Response to Referee #2 on *"Accuracy of the Mean Sea Level Continuous record with future altimetric missions: Jason-3 vs. Sentinel-3a"* by L. Zawadzki and M. Ablain, received and published the 22nd September 2015

Thank you for your very interesting review of our paper. Please find below our answers in red with the corresponding modifications. In the modifications, I refer to the lines in the version you reviewed, not the new one.

1. Page 1513 lines 13-21. This seems to prejudge the answer to the question posed on page 1514 lines 5-7.

    Response 1:
    I see your point. In this paper, the question is more "what would be the accuracy level if we were to link S3a instead of J3 ".
    Modification 1: (Page 1514 Lines 5-7)
    Although, the Sentinel-3a ground track will be different from the historical TOPEX one (27-day repeat cycle instead of 9.91), it is relevant to know if it would be possible or not to extend the MSL time series with Sentinel-3a instead of Jason-3, <u>meeting the climate users requirements</u>.

2. Page 1515. Line 15. Please say what "errors" you are considering here (c.f. comment on page 1519 lines 16-17).
    Response 2:
    Basically all the errors coming from the measurements and all SSH corrections (see l19): propagation, electromagnetic, geophysical, atmospheric as long as they may have an effect from one cycle to another.
    Modification 2: (Page 1515 Line 19)
    Therefore, because Jason-2 and Jason-3 are, during the calibration phase, on the same ground-track and spaced less than 1 minute apart, there is a significant correlation between their measurements errors and also between their corrections errors <u>(propagation, electromagnetic, geophysical, atmospheric) at the temporal scale of a cycle.</u>

3. Page 1515. Line 17. "only a few seconds" corresponds to tens of kilometres in which distance I believe atmospheric effects might change significantly.
    Response 3:
    Yes, but we first compute SLA average for each 10-day cycle, then we compute the bias with 9 cycles (so 9 values). Thus, we make the assumption –and we think it is reasonable – that a few seconds lag has a negligible impact.
    I rephrase.
    Modification 3: (Page 1515 Line 12)

The calibration phase between Jason-2 and Jason-3 allows us to make the reasonable assumption that the impact of oceanic variability sampling on each MSL series is negligible because both altimeters measure the same ocean at almost the same time (less than 1 minute) with respect to the MSL record temporal resolution (about 10 days).

4. Page 1515. Line 21. "our analyses show the correlation . ." If this is previous analysis, please give a reference. If this is analysis in this manuscript, please don't anticipate. How is it known whether what is being correlated is error or true variability?

Response 4:

We computed the correlation between TP/J1 and J1/J2 over their calibration phases to design the noise, this is part of the method because we need this information to tune the noise. What is correlated is indeed error and true variability. But we removed annual and semi-annual signals to estimate this correlation, otherwise you would of course get a very good correlation. We estimated this correlation over a large period (~100 cycles) so it remains significant, but the relative bias is then computed over 9 cycles so the annual and semi-annual cycles must not be taken into account.

Modification 4: (Page 1515 Lines 21-22)

Hence, the correlation between Jason-2 and Jason-3 global MSL time series over their calibration phase is strong. In the cases of TP/Jason-1 and Jason-1/Jason-2 calibration phases, we performed an analysis of the MSL time series, showing their correlation was close to 0.8 after removing annual and semi-annual signals.

5. Page 1516. Line 2. "considered identical" – not quite, see comment on page 1515 line 17.

Response 5:

See Response 3.

Modification 5: (Page 1516 Line 2)

which – in this paper – are considered identical during the calibration phase

6. Page 1516. Lines 6-8. It is important that the added "noise" represents all the sources of difference between the satellites; they are separated even if on the same track.

Response 6:

See Response 3.

7. Page 1516. Lines 10-11. It may be reasonable to aim for a similar correlation as between Jason-1 and Jason-2. However, the reader does not have the information about the character of error considered to be able to judge this. And the correlation might be affected by the "few seconds" separation if this is changed from Jason-1/Jason-2.

Response 7:

See Response 2-3.

8. Page 1516. Lines 12-13. Is Zawadzki and Ablain (2014, given via a Web address) refereed. If not, some detail should be given here to enable the reader to judge the noise specification.

Response 8:

Yes, I tried to give more details. The noises were all designed by analyzing real altimetric time series: Jason-1, Jason-2, Envisat. The criteria were:

- Variance of GLORYS_GMSL + Noise had to be consistent with real time series → amplitude of the noise
- Wavelet analysis of GLORYS_GMSL + Noise had to be consistent with real time series → noise correlation period

Modification 8: (Page 1516. Lines 12-13)

This noise was designed based on the analysis of real altimetric MSL records. Firstly, the variances of Jason-1, Jason-2 and Envisat “real” GMSL time series were estimated. Secondly, a spectral analysis of Jason-1, Jason-2 and Envisat GMSL time series was performed. The noise was then tuned so that the variance and correlation period of the simulated GMSL time series are consistent with these analysis, see Tab. 1. It is worth noting that its characteristics strongly impact the results and should be designed carefully

9. Page 1516. Line 14. Table 1 shows a correlation period 30 days which is very long for any atmospheric effect.

Response 9:

It is actually not directly the atmopheric effects but the errors on atmospheric effect we want to simulate. With respect to sea level, these errors are very low. In fact, there are also other types of errors that decorrelate over a short time-period. However, the sum of these errors (averaged over a cycle) is very low in the MSL signal (<1 mm) compared to well-known 60-day correlated errors for instance (2-3 mm). Thus errors below 30 days are negligible in our spectral (or wavelet) analysis, and that is why we did not simulate them.

10. Page 1516. Lines 28-29. “over a given location” Indeed, how close does Sentinel-3a necessarily go to any given Jason-2 location?

Response 10:

The sentence is indeed misleading. I rephrase.

Modification 10: (Page 1516. Lines 28-29)

First, the impact of oceanic variability may no longer be neglected because the two satellites do not observe it with the same space-time sampling.

11. Page 1517. Line 14. “need to be identical”. This certainly isolates the impact of SSH error decorrelation. But is the impact the same as with different ground tracks (without ocean variability)?

Response 11:

The decorrelation of SSH errors is supposed to ensure this. But that is of course an approximation. The fact that the uncertainty is close to Jason-1/Envisat suggests that it is reasonable.

12. Page 1517. Line 17. “without paying attention to” – maybe “but removing”?

Response 12:

Yes. The correlation between Jason-1 and Envisat is close to 0.4. So, I actually “payed attention to” the correlation between Jason-2 and Sentinel-3a but it was useless: I did not get any cases were the correlation was not around that. This is due to the fact that the MSL series have the same basis, GLORYS. It is sufficient to ensure this level of correlation.

Modifications 12:
(Page 1517. Line 3): For instance, we performed an analysis of Jason-1 and Envisat "real" GMSL time series, showing their correlation was close to 0.4 after removing annual and semi-annual signals.
(Page 1517. Line 16): Therefore the method used for Jason-2/Jason-3 is applied with a correlation between the series monitored around 0.4.

13. Page 1519. Line 17. "uncertainty on the relative bias". This suggests that aliasing error (sparse space-time coverage) is part of "SSH error" since Table 3 has "0" under "ocean variability sampling".

Response 13:
Here again, we considered a few seconds between Jason-2 and Jason-3 were not enough to have any significant impact on the oceanic variability sampling. In a strong current like the Gulf Stream, this would not be true along the track. However, here, we average the North Atlantic MSL over a whole cycle (10 days) so we think this is a reasonable hypothesis.

Modification 13 (taking into account also comments from second referee):
(Page 1517. Line 25)

Thus we adapted the methods for the global scale to refine the analyses at the regional scale with a focus on North Atlantic Basin. This region is dominated by mesoscale variability and is thus a challenging choice because it will increase both components of the relative bias uncertainty. The intention here was to estimate an "upper bound" of the uncertainty, but results could be more mitigated in other basins (e.g. East Pacific). However, because of this variability, the correlated noise used for this specific region is also more difficult to design, see **Erreur ! Source du renvoi introuvable.**.

14. Page 1521 line 6. I think equation (1) is OK for the trend uncertainty attributable to the bias. However, overall trend uncertainty is certainly not zero for $t < t_C$ owing to the finite series length and various measurement uncertainties. Would a more complete approach weight all the data (before and after tC) with an inverse error estimate so that weighting were reduced to represent increased uncertainty due to bias? Or, I believe there is a statistical approach to estimating steps in a time series. I guess (1) is a lower bound for the overall trend uncertainty. These comments are not from specific knowledge on my part but from a reluctance to accept that more information (from the successor satellite) should degrade the trend estimate if handled appropriately.

Response 14:
You are perfectly right and I did not treat this aspect in this paper as it could be the subject of another one. This formula represents the trend uncertainty induced by the relative bias uncertainty (RBU) **if** we compute the trend with a Ordinary Least Square regression (and this is what is generally done because it does not require to have an exhaustive knowledge of all uncertainties in the MSL records). A non-explicit aspect of this paper (but perhaps it should be explicit in the outlook) is to give estimates of the RBU in 2 common scenarii. These estimates contribute to modelling error-covariance matrices of the MSL record (in this case a block

diagonal matrix) and use a Generalized LSR, see Ablain et al., 2009 at the global scale, and soon Prandi et al., 2015-6 at the regional scale.

Modification 14: (outlook)
However, new methods could be designed to lower this uncertainty. The challenge is to find a method that uses a very limited period to avoid possible drifting issues (e.g. of the radiometer) between the consecutive missions. Using external MSL series (e.g. tidal gauges, a third mission) as a reference is also a challenge because it introduces a new source of uncertainty.

The intermission relative bias uncertainties estimated in this paper also contribute to better modelling error-covariance matrices of the continuous MSL record. With an accurate model of the MSL error budget, it is possible to access a very accurate estimate of the trend and most importantly its uncertainty, using a Generalized Least Square approach, see Ablain et al. (2009).