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.

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.

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 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.

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.

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"?

I. 41 Change 'data is' to 'data are'.

I. 44 'C3S' should be expanded at first use.

I. 88 Please expand on why data used for calibration purposes are discarded, as you do later use the calibration phase of S3B.

I. 101 Useful to also cite Frery et al., 2020.

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?

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).

I. 217 Change 'splitted' to 'split'.

I. 261 Should this be 'Aublanc, 2020'?

I. 262 Delete first instance of 'correction'.

I. 278 Should be 'S3B'.
I. 280-281 Needs revising in light of Dinardo's findings.

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.

I. 296 Suggest replace 'this time' with 'instead'.

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.

I. 376 Change 'march' to 'March'.

I. 377 In the light of Dinardo's findings, please revise, comment or remove this sentence.

I. 381 I think 'up to' should be replaced with 'over'.

I. 394 Need to use superscript (twice).

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.