Reviewer #1 (https://doi.org/10.5194/egusphere-2024-1802-RC1)

In this paper the authors present a novel validation method to assess the instrumental drifts in the global mean sea level trends based on the implementation of a second phase in which the two successive reference altimetry satellites fly in tandem a few years after the first. They quantify the uncertainties of the trend such method would achieve depending on the duration of the second tandem flight phase and the length of time between the two phases and compare the results with other validation methods, proving the usefulness of implementing the second tandem phase.

The manuscript represents a significant scientific contribution to achieving stable and consistent sea level measurements. However, while the methods applied in this study are most likely valid, it is hard to fully assess that because they are not presented in a clear way or sometimes at all, leaving some questions about them unanswered.

Major comments

1. It is hard to understand the methodology and follow the paper the way it is written because many parts are not explained, but only reference other papers where the method is described. It is perhaps not necessary to go into every detail as in the papers where the method is used first, but the reader should not need to read several other papers to understand what was done in this one. The paper should include a better explanation of how the covariance matrix for one tandem phase is created and how the number of independent observations is calculated, as well as what that number means for the calculation of the uncertainties. An additional problem is the way the manuscript is structured: Sect. 2 describes the method to quantify the uncertainty of the 2-tandem phase, which requires the error covariance matrix for each of the phases. It is not explained well enough how these error covariance matrices are calculated, but also the description of what they contain only comes after, in Sect. 3, which completely disrupts the flow of text.

In response to the major comments received, we have made the following improvements to our paper:

- **We have merged sections 2 and 3, presenting a sequential explanation of each step of the methodology proposed to calculate the uncertainty of the 2-tandem phase method. This reorganization aims to facilitate a clearer understanding of the methodology.**
- **In section 2.2, we have added detailed information on the calculation of error covariance matrices based on the work of Ablain et al., 2019.**

- **To enhance clarity, we have provided more detailed explanations of how the number of independent measurements is calculated and how this number is subsequently utilised to derive the uncertainty.**

2. It could be because of my misunderstanding due to lacking methodology explanations, but the sensitivity study and its conclusions do not seem convincing at all. You claim that the sensitivity tests showed that there is fairly low sensitivity of the uncertainty to the temporal correlation of errors and/or the variance because the uncertainty varies between 0.06-0.18 and 0.11-0.17, which correspond to 0.12 and 0.06 range, respectively. However, based on Fig. 2, the uncertainty varies between 0.12 and 0.18 depending on the duration of the second phase (with 2 years and 9 months between phases), so also only 0.06 difference. Same if you change the length between the two tandem phases from 3 to 6 years (double!), the uncertainty is only decreased by 0.06, from 0.13 to 0.07. To me that seems like the uncertainty is a lot more sensitive to the choice of decorrelation time than to the length of the second tandem phase or even the time between the two phases. Or am I misunderstanding something here? It would help to explain better how you chose the 1 month for the temporal correlation to make sure it is a good choice, since it seems to affect the results quite a lot.

The uncertainty we have estimated does indeed vary as a function of the length of the second tandem phase and the gap between the two tandem phases. One of the aims of this paper is therefore to determine these variations in uncertainty as a function of these two parameters, in order to be able to prescribe a relevant scenario for the second tandem phase between two successive missions. For example, the preliminary results of this study, which will be presented to OSTST in 2020, have been used to determine the minimum duration of the second tandem phase between S6 and J3 (which will take place in the first quarter of 2025).

However, the results obtained are sensitive to the uncertainty budget of the SSHA differences during a tandem phase, which we have estimated in this study (Section 3). In order to accurately estimate this uncertainty budget, we have analysed three different tandem phases, from which we have derived the standard deviation of the differences and the temporal correlation of the SSHA differences during a tandem phase (see Section 3). Each tandem phase gives close but not strictly identical results, which can be explained by the differences between the SSHAs of the different altimetry missions compared to each other, and by the relatively short duration of the tandem phases (about 6 to 9 months). Therefore, we thought it would be useful to analyse the sensitivity of our analyses by varying the uncertainty budget by taking the extreme values possible for the uncertainty budget (for example, by varying the temporal correlation from 15 days to 2 months). The aim of this sensitivity analysis is to demonstrate the robustness of the results of this study.

In the updated paper we have therefore improved the value of this sensitivity analysis in relation to the uncertainty budget. The fact that the methodological sections have been improved should also clarify this sensitivity analysis.

Other comments

L61-65: This is a very long and quite hard to read sentence.

Corrected : The sentence has been split.

L86: This seems like a wrong unit or a very wrong number.

Corrected

L106: dot missing

Corrected

L? (p5 bottom): "contain the along-track sea level anomaly at 1Hz (SLA, see Eq. (1))"

What does Eq. (1), which is for the estimator of beta, have to do with SLA? Seems like it is referring to the wrong equation.

Corrected

L? (p6 top): "The along-track SLA provided in the L2P products is derived from the following equation:

SLA = Orbit − Range − Σ_i Correction_i − MeanSeaSurface"

What is the meaning of each of the variables in this equation?

Corrected : The meaning of each variable in the equation has been defined

L137-145 It is not clear to me whether you here describe how the dataset you downloaded and used was created or the processing steps you applied to the dataset before using it. It is especially unclear what was done here considering the "regional scales" you refer to throughout the rest of the manuscript are using different longitude-latitude box sizes that the 1 degree latitude per 3 degrees longitude mentioned here. Additional minor comment: You here use degree (word), and in other places in the manuscript degree symbol when describing the size of the box. Please be consistent.

L2P products contain the SLA along the track at 1Hz, already homogenised in terms of geophysical and atmospheric correction. On our side, the work

consisted in calculating the GMSL differences between two altimeter missions during a tandem phase from the 1Hz SLA measurements along the track available in L2P. We applied the same method as in Henry et al. (2014), who recommend applying 1° latitude per 3° longitude to the calculated SLA grids to optimise the effect of sea level variability observed by reference altimeter missions in the GMSL.

In the revised paper, the applied method is better described. We have also used the degree symbol (°) rather than the word "degree".

L147-152 This is a general statement, not specifically for the global scale, so it should be in the previous subsection.

Corrected, the sentence has moved in new section 2.1

L151-152 Why exactly are you referring to the previous section here? Seems completely unnecessary and it disrupts the reading.

Corrected

L156-157 1. Is the dashed line the periodic signal or the differences before removing it? The way this sentence is phrased, it could also be understood as the dashed line being the difference with the periodic signal removed, which does not seem the case when looking at the figure. The explanation of what the grey dashed line is should also be in the figure caption; 2. How much did the removal of the signal reduce the standard deviation (or what was the standard deviation before removing this signal)?

Corrected: the dashed line is the difference before removing the periodic signal. Removing the signal reduces the standard deviation from 0.99 to 0.48 mm. It has been clarified in the text and added on the figure caption.

L164-167 This could be because there was no proper explanation of what is the meaning of n and its calculation, but I cannot follow this conclusion. Could you please explain the reasoning behind it?

The temporal signal of GMSL difference between Jason-3 and Sentinel6-MF consists of 20 measurements taken every 10 days. We used the first-lag autocorrelation coefficient, ρ_1, to estimate the number of independent measurements using the formula from Guerou et al. (2023):

$$n = \frac{(1-\rho_1)}{(1+\rho_1)} \times n_{\text{sample}}$$

where ρ_1 is the first-lag autocorrelation coefficient, which represents the correlation between consecutive measurements, and n_{sample} the total number of measurements of the sample. This analysis shows that 18 out of the 20 measurements are independent, implying that the signal decorrelates quickly after a few measurements. Based on the low autocorrelation, we can conclude that after approximately 3 measurements (around 30 days, as we have one measurement per cycle), the signal becomes largely decorrelated. In other words, measurements separated by more than one month have little to no correlation, meaning they can be treated as independent for analysis purposes.

We have better explain this method in the updated paper.

L180 In this section you seem to focus on the Jason-3 and S6-MF tandem phases, but you do not clearly state that, you only mention that you use the duration of the first tandem phase from those missions and refer to Tab. 1 for the uncertainty budget, which contains the uncertainties for all 3 pairs of satellites. Could you please make this clearer.

Corrected

L193 10.16 seems to be the wrong number, probably a typo.

Corrected

L194-195 This conclusion comes out of nowhere because the specific Jason-3/S6-MF 2-tandem phase scenario has only been mentioned once before in the manuscript, at the end of the introduction, where the duration of the second phase is not mentioned at all, just when the second phase would be. You need to elaborate this scenario better before discussing results and conclusions about it, not after. It would also be good to know how and why was this particular scenario chosen.

There is a misunderstanding here underlying the fact this section have to be improved. The conclusion "Following this analysis, a second tandem phase of 4 months is deemed sufficient to verify the instrumental stability on the global scale" is not related to the second tandem phase between S6 and Jason-3, but is derived from Fig. 3, which shows the evolution of the uncertainty of the trend of the GMSL differences as a function of the time period between the two tandem phases (between 1 and 6 years), for 4 different time spans of the second tandem phase (from 1 month to 6 months). The sentence has been reformulated to better understand the recommendation coming from the analysis presented in this paper.

The recommendations on the duration of the second tandem phase between S6 and Jason-3 are coming from preliminary results of the same study

presented in conference (OSTST 2020) but not not yet published in a peer review paper.

Note that the duration of the second tandem phase between Jason-3 and S6-MF (4 months) has been added at the end of the introduction as recommended.

L199-206 You might want to put the sensitivity tests into a separate paragraph to improve readability of the manuscript.

Done

L212 I do not understand what does “corresponding to regional scales” mean in this context.

Corrected: it is not relevant in this sentence, it has been removed.

L221 Nothing is actually marked as a results section and there are two sections that describe the results, the previous one (Sect. 4) and this one (Sect. 5).

Corrected: it has been replaced by “In this section”

L228 These are not the same satellites as in Sect. 4. Are you comparing the results for the without-tandem method with the results from Sect. 4. for the Jason-3 and S6-MF or are you re-calculating the 2-tandem uncertainties for other satellites with different short-term time-correlated errors?

Corrected. The uncertainty budget presented in Table B1 (from Jugier et al., 2022) is applicable to different altimetry missions.

L232 & L240 You refer to Tab. B1 and Tab. C1 before Tab. A1.

Corrected

L285 methods, there are two of them

Corrected

L301-302 "since" is written twice

Corrected

Fig. 2. What is the unit of lag? When discussing the time with no autocorrelation, you use months, so could you explicitly relate that to whatever is shown here. Also, please note clearly what is the grey line in (f) here, not just in the main text.

Corrected: the description of the grey line has been added to the figure description. The autocorrelation of the data is related to the time interval

between measurements, with one measurement taken per cycle (approximately every 10 days). It has been added to the x-label of autocorrelation plots.

Fig. 3. Should be space agency, not spatial.

Corrected

Fig. 4. Since in the text you sometimes refer to box sizes in kilometers, could you please add that in the figure, it would make it easier to follow.

Done

Tab. 2. Could you please also provide the values for other box sizes used in the study?

We provide here the value for other box sizes for information,but not added in the paper :

- **Uσ = 4.1 mm for box size 3°*3°**
- **Uσ = 2.7 mm for box size 6°*6°**
- **Uσ = 2.3 mm for box size 9°*9°**
- **Uσ = 2.1 mm for box size 12°*12°**
- **Uσ = 1.9 mm for box size 18°*18°**
- **Uσ = 1.6 mm for box size 36°*36°**

There is an error with the format of citations throughout the paper, most of them have too many brackets.
Done