://doi.org/10.3390/rs12111763